Rising like a Phoenix: Emerging from the Pandemic and Reshaping Human Endeavors with Digital Technologies ICIS 2023

AI in Business and Society

Dec 11th, 12:00 AM

AI Washing: The Framing Effect of Labels on Algorithmic Advice Utilization

Dirk Leffrang
Paderborn University, dirk.leffrang@uni-paderborn.de

Oliver Mueller
Paderborn University, oliver.mueller@uni-paderborn.de

AI Washing: The Framing Effect of Labels on Algorithmic Advice Utilization

Completed Research Paper

Dirk Leffrang
Paderborn University
dirk.leffrang@uni-paderborn.de

Oliver Müller
Paderborn University
oliver.mueller@uni-paderborn.de

Abstract

Many researchers and practitioners see artificial intelligence as a game changer compared to classical statistical models. However, some software providers engage in "AI washing", relabeling solutions that use simple statistical models as AI systems. By contrast, research on algorithm aversion unsystematically varied the labels for advisors and treated labels such as "artificial intelligence" and "statistical model" synonymously. This study investigates the effect of individual labels on users' actual advice utilization behavior. Through two incentivized online within-subjects experiments on regression tasks, we find that labeling human advisors with labels that suggest higher expertise leads to an increase in advice-taking, even though the content of the advice remains the same. In contrast, our results do not suggest such an expert effect for advice-taking from algorithms, despite differences in self-reported perception. These findings challenge the effectiveness of framing intelligent systems as AI-based systems and have important implications for both research and practice.

Keywords: Artificial intelligence, algorithm appreciation, framing, advice-taking, expertise

Introduction

The term Artificial intelligence (AI) has gained immense popularity in recent years, becoming a buzzword used by companies, newspapers, and researchers to describe a wide range of more or less intelligent systems and algorithms. A buzzword is "a catchword or expression currently fashionable; a term used more to impress than to inform" (Oxford University Press, 2022). For example, companies use buzzwords in the context of greenwashing, which refers to the "creation or propagation of an unfounded or misleading environmentalist image" (Oxford University Press, 2023b). Analogous to the concept of greenwashing, some companies engage in "AI washing" by exaggerating the AI content of their products to gain business advantages (Moore, 2017). For instance, some providers might label a supervised machine learning classifier based on logistic regression as AI, while the label "statistical model" would arguably be more adequate.

Similar to the question of the proper terminology for algorithmic advisors, the question whether people trust algorithmic advice at all has attracted a lot of research interest in recent years. In a series of seminal studies, Dietvorst and colleagues initiated a stream of research and coined the term algorithm aversion (e.g., Dietvorst et al., 2015; Prahl & Swol, 2017; Saragih & Morrison, 2022). Algorithm aversion refers to the observation that although algorithms are typically more accurate in estimating numerical short-term regression tasks than humans (see Grove et al., 2000), people still tend to prefer advice from humans compared to algorithms (Jussupow et al., 2020). This phenomenon has been observed in multiple manifestations and contexts (Mahmud et al., 2022). Even if people are not averse to algorithms they struggle to assess the capabilities of AI correctly (Fügener et al., 2022). However, there are also reports of an opposite phenomenon, namely algorithm appreciation (Logg et al., 2019). Other studies found no significant differences or increasing algorithm appreciation over time (Leffrang et al., 2023).

The inconclusive results from previous studies on algorithm aversion or appreciation suggest that various factors influence people's trust in algorithmic or human advisors. In their literature review, Jussupow et al. (2020) emphasized the expertise of a human advisor as an important factor in algorithm appreciation. Furthermore, they noted that technical labels used to describe algorithmic advisors have varied across experiments. For example, labels ranged from "sophisticated statistical model" (Dietvorst et al., 2015), over "decision support system" (Heßler et al., 2022), to "artificial intelligence" (Longoni et al., 2019). Yet, studies systematically examining the influence of these different labels on users' trust in advice are scarce. The only empirical study investigating this issue is a study by Langer et al. (2022). In a series of surveys, they found that the labels used to describe algorithms can significantly impact people's perceptions of the advisors, particularly in terms of perceived fairness and trust.

However, self-reported perceptions and opinions may not always align with actual behavior. For instance, individuals may express a willingness to purchase green products but exhibit different purchasing behavior (Moser, 2015). Therefore, in this paper, we seek to contribute to the existing literature by experimentally examining the impact of labels with varying positive or negative connotations of expertise (referred to as the framing effect) on actual user behavior. Expertise is a "[...] skill or expertness in a particular branch of study" (Oxford University Press, 2023a). While algorithmic advisors arguably do not develop expertise on their own, computational advancements allow experts to incorporate more complex forms of expertise into algorithmic advisors. Drawing on theoretical foundations and literature on framing (Tversky & Kahneman, 1981), we address the following research question:
What is the influence of expert framing on algorithmic advice utilization?

We tackled this research question with two online within-subjects experiments, in which people received advice on a car price regression task. In these experiments, we examined how individuals adjusted their initial car price estimation when receiving help from advisors described using different framings with varying positive or negative connotations. Specifically, we used four conditions: Two of these framings referred to human advisors, a negative one signaling low expertise (a student) and a positive one signaling high expertise (an industry expert). Likewise, two framings referred to algorithmic advisors, a negative one signaling low expertise (a statistical model) and a positive one signaling high expertise (an artificial intelligence). Notably, the advice provided by the advisors remained constant across all conditions, with only the labels used to describe the advisors varying. Through our analysis of the experimental results, we discovered evidence for algorithm appreciation (i.e., individuals utilized advice from algorithms more than from humans) and observed a framing effect for human advisors, but not for algorithmic advisors (i.e., differences in advice utilization between experts and students, but no differences between statistical models and AI).

This paper makes valuable contributions to the information systems domain by providing empirical insights into the behavioral effects of framing AI-based systems in business and society. Notably, this paper extends prior survey research by employing incentivized experiments to investigate the effects of terminological framing (high vs. low expertise) on both human and algorithmic advisors in a regression task. Based on our findings we suggest that product owners should focus on the functionality and user experience of their algorithms, rather than using buzzwords like AI.

Theoretical Background & Hypotheses Development

Framing in Decision-Making

Framing refers to the presentation of the same factual content in different formats and has been shown to have a significant impact on decision-making (Tversky & Kahneman, 1981). The pioneering work of Tversky & Kahneman (1981) demonstrated that people's risk preferences could be influenced by the framing of decision options. Specifically, people were more risk-seeking when a frame focused on possible negative consequences of a decision (e.g., 10 out of 100 lives will be lost). If they observed positive consequences (e.g., 90 out of 100 lives will be saved), people were less willing to take risks. This observation has become commonly known as the framing effect. Framing effects are not limited to risk-taking but can take several forms, including attribute framing (Levin et al., 1998).

Attribute framing involves presenting the same content using different formulations for an attribute, which can result in different decision-making behaviors (Levin et al., 1998). For instance, people preferred the label "75% lean" beef over "25% fat" beef (Levin et al., 1998). The positive or negative association of an attribute can semantically and syntactically influence the cognitive representation of an entire statement, leading to different interpretations based on changing a single word or number (Levin et al., 1998).

Hypotheses Development

Research on attribute framing suggests that proper framing of an advisor can convey certain characteristics that influence the acceptance of advice. Existing literature suggests that people are generally more willing to seek advice from experts than from non-experts (Bonaccio & Dalal, 2006). This preference for expert advice is not necessarily a cognitive bias, but rather a finely-tuned intuition (Gigerenzer, 2018). Seeking advice from experts appears to be a reasonable approach, as users typically lack information about the quality of an advisor. They have to use signals like "doctor" to assess the quality of an advisor (see Spence, 1973).

Furthermore, Goldsmith & Fitch (1997) interviewed participants who considered advice from designated experts to be more helpful compared to advice from other sources such as roommates or family members. The media also often portrays experts as experienced and knowledgeable (Önkal et al., 2009) and people tend to use the advice of experienced and knowledgeable advisors more than the advice of other advisors (Harvey & Fischer, 1997). While this may seem reasonable, it is important to note that expertise itself does not necessarily guarantee better outcomes. For example, the factor experience did not improve participants' performance in predicting human behavior across different domains (Highhouse, 2008).

In summary, prior research found expert effects in advice-taking from humans. Our first objective is to extrapolate the concepts of expert effects and framing to the specific scenario, where the label of a human advisor is the only variable being manipulated:

Hypothesis 1: Participants are more likely to take the advice of a human advisor if the advisor's framing suggests more expertise.

Previous research examined the influence of textual descriptions on the perception of algorithms. For instance, Madhavan & Wiegmann (2007) used a public safety task to investigate how seven-sentence-long textual descriptions of expertise influenced algorithm aversion. The textual descriptions included unique characteristics for each advisor, such as their profession (e.g., "student", "expert in airport security", "computer system"), training, competence (e.g., "Novice" vs. "Expert"), and experience in the application area. Surprisingly, participants showed algorithm appreciation (i.e., preferred algorithms over humans) for non-expert advisors and algorithm aversion (i.e., preferred humans over algorithms) for expert advisors.

Similarly, Pearson & Mayhorn (2017) conducted a study on a routing problem with human and algorithmic advisors providing contradicting advice. They used advisor descriptions similar to Madhavan & Wiegmann (2007). They found that higher levels of expertise led to more trust, but they did not observe an interaction between advisor type (i.e., human vs. algorithm) and expertise.

In another study, Hou & Jung (2021) manipulated the textual descriptions of algorithmic and human advisors for creative and analytical tasks. When the human was described as a non-expert (crowdworker or person) and the algorithm as an expert (AI), participants exhibited algorithm appreciation. However, when the human was described as an expert and the algorithm as a non-expert (algorithm or computer), participants showed algorithm aversion. The textual descriptions in this study consisted of one sentence and assigned unique individual characteristics to each advisor, including terminological differences, number of advisors (e.g., "A group of experts" vs. "An algorithm"), and information about the prior experience (e.g., "A group of experts [...] with their 20 years of experience" vs. "An algorithm [...] aggregating several MTurkers' responses") to differentiate the conditions. Participants' perceptions and preferences towards algorithms changed based on the framing of the descriptions.

While most of the previous studies utilized detailed textual descriptions of varying lengths and information, our study focuses on the effects of single words or labels. Providing different information, even with different levels of detail, can lead to unintended framing effects as readers draw conclusions from it (Gigerenzer, 2018). Yet, according to attribute framing, even a single property of an advisor can change the behavioral consequences of an entire sentence.

Previous research already observed that terminological differences can lead to varying perceptions of algorithms. For example, Shank et al. (2019) investigated the effect of technological terms on sentiments. They found that participants preferred technologies they were more familiar with. Additionally, as mentioned in the introduction, Langer et al. (2022) found that terminology influenced evaluations of trust and fairness, but not perceived abilities of algorithms when compared to humans. However, as they rightfully pointed out, focusing on perceptions allows only for tentative conclusions about behavior, and more research on the actual behavioral effects of terminology is needed.

In summary, prior research has found varying evaluations of algorithms when using descriptions of different lengths and with different amounts of information about the algorithms. Other strands of research found that individual labels are sufficient to influence the perceptions of algorithms. However, changes in perception do not necessarily lead to changes in behavior. To the best of our knowledge, the behavioral effects of terminology on algorithmic advice utilization have not been studied explicitly in an experiment. Accordingly, our second hypothesis in this paper is:

Hypothesis 2: Participants are more likely to take the advice of an algorithmic advisor if the advisor's framing suggests more expertise.

Research Design

In order to investigate the effect of framing on algorithm aversion, we conducted two experiments. Study 1 focused on participants who had a particular interest in algorithms. In Study 2, we aimed to increase the generalizability of our results by including a more diverse population. We pre-registered both experiments before we ran them[1].

Study 1

Conditions

We utilized a 2 x 2 repeated measures within-subject design, employing four conditions that varied the advisor type (human vs. algorithm) and the framing of the advisor's expertise (high vs. low). For each task, we assigned participants to one of the four possible combinations of advisor type and expertise framing. Notably, the only difference between the conditions was the label used to denote the advisor. The content of the advice - i.e., the numerical estimation - remained constant.

To determine appropriate labels for the algorithmic advisor type, we conducted a review of existing literature. We drew on prior reviews on algorithm aversion to select appropriate labels for the algorithmic advisor type (Langer et al., 2022; Langer & Landers, 2019). Table 1 displays the identified labels.

Label	Exemplary Study
Algorithm	(Lee, 2018)
Automated system	(Keel et al., 2018)
Artificial intelligence	(Leyer & Schneider, 2019)
Computer	(Langer et al., 2020)
Computer program	(Grgic-Hlaca et al., 2019)
Decision support system	(Shibl et al., 2013)
Machine learning	(Gonzalez et al., 2019)
Technical system	(Montague et al., 2009)
Robot	(Ötting & Maier, 2018)
Sophisticated statistical model	(Dietvorst et al., 2015)

Table 1. Terminological Differences for Algorithms (see Langer et al., 2022)

[1] We pre-registered the research question, our hypotheses, the dependent variable, our analysis, a plan to exclude data, and the factors determining the sample size. Please note that we reworded the research question to provide a better reading flow, but did not change its meaning. First experiment: https://aspredicted.org/55H_D8X Second experiment: https://aspredicted.org/BXN_9FQ

Among these candidates, ”(sophisticated) statistical model” has probably been in use for the longest period of time, even predating the computer era. ”Artificial intelligence”, in contrast, has gained prominence as a buzzword in recent years and has received significant media attention. AI is different from previous generations of approaches as it is autonomous, learning, and often inscrutable (Berente et al., 2021). The differences between the two labels have also been identified in the study of Langer et al. (2022), who semantically clustered different labels for algorithmic advisors and found the highest distance between the labels artificial intelligence and sophisticated statistical model. Therefore, we selected these two framings as conditions for our experiments.[2]

Label	Expertise	Exemplary Study
Industry expert	Expert	(Berger et al., 2021)
Colleague	Non-expert	(Leyer & Schneider, 2019)
Human	Non-expert	(Kramer et al., 2018)
Student	Non-expert	(Madhavan & Wiegmann, 2007)
You	Non-expert	(Dietvorst et al., 2015)

Table 2. Terminological Differences for Humans (see Jussupow et al., 2020)

For the framing of the human condition, we referred to another literature review on algorithm aversion by Jussupow et al. (2020). Table 2 presents the labels used in prior studies, their expertise according to the authors, and their respective origins. Based on the findings from the review, we chose the framing ”industry expert” to communicate high human expertise and ”student” to communicate low human expertise, which aligns with prior studies such as Madhavan & Wiegmann (2007).

Sample

We offered an online experiment to 127 first-year bachelor’s students studying Information Systems at a medium-sized, IT-focused university. Students could participate in the online experiment from July 1 to July 11, 2022. They received a performance-based incentive in the form of exam bonus points for their participation, with no penalty for non-participation or incomplete responses. Out of the invited students, 60 participated in the experiment.

We excluded 13 participants because their data was incomplete, and 7 participants failed the attention checks or made constant estimations (e.g., always entering the same number). We also controlled for participants who finished too quickly (i.e., less than 90 seconds for all tasks) or too slowly (i.e., more than two standard deviations above the mean participation time) to prevent ”clicking-through” the experiment, but no participants were excluded based on these criteria. In total, our final sample consisted of 40 participants.

Because participants’ interest in algorithms might influence our results, we included participants’ affinity for technology (ATI) as a control variable, consistent with prior research on terminology (Franke et al., 2019; Langer et al., 2022). The ATI scale is a 9-item scale ranging from 1 to 6, with higher values coded as higher affinity for technology. The mean value on the ATI scale was 4.49 (SD = 0.58). Notably, our mean ATI value was more than one point higher compared to the survey of Langer et al. (2022), indicating a more technology-savvy population in our study.

Procedure & Materials

In order to investigate the effect of framing on advice-taking, we conducted an incentivized online experiment. Incentivization is an effective setup for estimation tasks, which respond to effort (Camerer & Hogarth, 1999). Upon accepting the invitation to the experiment, sent in the form of a link in the learning platform of the university, they were directed to the instructions page in the online survey tool (LimeSurvey). In the instructions, we informed the participants that they would be estimating the price of a car. If their estimation accuracy was good enough, we compensated them with bonus points for the final exam. The instructions informed participants of attention checks to ensure careful reading.

[2]To avoid potential unintended bias associated with the adjective ”sophisticated”, we omitted it. Reproducing the procedure of (Langer et al., 2022), the clusters still joined last.

A paradigm commonly used to investigate such forms of advice-taking differences is the judge-advisor system (JAS) (Bonaccio & Dalal, 2006). In a JAS, the advice-taking process involves two roles: a judge and an advisor. The judge starts with an initial estimation but can receive advice from an advisor. The judge then weighs the advice in light of their initial estimation to arrive at a final decision.

During the experiment, participants took the role of the judge and received help from an advisor, which - according to our 2x2 experimental design - could be an artificial intelligence, an industry expert, a statistical model, or a student. At the beginning of the experiment, we placed an attention check. Afterward, participants provided self-reported age and gender. Next, participants proceeded to complete eight regression tasks, each consisting of the following steps:

1. Participants saw data on a used car (e.g., brand, model, age, mileage). Based on this information, they had to estimate the price of the car.
2. After submitting their initial estimation, they saw a price estimation from an advisor (AI, statistical model, industry expert, or student). Now they had the opportunity to adjust their initial estimation.
3. After submitting their final estimation, they learned about the actual price of the car, the mean absolute error (MAE) of the advisor's estimation, and their own MAE for the final estimate.

Figure 1 shows an example of the user interface from Step 2 for the AI condition. A table showed mileage, brand, model, fuel, gear, the type of offer, horsepower, and the vehicle registration year of the car. Note that the only difference between the four conditions was the label of the advisor, while the advice itself remained the same. To avoid anchoring bias (Tversky & Kahneman, 1974), we did not specify a default value for the user's estimation. Figure 2 shows an example of the feedback provided in Step 3.

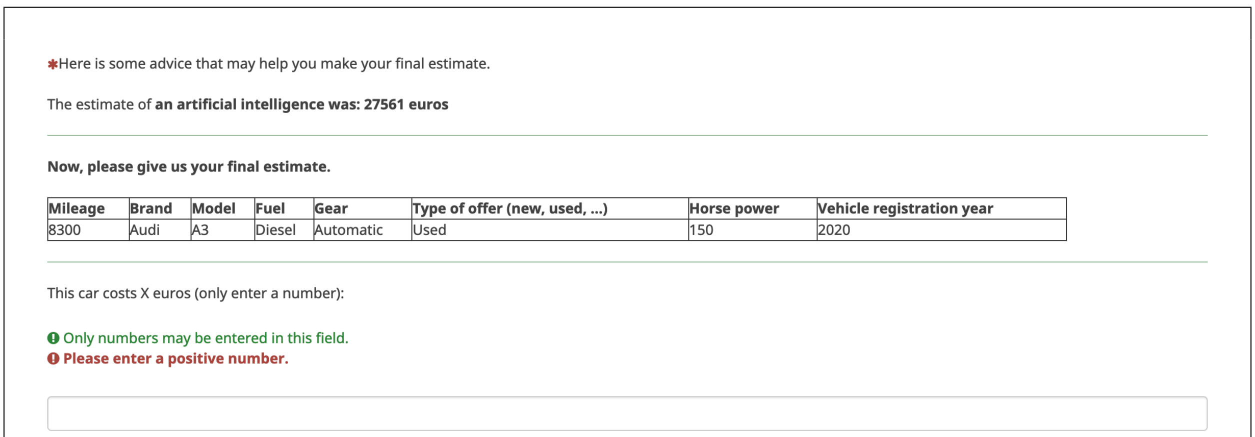

Figure 1. User Interface of Step 2 for the "Artificial Intelligence" Condition

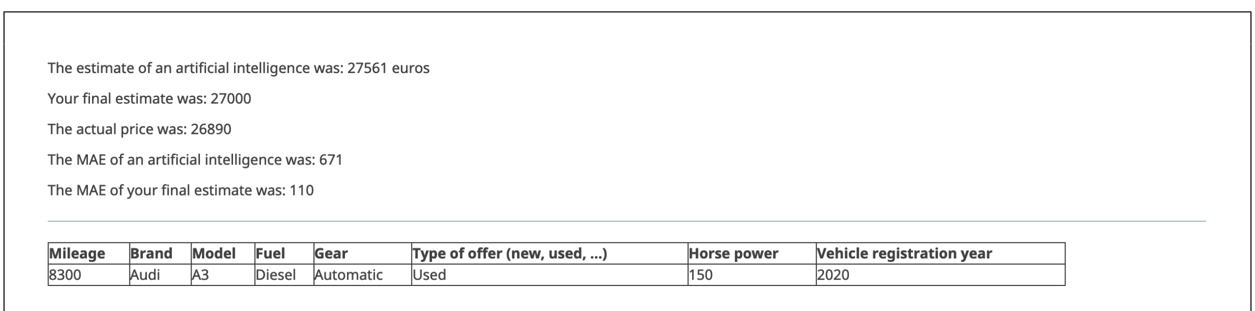

Figure 2. User Interface of Step 3 for the "Artificial Intelligence" Condition

To familiarize participants with the advisors and the user interface, participants received feedback on the actual price of the cars for the first four example cars (warm-up). For the next four cars, participants only conducted the first two steps and did not receive feedback in Step 3. These vehicles formed the basis of our analysis. Participants saw each condition (AI, statistical model, industry expert, and student) once in a randomized order in both the warm-up and actual regression tasks.

We chose a car regression task to create a situation that participants could realistically encounter. Additionally, both human and algorithmic advisors can provide reasonable advice in such tasks. The data provided was based on real-world data from an online marketplace for new and used cars sold in Europe. To ensure comparability of the cars, we randomly selected used cars that were sold between 2020 and 2021, had a mileage of at least 1000 km, and a price between 10,000€ and 50,000€. To control for varying the quality of the advice, we generated the advice artificially with a random mean percentage error of up to 5%.

As framing is not the only factor influencing algorithm aversion, we did not provide feedback on the general accuracy of all advisors and kept the order of the tasks constant. To avoid ordering and learning effects, we assigned a randomized condition to each participant-task combination in which we presented the same advice. Additionally, we included a second attention check to mitigate fatigue biases and controlled for multiple additional factors, which will be described in the model specification section.

Finally, we assessed the robustness of our assumptions by asking participants to rate the expertise of each of the four advisors separately, using an 11-point Likert scale (0 = least expertise, 10 = highest expertise). This additional measure allowed us to supplement our observations of participants' actual behavior by perceived differences in expertise between the different advisor conditions.

Measures

The dependent variable was the weight participants assigned to the advisor's recommendation. Weight of advice (WOA) captures the degree to which participants adjust their final estimation towards the advisor's recommendation and is commonly used in the JAS paradigm (Bonaccio & Dalal, 2006; Harvey & Fischer, 1997). We used the following formula to calculate WOA:

$$WOA = \frac{|final\ estimation - initial\ estimation|}{|advisor's\ estimation - initial\ estimation|} \tag{1}$$

A value of 0 corresponded to no change in the initial estimation after receiving advice, which indicated no advice-taking. An exception to this was when the original estimation was equivalent to the advice, in which case WOA was not defined. A WOA value of 1 meant that the judge fully adjusted his or her original estimation to the value of the advisor. Consistent with prior studies (e.g., Gino & Moore, 2007; Logg et al., 2019), we winsorized any WOA values that were greater than 1 to mitigate outliers due to overshooting.

Model Specification

To test our hypotheses, we employed a linear mixed-effects regression approach. The unit of analysis was the combination of participant and task. The four main variables of interest were Student, Industry Expert, Statistical Model and Artificial Intelligence. These took the value 1 if participants saw the respective label.

We also controlled for between-subject factors from prior research on algorithm aversion. As mentioned in the sample section, we included the ATI scale. Additionally, we assessed participants' domain knowledge about cars using an 11-point Likert scale (0 = no knowledge, 10 = very good knowledge), since higher domain knowledge has been found to increase algorithm aversion (Logg et al., 2019). Age and a dummy variable for Female gender served as additional control variables. Given the repeated measures within-subject design of our experiment (i.e., every participant solves multiple tasks), we included random intercepts for Participant and Task in the model to account for potential between- and within-subject variability.

Our hypotheses aimed at pairwise comparisons, so we used Student as reference level to test our first hypothesis. Statistical Model served as reference level to test our second hypothesis. In this case, we replaced the variable "Student" with "Statistical Model" to avoid perfect multicollinearity.

Study 2

To examine the generalizability of our findings to a broader population, we conducted a second study on car price regression tasks. This was motivated by the findings of Lee & Rich (2021), who reported differences in algorithm aversion across social groups. They found that prior experiences with algorithms influenced the evaluation of algorithms. Additionally, Kramer et al. (2018) found that self-reported prior experience with algorithms correlated with participants' decision-making regarding human versus algorithmic decision-makers. This suggests that specific social groups may differ in their propensity to take advice from algorithms. Except for the changes outlined below, we followed the same procedure as in Study 1.

Sample

We offered the experiment on Prolific until 80 participants from the US, balanced on gender, completed the experiment successfully. Successful completion was defined as passing all attention checks. Participants could participate between November 21 and November 29, 2022. The data collection period took place between November 21 and November 29, 2022. Participants received a base payment of 3£ for their successful participation, and a performance incentive of up to 2£ depending on their relative performance compared to other participants.

A total of 115 Prolific workers participated in the study, but 35 were excluded due to failing attention checks. The remaining 80 participants had a mean value of 3.65 (SD = 0.98) on the ATI scale, indicating a less technology-savvy population compared to Study 1 (mean = 4.49, SD = 0.58).

Results

Model-Free Evidence

Figure 3 displays box plots presenting model-free evidence from Study 1 and 2. In Study 1, the median values for the expert conditions (artificial intelligence: 0.90 and industry expert: 0.68) were associated with higher WOA levels than the non-expert conditions (statistical model: 0.80 and student: 0.52). Additionally, both algorithmic frames (artificial intelligence and statistical model) had higher median values than the human conditions (expert and student). Overall, the algorithmic expert condition had the highest median.

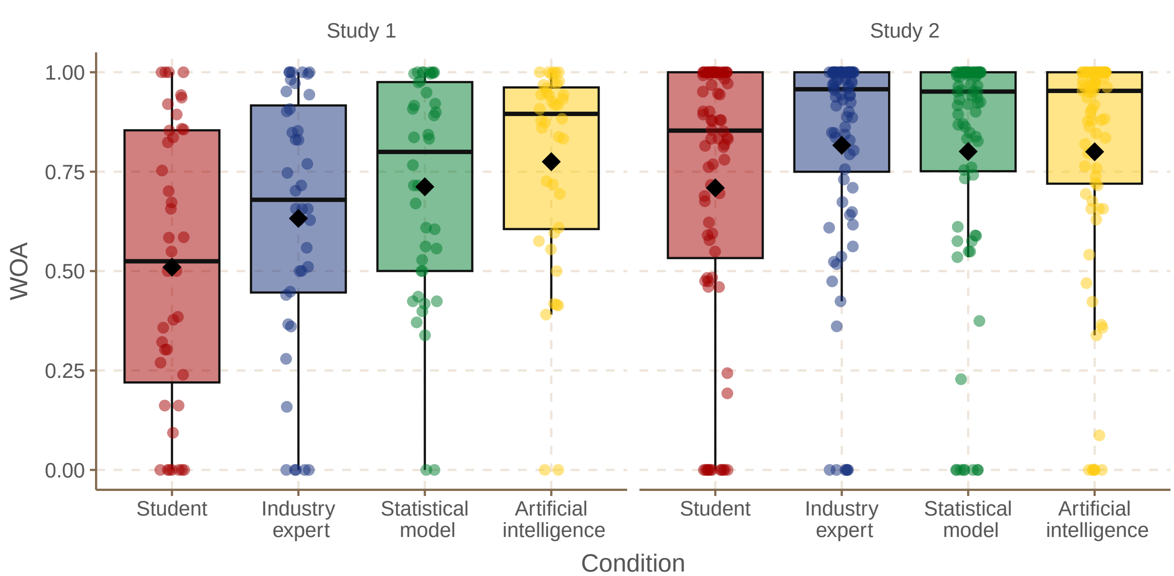

Figure 3. Boxplots of WOA in Study 1 and 2 With Means as Black Squares

In Study 2, the model-free evidence revealed similar median values for artificial intelligence (0.95), statistical model (0.95), and industry expert (0.96), along with similar distributions. However, the student condition had a lower median WOA (0.85). Notably, WOA occurred more frequently with a value of 1 compared to Study 1, suggesting a higher likelihood of participants relying entirely on the advice.

Estimation Results

Table 3 presents the results of the regression models specified above. The first two models present the results from Study 1 (S1_Student & S1_Stat Model) while the last two models represent the results from Study 2 (S2_Student & S2_Stat Model).

	Dependent Variable:			
	WOA			
	Study 1		Study 2	
	(S1_Student)	(S1_Stat Model)	(S2_Student)	(S2_Stat Model)
Constant	0.706* (0.394)	0.912** (0.394)	0.781*** (0.149)	0.876*** (0.149)
Student		−0.206*** (0.061)		−0.095** (0.038)
Industry Expert	$\boxed{0.125^{**}\ (0.060)}$	−0.080 (0.061)	$\boxed{0.109^{***}\ (0.038)}$	0.014 (0.038)
Statistical Model	0.206*** (0.061)		0.095** (0.038)	
Artificial Intelligence	0.266*** (0.061)	$\boxed{0.060\ (0.061)}$	0.099** (0.038)	$\boxed{0.004\ (0.038)}$
ATI	−0.058 (0.055)	−0.058 (0.055)	0.001 (0.030)	0.001 (0.030)
Domain Knowledge	−0.019 (0.018)	−0.019 (0.018)	−0.025* (0.014)	−0.025* (0.014)
Age	0.008 (0.012)	0.008 (0.012)	0.002 (0.002)	0.002 (0.002)
Female	−0.039 (0.129)	−0.039 (0.129)	−0.041 (0.057)	−0.041 (0.057)
Participant?	✓	✓	✓	✓
Round?	✓	✓	✓	✓
Observations	160	160	320	320
Log Likelihood	-48.702	-48.702	-73.221	-73.221
Akaike Inf. Crit.	117.405	117.405	166.442	166.442
Bayesian Inf. Crit.	148.157	148.157	204.125	204.125
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3. Regression Results

Study 1

We investigated whether participants were more likely to take the advice of a human (H1) or an algorithmic (H2) advisor based on the advisor's framing. In our first model (S1_Student) in Table 3, we regressed WOA on our four conditions, using the student condition (i.e., low-expertise human advisor) as the reference group to test our first hypothesis. To test our second hypothesis, we constructed a similar model (S1_Stat Model) using the statistical model condition (i.e., low expertise algorithmic advisor) as the reference group. We statistically controlled for ATI, domain knowledge, age and gender, and included random intercepts for participant and task.

In model (S1_Student), Industry Expert (i.e., high expertise human advisor) had a 0.125 higher WOA on average compared to Student. This association was significant ($p < 0.05$). A 12.5%-point increase in WOA due to high expertise for the human advisor appeared to be practically meaningful. Thus, we found support for our first hypothesis that expertise increases advice-taking for humans.

In model (S1_Stat Model), the AI condition was associated with a 0.06 higher WOA on average compared to the statistical model condition. However, due to the lack of statistical significance ($p > 0.1$), we could not support our second hypothesis that expertise increases advice-taking for algorithms.

In both models, statistical models had a 0.206 higher WOA compared to students ($p < 0.01$). The coefficient of model (S1_Student) indicated a 0.266 higher WOA for the AI condition compared to the student condition ($p < 0.01$). This indicated algorithm appreciation for the non-expert case.

We did not find statistically significant coefficients for ATI in both models ($p > 0.1$). Our results suggested no significant differences based on domain knowledge in the specified models ($p > 0.1$). The results for age and the gender female themselves did not indicate a significant effect on WOA ($p > 0.1$). Omitting the ATI, domain knowledge, age, and female did not change the interpretation of the regression results for our variables of interest.

The intra-class correlation (ICC) for Participant was 0.22 in both models. For Task, it was 0.02. Due to the small ICC ($\rho < 0.05$), the task could arguably be omitted. However, this only changed the coefficients in the third decimal place and caused no changes with respect to the significance thresholds. In summary, we find support for a human expert effect in advice-taking in Study 1, but cannot confirm an expert effect between the algorithmic conditions in an algorithm appreciation scenario.

Study 2

Table 3 shows the results of Study 2 in the models (S2_Student) and (S2_Stat Model). Expertise increased WOA by 0.109 for the human condition in model (S2_Student). This coefficient was statistically significant ($p < 0.01$). A 10.9%-point increase in WOA due to high expertise appeared to be practically meaningful. Thus, we found support for our first hypothesis.

Our results from model (S2_Stat Model) indicated no significant difference between the statistical model and the AI ($p > 0.1$). A 0.4%-point increase in WOA due to higher expertise in the algorithmic condition does not appear to be practically meaningful either. Thus, we could not confirm our second hypothesis that an expert effect is prevalent for algorithmic labels in a general population.

Similar to Study 1, the results indicated a 0.095 higher WOA for statistical models compared to students ($p < 0.05$). The coefficient for the AI in model (S2_Student) indicated algorithm appreciation in the form of an increase in advice-taking by 0.099 if the advice came from an AI compared to advice from a student ($p < 0.05$). However, the results revealed no statistically significant difference between the statistical model and the industry expert ($p < 0.1$). This confirmed the observation of algorithm appreciation in the non-expert case from Study 1.

We did not find statistically significant coefficients for ATI, age, and a female gender ($p > 0.1$). However, our results suggested a significant decrease in WOA of 0.025 based on domain knowledge in the specified model ($p < 0.1$). Just as in Study 1, omitting the control variable did not change the interpretation of the regression results for the variables of interest.

The ICC for Participant was 0.40 and 0.02 for Task. Omitting Task (as $\rho < 0.05$) did not change coefficients except for the third decimal place. In summary, we found support for a human expert effect in advice-taking but not for an algorithmic expert effect in an algorithm appreciation scenario in Study 2.

Robustness Checks

We've included the results of an alternative formulation of our regression in Figure 4. The two main variables of interest in the models are High Expertise and Algorithmic Advisor. The binary dummy variable High Expertise took the value 1 for high expertise (AI, industry expert) and 0 for low expertise (statistical model, student). Similarly, the binary dummy variable Algorithmic Advisor took the value 1 for the algorithmic condition (AI, statistical model) and 0 for the human condition (industry expert, student). The interaction coefficient between these two variables controlled whether the effect of one variable depended on the other.

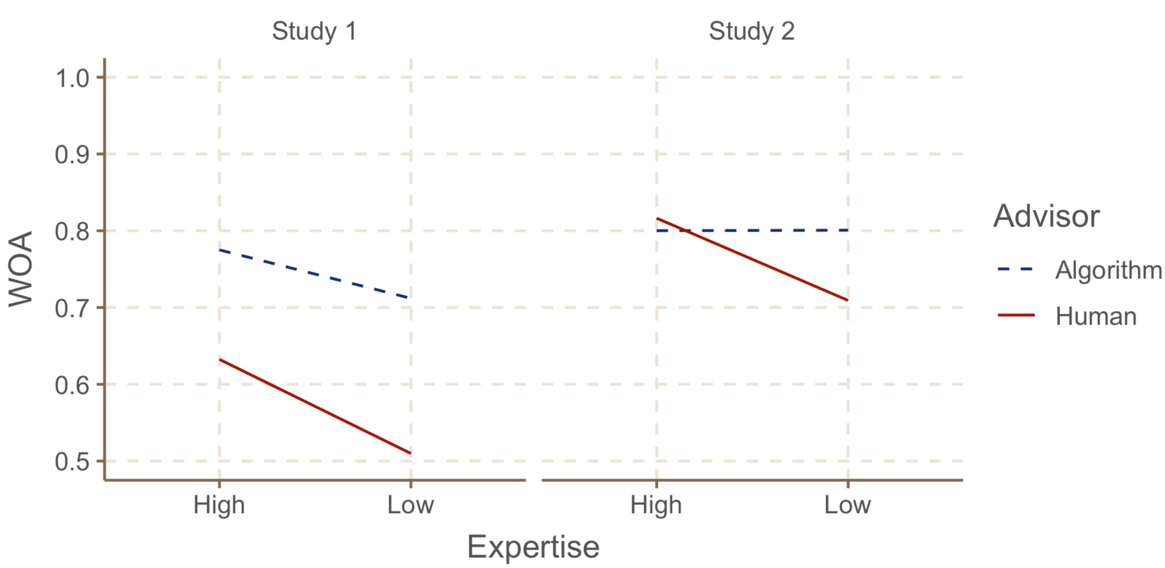

Figure 4. Interaction Plot of Expertise and WOA for the Two Advisor Types

In Study 1, the results indicated a 0.125 higher WOA on average for the expertise condition compared to the non-expert condition. This coefficient was statistically significant ($p < 0.05$). A 12.5%-point increase in WOA due to high expertise appeared to be practically meaningful. The coefficients for the algorithmic condition indicated an increase of 20.6%-points if the advice came from an algorithm compared to advice that came from a human ($p < 0.01$). An insignificant interaction of High Expertise and Algorithmic Advisor of -0.066 ($p > 0.1$) indicated no significant dependence between the algorithmic and expertise factors. Thus, the results from our robustness check imply an expert effect across both human and algorithmic conditions in an algorithm appreciation scenario. However, the measured interaction term is negative and a non-significant interaction term does not provide evidence against the results from Study 1.

Similar to Study 1, the results from Study 2 indicated an expert effect of 10.9% ($p < 0.01$) and algorithm appreciation of 9.5% ($p < 0.05$). However, the coefficients implied an interaction coefficient between high expertise and algorithmic advisors of -10.5% ($p < 0.05$). Thus, this interaction coefficient mediates the positive association of expertise for the algorithmic condition, confirming the results from our previous analyses. In summary, our robustness checks support our results from Study 1 and 2.

Post-hoc analysis

We also investigated perceived differences in expertise. Figure 5 shows the results of the post-experiment questionnaire. In study 1, the student condition had the lowest expertise rating of 4.53 on average. The industry expert had the highest rating, with an average rating of 8.58. We conducted a Wilcoxon signed-rank test, which indicated that the location shift is not equal to 0 ($p < 0.01$). We found similar results in study 2 ($p < 0.01$). Thus, these results supported our initial assumption that industry experts are associated with more expertise than students.

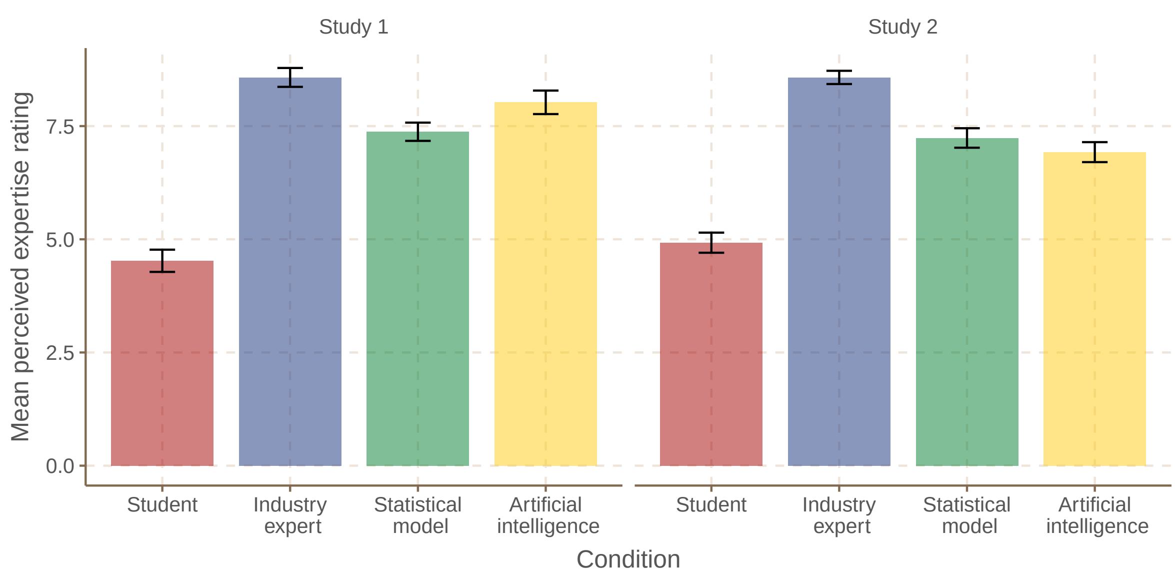

Figure 5. Self-Reported Perceptions by Conditions

In Study 1, the statistical model had an average rating of 7.38. AI had an average perceived expertise rating of 8.03. We conducted a Wilcoxon signed-rank test, which indicated that the true location shift is not equal to 0 ($p = 0.05$). In study 2, the statistical model received a 7.24 average rating. The average perceived level of expertise for AI was 6.93. A Wilcoxon signed-rank test was not significant in study 2 ($p > 0.1$). Overall, both post-hoc analyses confirmed a perceived expert association for human advisors. The results of the first post-hoc analysis confirmed our original assumption that AI is associated with more expertise compared to a statistical model. Based on the results from post-hoc analysis of Study 2, we could not confirm a perceived expertise difference between these conditions.

Discussion

The objective of this paper was to investigate the impact of expert framing on algorithmic advice utilization. As prior research on this topic used surveys with self-reported measures and focused on perceived intentions (e.g., Langer et al., 2022), our results from two incentivized experiments on regression tasks add new and robust empirical evidence on the question whether the labeling of algorithmic advisor matters to users. In a nutshell, we found an expert effect for human advisors in our experiments, but no expert effect for algorithmic advisors. Hence, our research implies that product owners should prioritize communicating the functional and non-functional attributes of their systems, rather than relying on buzzwords like AI.

Our experiments confirmed our first hypothesis that framing *human* advisors as experts increases advice utilization. This finding is in line with prior empirical studies (Madhavan & Wiegmann, 2007; Pearson & Mayhorn, 2017). Surprisingly and contrary to conjectures in prior work (Langer & Landers, 2019), we could not confirm our second hypothesis that there is a similar expert framing effect for *algorithmic* advisors.

Regarding self-reported perception of expertise, the technology-savvy population in Study 1 rated the perceived expertise of AI significantly higher than the perceived expertise of statistical models. In contrast, for the more general population in Study 2, we did not observe significant differences in perceived expertise between statistical models and AI. Taken together, these results suggest that ”AI washing” may not be effective, or may only affect the perceptions of technology-savvy users, but not their actions.

This study further extends the existing literature on algorithm aversion by experimentally examining the effect of *single words or labels* on advice-taking. Prior research has noticed differences in algorithm advice utilization depending on the framing of advisors (Jussupow et al., 2020). But previous studies have used lengthy textual framings as a factor to explain differences in advice-taking and simultaneously modified various advisor characteristics. Examples of such characteristics are years of experience, education, and amount of advisors (e.g., Hou & Jung, 2021; Madhavan & Wiegmann, 2007; Pearson & Mayhorn, 2017). Extending these insights, our results provide empirical evidence that even single labels are sufficient to trigger such an expert effect.

Regarding our second hypothesis, we could not confirm the conjecture of Langer & Landers (2019) that labels for algorithms have an impact on advice utilization. We used a metric frequently employed in research on algorithm aversion, namely WOA (e.g., Berger et al., 2021; Logg et al., 2019). Thus, this study extends the survey of Langer et al. (2022) on algorithmic perception by providing experimental evidence of non-significant differences in advice-taking from algorithmic labels for a within-subjects regression task.

One possible explanation for an expert effect is the portrayal of experience and knowledge in the media (see Önkal et al., 2009). Recently, AI has received significant attention in the media (e.g., Tarrant, 2023). However, we only observed higher perceived expertise associated with AI in Study 1, and we did not observe differences in advice-taking.

People tend to prefer labels they are more familiar with (Shank et al., 2019). A ”highly complex algorithm” can lead to more algorithm aversion compared to a ”simple algorithm” (see Ganbold et al., 2022). Deterrence of the unknown could explain the differences in perceived expertise between the technology-savvy population in Study 1 and the more general population in Study 2. Furthermore, the ongoing debate about the limitations of AI (e.g., Lebovitz et al., 2021), could explain the lack of significant differences in actual advice-taking. Due to a lack of experience with AI, individuals are unable to overcome their unconscious bias towards AI (Turel & Kalhan, 2023).

The question at hand is why companies involve in AI-washing when it fails to contribute substantial value. Firstly, it's important to note that insignificant results do not necessarily indicate that the AI label is inferior to the label ”statistical model”. Secondly, one plausible explanation could be that while the AI label may not directly affect advice utilization, it may have the potential to shape the perceptions of the product's or company's novelty and computational advancement.

Nonetheless, AI is not without controversy, exemplified by the AI Act, which the European Parliament recently passed. The AI Act classifies AI applications into certain risk classes, determining the extent of legal obligations imposed. Consequently, a situation opposite to AI-washing could emerge: Companies might be inclined to conceal the AI content of their applications. While our study's findings suggest that such reverse AI-washing does not notably impact advice-taking, we emphasize the need for caution against such practices. This caution is rooted in the potential legal consequences that render such washing both unprofitable and unethical.

In both studies, we observed algorithm appreciation or no difference in advice utilization from algorithmic or human advisors for a regression task. This finding aligns with previous research that observed greater algorithm appreciation for objective tasks compared to subjective tasks (Castelo et al., 2019). Taking advice from algorithms in more subjective tasks can lead to a dehumanizing experience, where people associate algorithms with a lack of intuitive and subjective judgment capabilities, rather than recognizing their efficiency and objectivity (Lee, 2018). Consequently, the level of algorithm appreciation might diminish for non-regression tasks. It is worth noting that the data in our study may be more comprehensive compared to other applications, which could suffer from more unique characteristics (e.g., Longoni et al., 2019). However, regardless of the application area, evaluations of algorithms and research on algorithmic advice utilization should be aware of single labels turning an algorithm appreciation situation into non-significant differences in advice utilization.

People sometimes don't use what they don't know and product owners should take this into account when designing intelligent systems. Although hopes in AI systems can be high in the development phase, with developers perceiving the potential of AI to be greater than conventional statistical models, it is important to note that people may not rely on AI advice until its performance is demonstrated to be better than other alternatives. Therefore, AI should not be used because it is a buzzword, but in situations where it brings added value. Simple AI washing without real substance does not work.

Limitations & Outlook

Exploring the effects of labels on algorithmic advice utilization is a promising avenue for future research. For example, the observed effects could be larger for individuals with a negative attitude towards algorithms (Langer et al., 2022). We focused on competent advisors, while failing algorithms might disappoint (Prahl & Swol, 2017) or require a recovery from initial aversion (Leffrang et al., 2023). Additionally, terminology could differ in dimensions, such as cognitive flexibility (Longoni et al., 2019) or task characteristics (Lee, 2018). Future work can examine whether people understand how expertise manifests in algorithms or whether additional cues on the sophistication of algorithms are necessary.

As not all factors influencing algorithm appreciation are known (Jussupow et al., 2020), we did not preregister hypotheses on algorithm appreciation. Other tasks and measures can enhance insights beyond regression. We acknowledge that WOA is a common dependent variable in the algorithm appreciation literature. However, it is important to note that WOA has limitations (Bonaccio & Dalal, 2006). While the evaluation of advice utilization among students and crowdworkers is a prevalent research design within the field of algorithm aversion (e.g., Fügener et al., 2022), subsequent studies could focus on more specific populations in diverse scenarios, like long-term predictions with many possible futures.

Finally, framing is not the only factor influencing algorithmic advice utilization. While our specific task involves users' estimation and decision-making, we recognize that the advisory process encompasses further stages like data collection, model training, and advice communication. We encourage further research to explore other factors, like different tasks and advisors' quality, and their interactions with framing to gain a more comprehensive understanding of algorithmic advice utilization in different contexts.

References

Berente, N., Gu, B., Recker, J. & Santhanam, R. (2021). "Managing artificial intelligence," *MIS Quarterly* (45:3), 1433–1450. http://doi.org/10.25300/MISQ/2021/16274.

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

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

Camerer, C. F. & Hogarth, R. M. (1999). "The effects of financial incentives in experiments: A review and capital-labor-production framework," *Journal of risk and uncertainty* (19:1), 7–42. http://doi.org/10.1023/A:1007850605129.

Castelo, N., Bos, M. W. & Lehmann, D. R. (2019). "Task-dependent algorithm aversion," *Journal of Marketing Research* (56:5), 809–825. http://doi.org/10.1177/0022243719851788.

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. http://doi.org/10.1037/xge0000033.

Franke, T., Attig, C. & Wessel, D. (2019). "A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ATI) scale," *International Journal of Human-Computer Interaction* (35:6), 456–467. http://doi.org/10.1080/10447318.2018.1456150.

Fügener, A., Grahl, J., Gupta, A. & Ketter, W. (2022). "Cognitive challenges in human–artificial intelligence collaboration: Investigating the path toward productive delegation," *Information Systems Research* (33:2), 678–696. http://doi.org/10.1287/isre.2021.1079.

Ganbold, O., Rose, A. M., Rose, J. M. & Rotaru, K. (2022). "Increasing reliance on financial advice with avatars: The effects of competence and complexity on algorithm aversion," *Journal of Information Systems* (36:1), 7–17. http://doi.org/10.2308/ISYS-2021-002.

Gigerenzer, G. (2018). "The bias bias in behavioral economics," *Review of Behavioral Economics* (5), 303–336. http://doi.org/10.1561/105.00000092.

Gino, F. & Moore, D. A. (2007). "Effects of task difficulty on use of advice," *Journal of Behavioral Decision Making* (20:1), 21–35. http://doi.org/10.1002/bdm.539.

Goldsmith, D. J. & Fitch, K. (1997). "The normative context of advice as social support," *Human Communication Research* (23:4), 454–476. http://doi.org/10.1111/j.1468-2958.1997.tb00406.x.

Gonzalez, M., Capman, J., Oswald, F., Theys, E. & Tomczak, D. (2019). "'Where's the I-O?' Artificial Intelligence and Machine Learning in Talent Management Systems," *Personnel Assessment and Decisions* (5:3), 33–44. http://doi.org/10.25035/pad.2019.03.005.

Grgic-Hlaca, N., Engel, C. & Gummadi, K. P. (2019). "Human decision making with machine advice: An experiment on bailing and jailing," in: *Proceedings of the ACM on Human-Computer Interaction,* 1–25. http://doi.org/10.1145/3359280.

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. http://doi.org/10.1037/1040-3590.12.1.19.

Harvey, N. & Fischer, I. (1997). "Taking advice accepting help, improving judgment," *Organizational Behavior and Human Decision Processes* (70:2), 117–133. http://doi.org/10.1006/obhd.1997.2697.

Heßler, P. O., Pfeiffer, J. & Hafenbrädl, S. (2022). "When self-humanization leads to algorithm aversion: What users want from decision support systems on prosocial microlending platforms," *Business and Information Systems Engineering* (64:3), 275–292. http://doi.org/10.1007/s12599-022-00754-y.

Highhouse, S. (2008). "Stubborn reliance on intuition and subjectivity in employee selection," *Industrial and Organizational Psychology* (1:3), 333–342. http://doi.org/10.1111/j.1754-9434.2008.00058.x.

Hou, Y. T. Y. & Jung, M. F. (2021). "Who is the expert? Reconciling algorithm aversion and algorithm appreciation in AI-supported decision making," in: *Proceedings of the ACM on Human-Computer Interaction,* 1–25. http://doi.org/10.1145/3479864.

Jussupow, E., Benbasat, I. & Heinzl, A. (2020). "Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion," in: *Proceedings of the European Conference on Information Systems (ECIS '20),* 1–16. https://aisel.aisnet.org/ecis2020_rp/168.

Keel, S., Lee, P. Y., Scheetz, J., Li, Z., Kotowicz, M. A., MacIsaac, R. J. & He, M. (2018). "Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: A pilot study," *Scientific Reports* (8:1), 8–13. http://doi.org/10.1038/s41598-018-22612-2.

Kramer, M. F., Schaich Borg, J., Conitzer, V. & Sinnott-Armstrong, W. (2018). "When do people want AI to make decisions?," in: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES 2018),* 204–209. http://doi.org/10.1145/3278721.3278752.

Langer, M., Hunsicker, T., Feldkamp, T., König, C. J. & Grgić-Hlača, N. (2022). ""Look! It's a computer program! It's an algorithm! It's AI!": Does terminology affect human perceptions and evaluations of algorithmic decision-making systems?," in: *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '22),* 1–28. http://doi.org/10.1145/3491102.3517527.

Langer, M., König, C. J. & Hemsing, V. (2020). "Is anybody listening? The impact of automatically evaluated job interviews on impression management and applicant reactions," *Journal of Managerial Psychology* (35:4), 271–284. http://doi.org/10.1108/JMP-03-2019-0156.

Langer, M. & Landers, R. N. (2019). "The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers," *Computers in Human Behavior* (106878), 9–25. http://doi.org/10.1016/j.chb.2021.106878.

Lebovitz, S., Levina, N. & Lifshitz-Assaf, H. (2021). "Is AI ground truth really true? The dangers of training and evaluating AI tools based on experts' know-what," *MIS Quarterly* (45:3), 1501–1525. http://doi.org/10.25300/MISQ/2021/16564.

Lee, M. K. (2018). "Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management," *Big Data and Society* (5:1), 1–16. http://doi.org/10.1177/2053951718756684.

Lee, M. K. & Rich, K. (2021). "Who Is included in human perceptions of AI? Trust and perceived fairness around healthcare AI and cultural mistrust," in: *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '21),* http://doi.org/10.1145/3411764.3445570.

Leffrang, D., Bösch, K. & Müller, O. (2023). "Do people recover from algorithm aversion? An experimental study of algorithm aversion over time," in: *Proceedings of the Hawaii International Conference on System Sciences (HICSS '23),* 4016–4025. https://hdl.handle.net/10125/103122.

Levin, I. P., Schneider, S. L. & Gaeth, G. J. (1998). "All frames are not created equal: A typology and critical analysis of framing effects," *Organizational Behavior and Human Decision Processes* (76:2), 149–188. http://www.sciencedirect.com/science/article/pii/S0749597898928047.

Leyer, M. & Schneider, S. (2019). "Me, you or ai? How do we feel about delegation," in: *Proceedings of the European Conference on Information Systems (ECIS '19),* 1–17. https://aisel.aisnet.org/ecis2019_rp/36.

Logg, J. M., Minson, J. A., Moore, D. A. & States, U. (2019). "Algorithm appreciation: People prefer algorithmic to human judgment," *Organizational Behavior and Human Decision Processes* (151), 90–103. http://doi.org/10.1016/j.obhdp.2018.12.005.

Longoni, C., Bonezzi, A. & Morewedge, C. K. (2019). "Resistance to medical artificial intelligence," *Journal of Consumer Research* (46:4), 629–650. http://doi.org/10.1093/jcr/ucz013.

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

Mahmud, H., Islam, A. K., Ahmed, S. I. & Smolander, K. (2022). "What influences algorithmic decision-making? A systematic literature review on algorithm aversion," *Technological Forecasting and Social Change* (175), 1–26. http://doi.org/10.1016/j.techfore.2021.121390.

Montague, E. N., Kleiner, B. M. & Winchester, W. W. (2009). "Empirically understanding trust in medical technology," *International Journal of Industrial Ergonomics* (39:4), 628–634. http://doi.org/10.1016/j.ergon.2009.01.004.

Moore, S. (2017). *Gartner says AI technologies will be in almost every new software product by 2020*. Gartner Inc., July 18. (https://www.gartner.com/en/newsroom/press-releases/2017-07-18-gartner-says-ai-technologies-will-be-in-almost-every-new-software-product-by-2020, accessed September 08, 2022).

Moser, A. K. (2015). "Thinking green, buying green? Drivers of pro-environmental purchasing behavior," *Journal of Consumer Marketing* (32:3), 167–175. http://doi.org/10.1108/JCM-10-2014-1179.

Önkal, D., Goodwin, P., Thomson, M., Gönül, S. & Pollock, A. (2009). "The relative influence of advice from human experts and statistical methods on forecast adjustments," *Journal of Behavioral Decision Making* (22:4), 390–409. http://doi.org/10.1002/bdm.637.

Ötting, S. K. & Maier, G. W. (2018). "The importance of procedural justice in human–machine interactions: Intelligent systems as new decision agents in organizations," *Computers in Human Behavior* (89), 27–39. http://doi.org/10.1016/j.chb.2018.07.022.

Oxford University Press (2022). "Buzzword," in: *Oxford English Dictionary,* http://doi.org/10.1093/OED/3991639983.

Oxford University Press (2023a). "Expertise," in: *Oxford English Dictionary,* http://doi.org/10.1093/OED/3325950481.

Oxford University Press (2023b). "Greenwashing," in: *Oxford English Dictionary,* http://doi.org/10.1093/OED/3883558326.

Pearson, C. J. & Mayhorn, C. B. (2017). "The effects of pedigree and source type on trust in a dual adviser context," in: *Proceedings of the Human Factors and Ergonomics Society,* 319–323. http://doi.org/10.1177/1541931213601561.

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

Saragih, M. & Morrison, B. W. (2022). "The effect of past algorithmic performance and decision signifi-cance on algorithmic advice acceptance," *International Journal of Human-Computer Interaction* (38:13), 1228–1237. http://doi.org/10.1080/10447318.2021.1990518.

Shank, D. B., Burns, A., Rodriguez, S. & Bowen, M. (2019). "Software program, bot, or artificial intelligence? Affective sentiments across general technology labels," *Current Research in Social Psychology* (28), 32–41.

Shibl, R., Lawley, M. & Debuse, J. (2013). "Factors influencing decision support system acceptance," *Decision Support Systems* (54:2), 953–961. http://doi.org/10.1016/j.dss.2012.09.018.

Spence, M. (1973). "Job market signaling," *The Quarterly Journal of Economics* (87:3), 355. http://doi.org/10.2307/1882010.

Tarrant, G. (2023). *Generative AI is already changing white-collar work as we know it*. The Wall Street Journal, March 29. (https://www.wsj.com/articles/generative-ai-is-already-changing-white-collar-work-as-we-know-it-58b53918, accessed April 13, 2023).

Turel, O. & Kalhan, S. (2023). "Prejudiced against the machine? Implicit associations and the transience of algorithm aversion," *MIS Quarterly* (Forthcoming), 1–26. http://doi.org/10.25300/MISQ/2022/17961.

Tversky, A. & Kahneman, D. (1974). "Judgment under uncertainty: Heuristics and biases," *Science* (185:4157), 1124–1131. http://doi.org/10.1126/science.185.4157.1124.

Tversky, A. & Kahneman, D. (1981). "The framing of decisions and the psychology of choice," *Experiments in Environmental Economics* (211:4481), 453–458. http://doi.org/10.1007/978-1-4613-2391-4_2.