Ostfalia
University of
Applied Sciences
Wolfsburg
Faculty of Business

zeGI
ZENTRUM
FÜR GESELL
SCHAFTLICHE
INNO
VATION

WWP

Wolfsburg Working Papers

No. 24-04

Choice Overload as an antidote to algorithm aversion: ineffective for men and exacerbating for women

Ibrahim Filiz, Florian Kirchhoff, Thomas Nahmer and Markus Spiwoks, November 2024

Choice Overload as an antidote to algorithm aversion: ineffective for men and exacerbating for women

Ibrahim Filiz, Florian Kirchhoff, Thomas Nahmer and Markus Spiwoks

JEL codes: D81, D91, G41, I12, J16, O33

Key Words: Algorithm aversion, choice overload, gender differences, decision making under risk, behavioral economics, experiments.

Abstract:

What influence does the number of action alternatives have on the extent of algorithm aversion? This constitutes the research question of the present study. Research results in the area of choice overload indicate that a large number of alternatives often leads subjects to select an easily justifiable, i.e. obviously expedient, alternative. In laboratory experiments on algorithm aversion, subjects often have to decide whether they want to have a task performed by an expert or by a superior algorithm. If instead there are several experts and several superior algorithms to choose from, subjects may feel compelled to choose an easily justifiable alternative. In view of the higher success rates of the algorithms, this would be a decision in favor of an algorithm. Choice overload could thus reduce the tendency towards algorithm aversion. However, the results of the laboratory experiment conducted do not support this assumption. While the number of alternatives had no effect on the male subjects, the opposite effect was observed in the female subjects. A larger number of alternatives does not dampen the tendency towards algorithm aversion in women, but actually increases it significantly.

Ibrahim Filiz, Ostfalia University of Applied Sciences, Faculty of Business, Siegfried-Ehlers-Str. 1, D-38440 Wolfsburg, Germany, e-mail: ibrahim.filiz@ostfalia.de

Florian Kirchhoff, Ostfalia University of Applied Sciences, Faculty of Business, Siegfried-Ehlers-Str. 1, D-38440 Wolfsburg, Germany, e-mail: fl.kirchhoff@ostfalia.de, https://orcid.org/0009-0009-2482-1560

Thomas Nahmer, Ostfalia University of Applied Sciences, Faculty of Business, Siegfried-Ehlers-Str. 1, D-38440 Wolfsburg, Germany, e-mail: th.nahmer@ostfalia.de

Markus Spiwoks, Ostfalia University of Applied Sciences, Faculty of Business, Siegfried-Ehlers-Str. 1, D-38440 Wolfsburg, Germany, Tel.: +49 5361 892 225 100, e-mail: m.spiwoks@ostfalia.de

Introduction

For many tasks that are traditionally performed by experts, there are now algorithms available that can perform these tasks with the same or even greater reliability. Nevertheless, unless these algorithms achieve a 100% success rate, many subjects remain skeptical about employing them. This is often the case even when it is evident that the algorithm's success rate is significantly higher than that of the expert. This behavioral anomaly is referred to as algorithm aversion (Dietvorst, Simmons & Massey, 2015).

The phenomenon of algorithm aversion has been studied intensively over the past ten years. Laboratory experiments play a prominent role in this research. In experimental decision-making situations, subjects are typically presented with a choice between performing a task themselves or having it performed by an algorithm. Alternatively, they may choose to delegate the task to an expert or to an algorithm.

However, empirical consumer research demonstrates that the decision-making behavior of subjects is also influenced by the number of available alternatives. When consumers are required to select from a large number of product alternatives, the decision-making process is often considerably more challenging than when there is only a small selection available. Indeed, research has demonstrated that when consumers are presented with an excess of options, they are often unable to reach a purchase decision at all. This phenomenon is referred to as choice overload (see Iyengar & Lepper, 2000; Scheibehenne, Greifeneder & Todd, 2010; Park & Jang, 2013; Chernev, Böckenholt & Goodman, 2015; Thai & Yuksel, 2017; Jacob, Thomas & Joseph, 2024).

In an analysis of the choice overload anomaly, Sela, Berger & Liu (2009) reached an interesting conclusion: individuals facing a multitude of options tend to select a decision that is relatively easy to justify.

This may serve as a potential measure to assist in the mitigation of algorithm aversion. If there is not just one algorithm or one expert to choose from, but several algorithms and several experts, it is possible that a decision that is relatively straightforward to justify may be reached with greater frequency. If all algorithms are clearly superior to all experts in terms of their probability of success, the decision in favor of one of the algorithms would be easy to justify. If there are several experts and several superior algorithms to choose from, this could contribute to a reduction of algorithm aversion. The objective of this study is to investigate whether this is, in fact, the case.

Literature review

Algorithm aversion

Automated decision-making processes or decision-making aids, also known as algorithms, are becoming increasingly important. Powerful algorithms have been available for decades (cf. e.g. Dawes, Faust & Meehl, 1989). However, some subjects exhibit a negative attitude towards algorithms, even if it is clearly recognizable that algorithm-based decisions are associated with a significantly higher success rate than their own decisions or expert decisions. This phenomenon is referred to as algorithm aversion (Dietvorst, Simmons & Massey, 2015). However, there are differences in the definition of the term. Filiz et al. (2023) and Jussupow, Benbasat & Heinzl (2020) highlight the following aspects of differing understandings: the experience of the superiority of the algorithm (only perceptible or clearly recognizable and/or quantifiable), the reliability of the algorithm (perfect or imperfect), the observation of reliability (visible occurrence of errors or not), the frequency of highlighting success rates (once or repeatedly), the use of historical data (only for the creation or also in the operation of the algorithm), the setting (algorithm vs. expert, algorithm vs. layperson or algorithm vs. acting subject), and the extent of algorithmic intervention (algorithm only acts as a decision-making aid or carries out the activity independently).

Despite all the differences in detail, there is consensus that the behavioral anomaly of algorithm aversion can result in significant economic disadvantages. In view of its enormous practical relevance, a lively scientific debate about the causes of algorithm aversion has therefore developed over the past ten years (cf. e.g. Sunstein & Gaffe, 2024), but also about ways to reduce algorithm aversion. Dietvorst, Simmons & Massey (2018), for example, indicate that the tendency towards algorithm aversion can be significantly reduced when the individuals involved are provided the chance to alter the outcome of the algorithm - even if only to a minor extent (cf. also Gubaydullina et al., 2022). Filiz et al. (2021) demonstrate that as individuals gain experience with an algorithm through regular feedback on its success rates, they may gradually become less averse to it (cf. also Leffrang, Bösch & Müller, 2023). Watson (2024) emphasizes the importance of information regarding the reliability of the algorithm in mitigating algorithm aversion. As demonstrated by Berger et al. (2021), individuals engaged in economic activities who observe the learning progress of an algorithm are less prone to exhibit algorithm aversion (cf. also Reich, Kaju & Maglio, 2023). As indicated by Bogert, Schecter & Watson (2021), the greater the complexity and difficulty of the task at hand, the less reluctance there is to utilize algorithms. Furthermore, greater time pressure when completing tasks has been shown to contribute to a reduction in algorithm aversion (cf. Jung & Seiter, 2021). Castelo, Bos & Lehmann (2019) posit that an increase in the perceived affective (emotional) human likeness of algorithms is suitable for reducing algorithm aversion, particularly in tasks that are open to interpretation and based on personal opinion or intuition (cf. also Zhao et al., 2024). As Judek (2024) notes, knowledge about a high level of acceptance by other users contributes to a reduction in reluctance to use an algorithm. Efendic, Van de Calseyde & Evans (2020) demonstrate that reservations about algorithms are lower when they reach a result rapidly. As indicated by Kim, Giroux & Lee (2021), the presentation of a result with numerous decimal places is associated with greater trust in the algorithm, which in turn counteracts algorithm aversion.

Choice Overload

In three field and laboratory experiments, Iyengar & Lepper (2000) demonstrate that the number of choices available can significantly influence the decision-making behavior of subjects. It is noteworthy that both the willingness to make a decision and the subjective satisfaction with a decision made decrease when a larger number of alternatives are available. This phenomenon has been termed *choice overload* by the researchers. The term is also employed by Mogilner, Rudnick & Iyengar (2008), Diehl & Poynor (2010), Scheibehenne, Greifeneder & Todd (2010), Park & Jang (2013), Chernev, Böckenholt & Goodman (2015), Thai & Yuksel (2017) and Jacob, Thomas & Joseph (2024). However, slightly different terms are also used in the literature, including the *over choice effect* (Gourville & Soman, 2005), the *problem of too much choice* (Fasolo, McClelland & Todd, 2007), the *tyranny of choice* (Schwartz, 2000), the *too-much-choice effect* (Scheibehenne, Greifeneder & Todd, 2009), *consumer hyperchoice* (Mick, Broniarczyk & Haidt, 2004), or the *paradox of choice* (Schwartz, 2004).

To date, around 100 studies have been conducted on the phenomenon of choice overload (cf. Jacob, Thomas & Joseph, 2024). It is evident that the empirical findings can vary significantly depending on the contextual factors, including, for example, the number of alternatives, the number of evaluation criteria for the alternatives, the abundance of information, and the time budget for the potential selection decision. The possible links to existing psychological and/or economic theories (e.g. *construal level theory*, *diminishing marginal utility theory*, *cognitive load theory*, *cognitive dissonance theory*, *structural alignment theory*, *unconscious thought theory*, *lay theory, prospect theory*, *regulatory focus* *theory*, *action identification theory*, *dual coding theory*, *regulatory mode theory*, *attentional resource theory*, *search cost theory*, *complexity theory*, *decision field theory*) remain a topic of intense debate (Jacob, Thomas & Joseph, 2024). Nevertheless, one outcome can be regarded as definitive: the number of alternatives can exert a significant influence on the decision-making behavior of subjects.

One particularly noteworthy finding from the research merits further examination. Sela, Berger & Liu (2009) demonstrate in five experimental surveys that the presence of a large number of alternatives contributes to the selection of easily justifiable alternatives with greater frequency. This increases the frequency with which virtuous and expedient alternatives are chosen. In contrast, purely pleasurable, hedonistic alternatives are chosen less frequently. Consequently, an increase in the number of alternatives tends to result in fewer profligate and more virtuous choices being made, and to more attention being paid to expediency than to the promise of pleasure.

Additionally, there is a debate as to whether choice overload affects women and men in different ways. A greater emphasis on detail in women (cf. Kempf, Laczniak & Smith, 2006) results in a higher incidence of choice overload and a more pronounced impact in women than in men (cf. e.g. Misuraca et al., 2024; Ji, Ha & Sypher, 2014; Williamson & Eaker, 2012).

Hypotheses and experimental design

We situate the experiment in the context of a medical decision-making process. There are two reasons for this. On the one hand, there are highly sophisticated algorithms available, particularly in the medical domain, which elicit considerable scepticism from both patients and medical practitioners (Longoni, Bonezzi & Morewedge, 2019; Yun, Eun-Ju & Dong, 2021; Robertson et al, 2023; Bankuoru Egala & Liang, 2023; Mellers, Lu & McCoy, 2023; Xue et a., 2024). On the other hand, Filiz et al. (2023) and Filiz et al. (2024) have demonstrated that algorithm aversion is particularly prevalent when there is a risk of severe consequences if an erroneous decision is made. This is often the case with decisions in the healthcare sector.

This is a vignette study in which the subjects act as health advisors for a patient. The patient has undergone a surgical procedure to remove a malignant melanoma, a form of skin cancer with a propensity for metastasis. To exclude the possibility that the skin cancer had spread to other parts of the patient's body, a comprehensive examination was conducted. This entailed the utilization of a magnetic resonance imaging (MRI) scanner in conjunction with other diagnostic techniques. The question now arises as to who should be responsible for the final assessment of all the examination results.

Two alternatives are available for <u>Treatment 1:</u>

- An experienced doctor who is both an oncologist (doctor specializing in cancer treatment) and a radiologist (doctor specializing in imaging techniques) evaluates all the results (option A). In 60% of cases, the doctor obtains an examination result that is also confirmed in the medium term.
- A specialized computer program evaluates all examination results. It is designed in particular for pattern recognition in MRI scans (option B). In 70% of cases, the computer program achieves an examination result that is also confirmed in the medium term.

A total of six alternatives are available in <u>Treatment 2:</u>

- An experienced doctor who is both an oncologist (doctor specializing in cancer treatment) and a radiologist (doctor specializing in imaging techniques) evaluates all the results (option A). In

60% of cases, the doctor obtains an examination result that is also confirmed in the medium term.

- An experienced doctor who is both an oncologist (doctor specializing in cancer treatment) and a radiologist (doctor specializing in imaging techniques) evaluates all results together with a colleague. The dual control principle is intended to avoid careless errors (option B). In 60% of cases, the medical team achieves an examination result that is also confirmed in the medium term.
- An experienced doctor, who is both an oncologist (doctor specializing in cancer treatment) and a radiologist (doctor specializing in imaging techniques), evaluates all results. In doing so, he strictly follows a diagnostic plan that has been drawn up by an international medical commission. The diagnostic plan is intended to prevent decisions being made on the basis of incomplete data (option C). In 60% of cases, the doctor achieves an examination result that is also confirmed in the medium term.
- A specialized computer program evaluates all examination results. It is designed in particular for pattern recognition in MRI scans (option D). In 70% of cases, the computer program achieves an examination result that is also confirmed in the medium term.
- A specialized computer program evaluates all examination results, but places particular emphasis on the inclusion of electron and fluorescence microscopy results (option E). In 70% of cases, the computer program achieves an examination result that is also confirmed in the medium term.
- A specialized computer program evaluates all examination results, but places particular emphasis on using neural networks to better evaluate the raw data space (k-space) in order to avoid image noise and artefacts (option F). In 70% of cases, the computer program achieves an examination result that is also confirmed in the medium term.

Treatment 1 reflects the typical decision situation encountered in experimental studies on algorithm aversion, in which the test person is required to select between an expert and an algorithm. Treatment 2 presents a broader range of alternatives, requiring the respondent to select one of three experts or one of three algorithms. Consequently, there are six potential options for consideration.

If Sela, Berger & Liu (2009) are correct in their assessment that economic agents frequently select an easily justifiable alternative when presented with a multitude of options, the incidence of algorithm aversion should be lower in Treatment 2 than in Treatment 1. Given that all algorithms outperform all experts in terms of success rate, an emphasis on expediency, which can result from choice overload, would lead to a reduction in algorithm aversion in Treatment 2.

<u>Hypothesis 1</u> therefore is: The frequency with which an algorithm is selected is significantly greater in Treatment 2 (six alternatives) than in Treatment 1 (two alternatives).

<u>Null hypothesis 1</u> then reads: The frequency with which an algorithm is selected is not significantly greater in Treatment 2 (six alternatives) than in Treatment 1 (two alternatives).

If the findings of Misuraca et al. (2024), Ji et al. (2014) and Williamson & Eaker (2012) that women are more susceptible to choice overload than men are accurate, and if it is also true that choice overload has a dampening effect on algorithm aversion, then it can be expected that women in Treatment 2 (six alternatives) will demonstrate even lower levels of algorithm aversion than men.

<u>Hypothesis 2</u> is therefore: Women select an algorithm significantly more often than men in Treatment 2 (six alternatives).

Null hypothesis 2 is therefore: Women do not select an algorithm significantly more often than men in Treatment 2 (six alternatives).

The subjects receive a payment of € 4.00 if a diagnosis is made that is also confirmed in the medium term. If a subject opts for one of the algorithms, the expected value of their remuneration is € 2.80 (0.7 · € 4.00). If a subject opts for one of the experts, the expected value of the remuneration is €2.40 (0.6 · €4.00). Accordingly, the homo economicus would therefore have to choose one of the algorithms. Every decision in favor of one of the doctors must be interpreted as a case of algorithm aversion.

The specific remuneration for each individual subject is determined through the use of a random number generator. In 60% of cases, decisions in favor of a doctor result in remuneration of €4.00, while in 70% of cases, decisions in favor of an algorithm result in remuneration of €4.00.

The experiment employs a between-subject design. Thus, each subject engages with only one of the two treatments. The subjects are randomly assigned to one of the two treatments. The experiment is programmed in z-Tree (cf. Fischbacher, 2007).

Results

The experimental survey was conducted at the Ostfalia Laboratory for Experimental Economic Research (OLEW) in Wolfsburg between April 24 and May 8, 2024. The survey was completed by 200 students from Ostfalia University of Applied Sciences. A total of 100 subjects completed Treatment 1 and 100 subjects completed Treatment 2.

Of the total number of subjects, 87 (43.5%) identified as female and 113 subjects (56.5%) identified as male. A total of 134 subjects (67.0%) are affiliated with the Faculty of Business, 55 (27.5%) with the Faculty of Automotive Engineering and 11 (5.5%) with other faculties. The majority of the subjects (n = 194, 97.0%) are enrolled in a Bachelor's degree program, while a smaller proportion (n = 6, 3.0%) are pursuing a Master's degree. The subjects have, on average, already studied for 4.22 semesters and have an average age of 22.3 years.

38.5% of the subjects choose one of the doctors, although this is associated with a lower probability of success and a lower expected value of the remuneration. This demonstrates that 38.5% of the subjects exhibit the behavioral anomaly of algorithm aversion (Table 1).

Table 1: Decisions of the 200 subjects in favor of an expert or an algorithm depending on treatments

	Decisions in favor of		
	Expert	Algorithm	
Treatment 1 (Two alternatives)	33 (33.0%)	67 (67.0%)	100
Treatment 2 (Six alternatives)	44 (44.0%)	56 (56.0%)	100
Total	77 (38.5%)	123 (61.5%)	200

Indeed, there is a difference between the two treatments with regard to the frequency of algorithm aversion. However, this distinction is not aligned with the predictions derived from the choice overload hypothesis. Sela, Berger & Liu (2009) have previously highlighted that an increase in the number of available alternatives tends to result in a greater inclination towards expediency. The ambiguity associated with a multitude of options frequently results in the selection of a readily justifiable, expedient alternative.

In light of this study's experimental design, which positions all algorithms as inherently superior to all experts, the most readily justifiable and appropriate course of action would be to select one of the algorithms. It was thus anticipated that one algorithm would be selected with greater frequency in Treatment 2 (six alternatives) than in Treatment 1 (two alternatives).

However, the experimental survey yields the opposite result: While the expert is preferred in only 33.0% of the cases in Treatment 1 (two alternatives), an expert is preferred in even 44.0% of the cases in Treatment 2 (six alternatives) (Table 1). The frequency of algorithm aversion does not decline with an increase in the number of alternatives; it actually rises by 11.0 percentage points. However, this difference is not statistically significant in the chi-square test (p-value = 0.1099).

Consequently, the null hypothesis 1 cannot be rejected. The assumed correlation between choice overload on the one hand and algorithm aversion on the other cannot be proven in our experiment. Therefore, it cannot be assumed that a reduction in algorithm aversion can be achieved by increasing the number of alternatives.

When looking at any differences between genders, however, a surprising result emerges (Tables 2 and 3).

Table 2: Decisions of the 113 male subjects in favor of an expert or an algorithm depending on treatments

	Decisions in favor of		
	Expert	Algorithm	
Treatment 1 (Two alternatives)	20 (35.7%)	36 (64.3%)	56
Treatment 2 (Six alternatives)	21 (36.8%)	36 (63.2%)	57
Total	41 (36.3%)	72 (63.7%)	113

The male subjects were completely unimpressed by the different number of alternatives in Treatment 1 (two alternatives) and Treatment 2 (six alternatives) (see Table 2). Approximately one-third of the male subjects demonstrated algorithm aversion, with rates of 35.7% in Treatment 1 and 36.8% in Treatment 2. Apparently, there is no recognizable correlation between choice overload on the one hand and algorithm aversion on the other.

Table 3: Decisions of the 87 female subjects in favor of an expert or an algorithm depending on treatments

	Decisions in favor of		
	Expert	Algorithm	
Treatment 1 (Two alternatives)	13 (29.5%)	31 (70.5%)	44
Treatment 2 (Six alternatives)	23 (53.5%)	20 (46.5%)	43
Total	36 (41.4%)	51 (58.6%)	87

The situation differs for the female subjects (see Table 3). Their behavior is strongly influenced by the number of alternatives. However, the correlation is not as initially assumed. In Treatment 1 (two alternatives), 29.5% of the female subjects chose the doctor, thereby exhibiting algorithm aversion. In Treatment 2 (six alternatives), 53.5% of the female subjects selected a doctor and thus exhibited algorithm aversion. The discrepancy between treatments 1 and 2 is statistically significant, with a probability of error of less than 5% (p-value = 0.0234).

The greater number of alternatives in Treatment 2 does not result in a lower degree of algorithm aversion among female subjects, but in a significantly greater degree of algorithm aversion.

Therefore, the null hypothesis 2 cannot be rejected. The assumption that choice overload has a significantly inhibitory effect on the incidence of algorithm aversion, particularly among female subjects, was not substantiated. Indeed, the opposite is true. The greater number of alternatives in Treatment 2 did not mitigate the tendency towards algorithm aversion in female subjects, but actually increased it.

In light of these unexpected outcomes, it is worthwhile to examine the subjects' behavior in greater detail. In particular, it should be discussed whether the six alternatives presented in Treatment 2 are sufficient in comparison to the two alternatives presented in Treatment 1 to cause a significant difficulty in decision-making. In the present experiment, the time required by the subjects to make their decision was recorded. The time survey commences only after the instructions for the game have been read, the decision alternatives have been carefully considered, and the control questions have been answered. The time survey concludes with the selection of an alternative and subsequent confirmation via clicking the "OK" button. This marks the end of the decision-making process, after which the subjects proceed to answer demographic questions. The recorded time span thus reflects the duration of the decision-making process.

Both women and men require a significantly longer time to make a decision in Treatment 2 than in Treatment 1 (Table 4).

Table 4: Average duration of decision making in seconds (standard deviation) depending on treatment and gender

	Women	Men
Treatment 1 (Two alternatives)	15.50 (12.06)	12.29 (7.51)
Treatment 2 (Six alternatives)	33.07 (21.24)	23.19 (14.49)

Women need an average of 15.50 seconds to decide between two alternatives and an average of 33.07 seconds to decide between six alternatives. This discrepancy is statistically significant (p-value < 0.001) when evaluated using the Mann-Whitney U test. The mean time required for men to make a decision between two alternatives is 12.29 seconds, while the mean time required for them to make a decision between six alternatives is 23.19 seconds. This difference also yields a highly significant result when subjected to the Mann-Whitney U-test (p-value < 0.001). It can thus be concluded that the decision in Treatment 2 (six alternatives) was more challenging for both women and men. These findings align with those of previous studies on choice overload.

It is of particular interest to note the differences between women and men in the two treatments. In Treatment 1, the discrepancy in the mean duration of decision-making (15.50 seconds for women and 12.29 seconds for men) was not statistically significant (p-value = 0.179) when assessed using the Mann-Whitney U test. In contrast, the data from Treatment 2 demonstrate a different pattern. The mean time required for women to select from the six alternatives is 33.07 seconds. The average time required for men to make the same decision is only 23.19 seconds. The observed difference is statistically significant at the 5% level (p-value = 0.012) when assessed using the Mann-Whitney U-test. This finding is consistent with previous research on choice overload, which suggests that women are more susceptible to an increase in the number of alternatives (see, for example, Misuraca et al., 2024; Ji, Ha & Sypher, 2014; Williamson & Eaker, 2012).

However, this does not provide an explanation as to why only the group of female subjects exhibited a significantly greater tendency towards algorithm aversion in Treatment 2 (six alternatives) than in Treatment 1 (two alternatives). Further research efforts in the border area of algorithm aversion and choice overload are needed to clarify the background of this surprising empirical finding.

In conclusion, it can be stated that an increase in the number of decision alternatives does not contribute to a reduction in algorithm aversion. For men, an increase in the number of decision alternatives has no discernible influence on their tendency towards algorithm aversion. However, for women, an increase in the number of decision alternatives actually leads to an increase in the tendency towards algorithm aversion.

Summary

Algorithm aversion represents a damaging behavioral anomaly, as it contributes to a reduction in the probability of success. Consequently, extensive research is being undertaken to identify methods for reducing algorithm aversion. A controlled laboratory experiment is being conducted to investigate whether increasing the number of alternatives can mitigate algorithm aversion. Prior research on choice overload has demonstrated that economic agents tend to make easily justifiable decisions when they must choose from a large number of alternatives. Consequently, when faced with a large choice, they are more likely to select an alternative that is obviously expedient.

In this vignette study, the subjects are instructed to assume the role of a health advisor who is tasked with making a recommendation to a patient. In Treatment 1, subjects are presented with a choice between a doctor and an algorithm for making a diagnosis. The algorithm is associated with a 10 percentage point higher probability of success than the doctor. In Treatment 2, subjects are presented with three doctors with different approaches and three algorithms, each with a different focus. The success rate of all algorithms is 10 percentage points higher than the success rates of all doctors.

If a greater number of options increases the likelihood of an easily justifiable decision being made and expediency becoming the primary focus of the decision-making process, then a lower degree of algorithm aversion should be observed in Treatment 2 than in Treatment 1.

However, the opposite is in fact the case. In Treatment 2 (six alternatives), the superior and therefore expedient algorithms are used less frequently than the one in Treatment 1 (two alternatives). However, the difference between the two treatments does not prove to be statistically significant.

The results are somewhat different when the analysis is restricted to the female subjects. There is evidence from previous studies that women are even more affected by choice overload than men. If choice overload has a mitigating effect on algorithm aversion, this effect should occur particularly in female subjects. However, the opposite is true. In Treatment 2 (six alternatives), the female subjects were statistically significantly more likely to exhibit algorithm aversion than in Treatment 1 (two alternatives).

Finally, it has to be concluded that an increase in the number of decision alternatives (for example, more experts or more algorithms) is not expected to mitigate algorithm aversion.

Acknowledgement

We express our gratitude to Peter von Holten and Laura Narjes for their valuable feedback and suggestions.

Literature

Bankuoru Egala, S. & Liang, D., 2023. Algorithm aversion to mobile clinical decision support among clinicians: a choice-based conjoint analysis. *European Journal of Information Systems*, 1–17.

Berger, B., Adam, M., Rühr, A., Benlian, A., 2021. Watch Me Improve—Algorithm Aversion and Demonstrating the Ability to Learn. *Business & Information Systems Engineering*, 63, 55–68.

Bogert, E., Schecter, A. & Watson, R.T., 2021. Humans rely more on algorithms than social influence as a task becomes more difficult. *Nature Scientific Reports,* 11, 8028.

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

Chernev, A., Böckenholt, U. & Goodman, J., 2015. Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, 25(2), 333–358.

Dawes, R., Faust, D. & Meehl, P., 1989. Clinical versus actuarial judgment, *Science*, 243(4899), 1668–1674.

Diehl, K. & Poynor, C., 2010. Great expectations?! Assortment size, expectations, and satisfaction. *Journal of Marketing Research*, 47(2), 312–322.

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.

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.

Efendic, E., Van de Calseyde, P. P. & Evans, A. M., 2020. Slow response times undermine trust in algorithmic (but not human) predictions. *Organizational Behavior and Human Decision Processes*, 157, 103–114.

Fasolo, B., McClelland, G. H. & Todd, P. M., 2007. Escaping the tyranny of choice: When fewer attributes make choice easier. *Marketing Theory*, 7(1), 13–26.

Filiz, I., Judek, J. R., Lorenz, M. & Spiwoks, M., 2021. Reducing algorithm aversion through experience, *Journal of Behavioral and Experimental Finance*, 31, 1–8.

Filiz, I., Judek, J. R., Lorenz, M. & Spiwoks, M., 2023. The extent of algorithm aversion in decision-making situations with varying gravity, *PLoS ONE*, 18(2), 1–21.

Filiz, I., Kirchhoff, F., Nahmer T. & Spiwoks, M., 2024. When it really matters: Algorithm aversion occurs most often when it is most harmful. Wolfsburg Working Papers, No. 24-02.

Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics*, 10(2), 171–178.

Gourville, J. T. & Soman, D., 2005. Overchoice and assortment type: When and why variety backfires. *Marketing Science*, 24(3), 382–395.

Gubaydullina, Z., Judek, J. R., Lorenz, M. & Spiwoks, M., 2022. Comparing Different Kinds of Influence on an Algorithm in Its Forecasting Process and Their Impact on Algorithm aversion. *Businesses*, 2(4), 448–470.

Iyengar, S. & Lepper, M., 2000. When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6), 995–1006.

Jacob, B. M., Thomas, S. & Joseph, J., 2024. Over two decades of research on choice overload: An overview and research agenda. *International Journal of Consumer Studies*, 48(2), 1–30.

Ji, Q., Ha, L. & Sypher, U., 2014. The role of news media use and demographic characteristics in the prediction of information overload. *International Journal of Communication*, 8, 699–714.

Judek, J. R., 2024. Willingness to Use Algorithms Varies with Social Information on Weak vs. Strong Adoption: An Experimental Study on Algorithm Aversion. *FinTech*, 3(1), 55–65.

Jung, M. & Seiter, M., 2021. Towards a better understanding on mitigating algorithm aversion in forecasting: an experimental study. *Journal of Management Control*, 32, 495–516.

Jussupow, E., Benbasat, I. & Heinzl, A., 2020. Why Are We Averse Towards Algorithms? A Comprehensive Literature Review on Algorithm Aversion. Research Paper Series, ECIS Proceedings, AIS Electronic Library (AISeL).

Kempf, S. D., Laczniak, N. R. & Smith, R. E., 2006. The effects of gender on processing advertising and product trial information. *Marketing Letters*, 17(1), 5–16.

Kim, J., Giroux, M. & Lee, J. C., 2021. When do you trust AI? The effect of number presentation detail on consumer trust and acceptance of AI recommendations. *Psychology & Marketing*, 38(7), 1140–1155.

Leffrang, D., Bösch, K. & Müller, O., 2023. Do People Recover from Algorithm Aversion? An Experimental Study of Algorithm Aversion over Time. Proceedings of the 56th Hawaii International Conference on System Sciences, 4016–4025.

Longoni, C., Bonezzi, A. & Morewedge, C. K., 2019. Resistance to Medical Artificial Intelligence Get access Arrow. *Journal of Consumer Research*, 46(4), 629–650.

Mellers, B. A., Lu, L. & McCoy, J. P., 2023. Predicting the future with humans and AI. *Consumer Psychology Review*, 6(1), 109–120.

Mick, D. G., Broniarczyk, S. M. & Haidt, J., 2004. Choose, choose, choose, choose, choose, choose, choose: Emerging and prospective research on the deleterious effects of living in consumer hyperchoice. *Journal of Business Ethics*, 52(2), 207–211.

Misuraca, R., Nixon, A. E., Miceli, S., Di Stefano, G. & Scaffidi Abbate, C., 2024. On the advantages and disadvantages of choice: future research directions in choice overload and its moderators. *Frontiers in Psychology*, 15.

Mogilner, C., Rudnick, T. & Iyengar, S. S., 2008. The mere categorization effect: How the presence of categories increases Choosers' perceptions of assortment variety and outcome satisfaction. *Journal of Consumer Research*, 35(2), 202–215.

Park, J.-Y., & Jang, S., 2013. Confused by too many choices? Choice overload in tourism. *Tourism Management*, 35, 1–12.

Reich, T., Kaju, A., & Maglio, S. J., 2023. How to overcome algorithm aversion: Learning from mistakes. *Journal of Consumer Psychology*, 33(2), 285–302.

Robertson, C., Woods, A., Bergstrand, K., Findley, J., Balser, C. & Slepian, M. J., 2023. Diverse patients' attitudes towards Artificial Intelligence (AI) in diagnosis. *PLOS Digital Health*, 2(5), 1–16.

Scheibehenne, B., Greifeneder, R. & Todd, P. M., 2009. What moderates the too-much-choice effect? *Psychology & Marketing*, 26(3), 229–253.

Scheibehenne, B., Greifeneder, R., & Todd, P. M., 2010. Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research*, 37(3), 409–425.

Schwartz, B., 2000. Self-determination: The tyranny of freedom. *American Psychologist*, 55(1), 79–88.

Schwartz, B., 2004. The paradox of choice: Why more is less. Ecco books, HarperCollins Publishers, New York.

Sela, A., Berger, J., & Liu, W., 2009. Variety, vice, and virtue: How assortment size influences option choice. *Journal of Consumer Research*, 35(6), 941–951.

Sunstein, C. R. & Gaffe, J., 2024. An Anatomy of Algorithm Aversion. SSRN-Paper, No. 4865492.

Thai, N. T., & Yuksel, U., 2017. Too many destinations to visit: Tourists' dilemma? *Annals of Tourism Research,* 62, 38–53.

Watson, D. E., 2024. Through the Looking Glass: Overcoming Algorithm Aversion in Accounting. USF Tampa Graduate Theses and Dissertations.

Williamson, J. & Eaker, P. E., 2012. The information overload scale. Paper presented at The Association for Information Science & Technology (ASIST), Baltimore, MD.

Xue, C., Kowshik, S. S., Lteif, D., Puducheri, S., Jasodanand, V. H., Zhou, O.T., Walia, A. S., Guney, O. B., Zhang, J. D., Pham, S. T., Kaliaev, A., Andreu-Arasa, V. C., Dwyer, B. C., Farris, C. W., Hao, H., Kedar, S., Mian, A. Z., Murman, D. L., O'Shea, S. A., Paul, A. B., Rohatgi, S., Saint-Hilaire, M.-H., Sartor, E. A., Setty, B. N., Small, J. E., Swaminathan, A., Taraschenko, O., Yuan, J., Zhou, Y., Zhu, S., Karjadi, C., Ang, T. F. A., Bargal, S. A., Plummer, B. A., Poston, K. L., Ahangaran, M., Au, R. & Kolachalama, V. B., 2024. AI-based differential diagnosis of dementia etiologies on multimodal data. *Nature Medicine*, 10.1038/s41591-024-03118-z, advance online publication, https://doi.org/10.1038/s41591-024-03118-z.

Yun, J. H., Eun-Ju, L. & Dong, H. K., 2021. Behavioral and neural evidence on consumer responses to human doctors and medical artificial intelligence. *Psychology and Marketing*, 38(1), 1–16.

Zhao, Y., Xu, L., Yu, F. & Jin, W., 2024. Perceived opacity leads to algorithm aversion in the workplace. *Acta Psychologica Sinica*, 56(4), 497–514.

Appendix 1: Task description, control questions and decision situation in Treatment 1

Task description

You are advising a patient who had to have a malignant melanoma (skin cancer with a high metastasis rate) removed. Numerous examinations were carried out to rule out the possibility that the skin cancer had spread throughout the body. In particular, the entire body was scanned with a magnetic resonance imaging (MRI) scanner.

There are two alternatives to choose from when evaluating the MRI scans and other examination results:

a) An experienced doctor who is both an oncologist (doctor specializing in cancer treatment) and a radiologist (doctor specializing in imaging techniques) evaluates all the results. In 60% of cases, the doctor obtains an examination result that is also confirmed in the medium term.
b) A specialized computer program evaluates all examination results. It is designed in particular for pattern recognition in MRI scans. In 70% of cases, the computer program achieves an examination result that is also confirmed in the medium term.

Your task is now to recommend one of these two alternatives to the patient. It is assumed that the patient will follow your recommendation.

Procedure: After reading these instructions and answering the control questions, you will be presented with the decision situation. In this situation, you must select one of the two decision options.

Payment: For participating in this task, you will receive a payment depending on the decision you make and a random principle based on the above-mentioned probabilities of occurrence. If you succeed in making a diagnosis that is confirmed in the medium term, you will receive €4.00. If this is not the case, you will not receive any remuneration.

Control questions

1. What is a malignant melanoma?

- ☐ A skin cancer with a high metastasis rate (correct)
- ☐ A radiological diagnostic procedure
- ☐ A pathological tendency to melancholy

2. Which alternatives are available here to achieve an examination result that is also confirmed in the medium term?

- ☐ One is an experienced doctor who is both an oncologist and a radiologist, and the other is a specialized computer program. (correct)
- ☐ Two experienced doctors on the one hand and four different specialized computer programs on the other.
- ☐ A specialized computer program on the one hand and a fortune teller on the other.

3. How often has it been possible in the past to achieve an examination result with the help of the specialized computer program that is also confirmed in the medium term?

- ☐ In 60% of cases
- ☐ In 70% of cases (correct)
- ☐ In 80% of cases

4. How high is your remuneration if you succeed in making a diagnosis that is confirmed in the medium term?

- ☐ 2.00 €
- ☐ 3.00 €
- ☐ 4.00 € (correct)

Decision situation

Now make your choice as to which option you recommend to the patient!

- ☐ Option A: I recommend having the diagnosis made by an experienced doctor who is both an oncologist and a radiologist.
- ☐ Option B: I recommend having the diagnosis made by the specialized computer program, which is designed in particular for pattern recognition in MRI scans.

Appendix 2: Task description, control questions and decision situation in Treatment 2

Task description

You are advising a patient who had to have a malignant melanoma (skin cancer with a high metastasis rate) removed. Numerous examinations were carried out to rule out the possibility that the skin cancer had spread throughout the body. In particular, the entire body was scanned with a magnetic resonance imaging (MRI) scanner.

There are six alternatives to choose from when evaluating the MRI scans and other examination results:

a) An experienced doctor who is both an oncologist (doctor specializing in cancer treatment) and a radiologist (doctor specializing in imaging techniques) evaluates all the results. In 60% of cases, the doctor obtains an examination result that is also confirmed in the medium term.
b) An experienced doctor who is both an oncologist (doctor specializing in cancer treatment) and a radiologist (doctor specializing in imaging techniques) evaluates all results together with a colleague. The dual control principle is intended to avoid careless errors. In 60% of cases, the medical team achieves an examination result that is also confirmed in the medium term.
c) An experienced doctor, who is both an oncologist (doctor specializing in cancer treatment) and a radiologist (doctor specializing in imaging techniques), evaluates all results. In doing so, he strictly follows a diagnostic plan that has been drawn up by an international medical commission. The diagnostic plan is intended to prevent decisions being made on the basis of incomplete data. In 60% of cases, the doctor achieves an examination result that is also confirmed in the medium term.
d) A specialized computer program evaluates all examination results. It is designed in particular for pattern recognition in MRI scans. In 70% of cases, the computer program achieves an examination result that is also confirmed in the medium term.
e) A specialized computer program evaluates all examination results, but places particular emphasis on the inclusion of electron and fluorescence microscopy results. In 70% of cases, the computer program achieves an examination result that is also confirmed in the medium term.
f) A specialized computer program evaluates all examination results, but places particular emphasis on using neural networks to better evaluate the raw data space (k-space) in order to avoid image noise and artefacts. In 70% of cases, the computer program achieves an examination result that is also confirmed in the medium term.

Your task is now to recommend one of these six alternatives to the patient. It is assumed that the patient will follow your recommendation.

Procedure: After reading these instructions and answering the control questions, you will be presented with the decision situation. In this situation, you must select one of the six decision options.

Payment: For participating in this task, you will receive a payment depending on the decision you make and a random principle based on the above-mentioned probabilities of occurrence. If you succeed in making a diagnosis that is confirmed in the medium term, you will receive €4.00. If this is not the case, you will not receive any remuneration.

Control questions

1. What is a malignant melanoma?

   - ☐ A skin cancer with a high metastasis rate (correct)
   - ☐ A radiological diagnostic procedure
   - ☐ A pathological tendency to melancholy

2. Which alternatives are available here to achieve an examination result that is also confirmed in the medium term?

   - ☐ There are three different doctors or teams of doctors to choose from, each with different specializations, as well as three different specialized computer programs with different focuses. (correct)
   - ☐ There is an experienced doctor and five different specialized computer programs with different focuses to choose from.
   - ☐ There are two experienced doctors to choose from, each with different specializations.

3. How often has it been possible in the past to achieve an examination result with the help of one of the specialized computer programs that is also confirmed in the medium term?

   - ☐ In 60% of cases
   - ☐ In 70% of cases (correct)
   - ☐ In 80% of cases

4. How high is your remuneration if you succeed in making a diagnosis that is confirmed in the medium term?

   - ☐ 2.00 €
   - ☐ 3.00 €
   - ☐ 4.00 € (correct)

Decision situation

Now make your choice as to which option you recommend to the patient!

- ☐ Option A: I recommend having the diagnosis made by an experienced doctor who is both an oncologist and a radiologist.
- ☐ Option B: I recommend having the diagnosis made by an experienced doctor who is both an oncologist and a radiologist and who, with the help of a colleague, relies on the dual control principle.
- ☐ Option C: I recommend having the diagnosis made by an experienced doctor who is both an oncologist and a radiologist and who strictly follows a diagnostic plan.
- ☐ Option D: I recommend having the diagnosis made by the specialized computer program, which is designed in particular for pattern recognition in MRI scans.

- ☐ Option E: I recommend having the diagnosis made by the specialized computer program, which places particular emphasis on the inclusion of electron and fluorescence microscopy results.
- ☐ Option F: I recommend having the diagnosis made by the specialized computer program, which places particular emphasis on better evaluating the raw data space (k-space) with the help of neural networks.