Ostfalia
University of
Applied Sciences
Wolfsburg
Faculty of Business

WWP

Wolfsburg Working Papers

No. 24-02

When it really matters:

Algorithm aversion occurs most often when it is most harmful

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

When it really matters

Algorithm aversion occurs most often when it is most harmful

Ibrahim Filiz, Florian Kirchhoff, Thomas Nahmer and Markus Spiwoks

JEL codes: D81, D91, G41, O33.

Key Words: Algorithm aversion, decision-making under risk, framing, behavioral economics, experiments.

Abstract:

A laboratory experiment is used to test whether algorithm aversion occurs particularly in decision-making situations where serious consequences are at stake. It is shown that the willingness to use an algorithm that is recognizably more powerful than a human expert decreases when the decision is particularly important.

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, Tel.: +49 5361 8922 25420, e-mail: fl.kirchhoff@ostfalia.de

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

Some economic entities exhibit a negative attitude towards algorithms. They tend to delegate tasks to human experts or perform them themselves. This is often the case even when it is evident that the use of algorithms would improve the quality of results compared to human experts. This behavioral anomaly is referred to as algorithm aversion (Dietvorst, Simmons & Massey, 2015).

Filiz, Judek, Lorenz & Spiwoks (2023) conclude that the tendency towards algorithm aversion is especially prevalent when there is a threat of particularly serious consequences for an upcoming decision. They contrast three decision-making situations with less serious potential consequences with three decision-making situations with more serious potential consequences.

The decisions with less serious potential consequences are: (1) finding a partner on a dating platform with the help of a specialized computer program (algorithm) or with the help of psychologically trained employees, (2) selecting recipes for cooking boxes to be delivered with the help of a specialized computer program (algorithm) or with the help of culinary-trained employees and (3) making weather forecasts with the help of a specialized computer program (algorithm) or with the help of experienced meteorologists. Even if these tasks are not performed satisfactorily, the consequences are not particularly severe. A date may turn out to be boring, the taste of a lunch may be disappointing or one may find oneself out in the rain without a coat. While these circumstances may be unpleasant, they are relatively easy to overcome.

In contrast, the decisions with more serious potential consequences are: (1) driving services with the help of autonomous vehicles (algorithm) or with the help of chauffeurs, (2) evaluation of MRI scans with the help of a specialized computer program (algorithm) or with the help of medical professionals and (3) evaluations of criminal case files with the help of a specialized computer program (algorithm) or with the help of lawyers. Driving services and the evaluation of MRI scans may pose a risk to the individual's health. The evaluation of documents in the context of criminal proceedings may involve serious restrictions of personal freedom. Consequently, these three decision-making situations may have serious consequences if the outcome is unfavorable.

It has often been criticized that this study presents concrete decision-making situations that can trigger an unmanageable number of associations.[1] These associations may seriously distort the results of the study under certain circumstances. It is possible that the subjects do not primarily recognize the seriousness of the possible consequences. Instead, a general aversion to autonomous driving or dating platforms may drive the results.

This criticism is certainly justified. Therefore, the question arises as to whether the results of Filiz, Judek, Lorenz & Spiwoks (2023) are also confirmed when comparing two decision situations that do not differ in any way except for the severity of the possible consequences.

[1] This critique was presented at the Economic Science Association (ESA) Global Online Around-the-Clock Meeting on 07.07.2021, at the annual conference of the German Association for Experimental Economic Research (GfeW) e.V. in Magdeburg, Germany, on 23.09.2021 and by an anonymous reviewer at the journal PLoS ONE.

Hypotheses and experimental design

The objective of this experimental study is to ascertain whether the findings of Filiz, Judek, Lorenz and Spiwoks (2023) are also confirmed when two decision scenarios are presented that differ solely in the severity of the potential consequences.

The subjects are asked to decide for a company whether a complex business decision should be based on the advice of an experienced management consultant (expert) or on the recommendation of a specialized computer program (algorithm). They are informed that the advice of the management consultant leads to successful decisions in 60% of the cases, while the recommendation of the computer program leads to successful decisions in 70% of the cases. Each subject receives a payment of €1 for a successful decision, regardless of whether it is based on the advice of the management consultant (expert) or the computer program (algorithm). A random generator determines whether a payment is made (with a 60% probability of success when choosing the expert and 70% probability of success when choosing the algorithm). Consequently, the expected value of the reward is 60 cents if the respondent follows the expert's recommendation and 70 cents if the respondent follows the algorithm's recommendation. A homo economicus would necessarily have to follow the algorithm's recommendation. However, if a respondent trusts the expert's advice, this must be regarded as algorithm aversion.

In treatment 1 (insignificant possible consequences), it is pointed out that a successful decision will only result in minimal positive consequences for the company and an unsuccessful decision will lead to minimal negative consequences. In treatment 2 (serious possible consequences), it is pointed out that a successful decision leads to strong positive consequences and an unsuccessful decision will result in dramatic negative consequences for the company. The wording of the two treatments can be found in Appendix 1. The experiment is based on a between-subjects design.

We expect the results of Filiz, Judek, Lorenz and Spiwoks (2023) to be confirmed even under these altered conditions.

Hypothesis 1 is therefore: The frequency of algorithm aversion is significantly higher in treatment 2 than in treatment 1.

Null hypothesis 1 therefore reads: The frequency of algorithm aversion is not significantly higher in treatment 2 than in treatment 1.

Results

The experimental survey was conducted at the Ostfalia Laboratory for Experimental Economic Research (OLEW) in Wolfsburg from April 24 to May 8, 2024. A total of 150 students from Ostfalia University of Applied Sciences participated in the survey. Of these, 75 subjects completed treatment 1 and 75 subjects completed treatment 2.

62 subjects (41.3%) are female and 88 subjects (58.7%) are male. A total of 99 subjects (66.0%) belong to the Faculty of Business, 41 (27.3%) to the Faculty of Automotive Engineering and 10 (6.7%) to other faculties. 145 subjects (96.7%) are Bachelor students and 5 (3.3%) are Master students. On average, the subjects have already studied for 4.56 semesters and have an average age of 22.5 years.

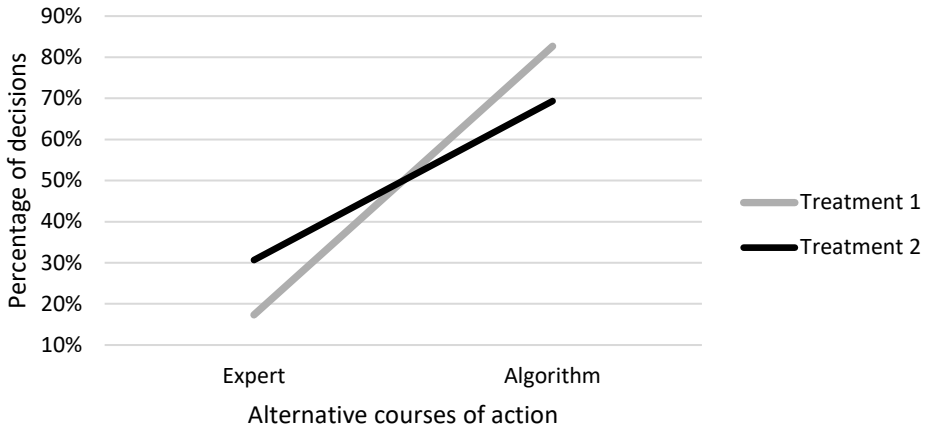
It can be seen that the behavioral anomaly of algorithm aversion occurs in only about a quarter of the subjects (24.0%) (see Table 1). The straightforward and transparent nature of the task, coupled with

the evident likelihood of success associated with the expert advice (60%) and the algorithmic recommendation (70%), may have played a pivotal role in this outcome.

Table 1: Decisions of the 150 subjects in favor of the expert or the algorithm by treatment

	Decision for		
	Expert	Algorithm	
Treatment 1 (insignificant consequences)	13 (17.33%)	62 (82.67%)	75
Treatment 2 (serious consequences)	23 (30.67%)	52 (69.33%)	75
Total	36 (24.00%)	114 (76.00%)	150

Figure 1: Decisions of the 150 subjects in favor of the expert or the algorithm by treatment



A comparison of the treatments (Table 1 and Figure 1) reveals that the expert is preferred more often in the case of serious consequences in treatment 2 (30.67%) than in the case of minor consequences in treatment 1 (17.33%). This difference is statistically significant, with a probability of error of <10% (p-value = 0.056), as demonstrated by the Pearson chi-square test.

The 99 business students, who may be more familiar with the business decision scenario than other students, demonstrate an even more pronounced result (Table 2). In the case of serious consequences in treatment 2, the expert is preferred considerably more often (36.17%) than in the case of insignificant consequences (11.54%) in treatment 1. The Pearson chi-square test indicates that this difference is statistically significant with a probability of error of <1% (p-value = 0.008). Consequently, null hypothesis 1 must be rejected.

Table 2: Decisions of the 99 business students in favor of the expert or the algorithm by treatment

	Decision for		
	Expert	Algorithm	
Treatment 1 (insignificant consequences)	6 (11.54%)	46 (88.46%)	52
Treatment 2 (serious consequences)	17 (36.17%)	30 (63.83%)	47
Total	23 (23.23%)	76 (76.77%)	99

It can be concluded that the "tragedy of algorithm aversion" described by Filiz, Judek, Lorenz & Spiwoks (2023) is indeed true: The behavioral anomaly of algorithm aversion is particularly prevalent at the very time when it really matters to make the most promising decision possible. This leads to a reduction in the chances of success, especially when making important decisions. Thus, algorithm aversion is an extremely harmful phenomenon. Any research that explores strategies to mitigate algorithm aversion is therefore worthwhile.

Summary

A laboratory experiment is used to test whether algorithm aversion occurs more often when there is a threat of serious consequences in the course of a decision-making situation. Filiz, Judek, Lorenz & Spiwoks (2023) come to this conclusion. However, their study is criticized for a certain methodological vagueness. The decision-making situations presented there do not differ exclusively in the range of possible consequences. This shortcoming is remedied in the present study.

It turns out that algorithm aversion actually occurs significantly more often when serious consequences are threatening. However, superior algorithms should be used, especially when important decisions have to be made. The fact that this is not the case can certainly be described as the "tragedy of algorithm aversion".

References

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.

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

Appendix 1: Game instructions for treatment 1 and treatment 2

Imagine you work in the management of a large company.

A complex decision is pending in your company, in which many influencing factors and many framework conditions must be taken into account. The decision is also made more difficult by the fact that random influences play a role.

Since the necessary expertise is not available within the company, advice should be sought from outside the company and the decision made in accordance with this advice. One option is an experienced and highly respected management consultant who has specialized in this type of decision for many years. Alternatively, you could use a specialized computer program that is designed for this type of decision.

Given the complexity of the decision and the importance of random influences, neither the help of the management consultant nor the computer program is associated with a 100% success rate. The experienced management consultant successfully handles such decisions in 60 out of 100 cases. The computer program successfully handles such decisions in 70 out of 100 cases.

> *Specific text module in Treatment 1*
>
> Although the decision-making situation is very complicated, the decision is not of great importance to the company. If a successful decision is made, the company's turnover and profit will improve minimally. If an unsuccessful decision is made, the company's turnover and profit will deteriorate minimally.
>
> *Specific text module in Treatment 2*
>
> The decision-making situation is very complicated and of very serious importance for the company. If a successful decision is made, the company's turnover and profit will improve significantly. If an unsuccessful decision is made, then the company's turnover and profit will deteriorate dramatically and even the continued existence of the company is threatened.

Your task now is to choose whether the business decision should be based on the advice of the management consultant or on the advice of the computer program.

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

For your participation in this task, you will receive a payout depending on your decision and a random principle based on the probabilities of occurrence mentioned above. If your decision is successful, you will receive a reward of €1.00. If your decision is unsuccessful, you will not receive any reward.

Appendix 2: Control questions

1. Who can you entrust with the preparation of the business decision?
   A) Specialized computer program or renowned fortune teller
   B) Specialized computer program or experienced management consultant (*correct*)
   C) Experienced management consultant or career starter in a management consultancy

2. What is the success rate of the specialized computer program?
   A) 65%
   B) 70% (*correct*)
   C) 75%
   D) 80%

3. What are the consequences for the company if an unsuccessful decision is made?
   A) There is a risk of a minimal decline in turnover and profit (*correct in treatment 1*)
   B) There is a risk of a dramatic decline in turnover and profits and even the demise of the company is possible (*correct in treatment 2*)

4. How high will your reward be if a successful decision is made?
   A) 0.50 €
   B) 1.00 € (*correct*)
   C) 3.00 €

Appendix 3: Decision situation

Now choose whose advice the business decision should be based on!

☐ I call in the experienced management consultant and make the decision based on his advice

☐ I use the specialized computer program and make the decision based on its advice