

Choice Overload als Gegengift zur Algorithmusaversion: Wirklos bei Männern und problemschärfend bei Frauen

Filiz, I.; Kirchhoff, F.; Nahmer, T.; Spiwoks, M.

sofia-Diskussionsbeitrag Nr. 24-4, Darmstadt 2024

ISBN 978-3-947850-09-9

Choice Overload als Gegengift zur Algorithmusaversion: Wirkungslos bei Männern und problemverschärfend bei Frauen

Ibrahim Filiz, Florian Kirchhoff, Thomas Nahmer und 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:

Welchen Einfluss hat die Anzahl der Handlungsalternativen auf das Ausmaß der Algorithmusaversion? Das ist die Forschungsfrage der vorliegenden Studie. Forschungsergebnisse im Bereich Choice Overload zeigen, dass eine Vielzahl von Alternativen häufig dazu führt, dass Wirtschaftssubjekte sich für eine leicht begründbare, zweckdienliche Alternative entscheiden. Choice Overload könnte somit die Neigung zur Algorithmusaversion dämpfen. Die Ergebnisse des vorliegenden Laborexperiments bestätigen diese Vermutung jedoch nicht. Während die Anzahl der Alternativen bei den männlichen Probanden keine Wirkung entfaltet, zeigt sich bei den weiblichen Probanden sogar der entgegengesetzte Effekt. Eine größere Zahl von Alternativen steigert bei Frauen die Neigung zur Algorithmusaversion signifikant.

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 892 225 420, 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: 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

Inhalt

1 Einführung	3
2 Literaturüberblick	4
3 Hypothesen und experimentelles Design	7
4 Ergebnisse	10
5 Zusammenfassung	14
Literatur	15
Anhänge	19

Tabellen

Tab. 1: Entscheidungen der 200 Probanden für einen Experten oder einen Algorithmus nach Treatments	10
Tab. 2: Entscheidungen der 113 männlichen Probanden für einen Experten oder einen Algorithmus nach Treatments	11
Tab. 3: Entscheidungen der 87 weiblichen Probanden für einen Experten oder einen Algorithmus nach Treatments	11
Tab. 4: Durchschnittliche Dauer der Entscheidungsfindung in Sekunden (Standardabweichung) nach Treatments und nach Geschlecht	12

1

Einführung

Für viele Aufgaben, die traditionell von Experten bewältigt werden, gibt es inzwischen Algorithmen, die diese Aufgaben genauso zuverlässig oder sogar zuverlässiger erledigen können. Sofern diese Algorithmen keine 100%ige Erfolgsquote aufweisen, bleiben viele Wirtschaftssubjekte gegenüber dem Einsatz dieser Algorithmen allerdings skeptisch. Dies ist häufig selbst dann der Fall, wenn klar erkennbar ist, dass die Erfolgsquote des Algorithmus deutlich höher liegt als die Erfolgsquote des Experten. Diese Verhaltensanomalie wird als Algorithmusaversion bezeichnet (Dietvorst, Simmons & Massey, 2015).

In den zurückliegenden zehn Jahren ist das Phänomen der Algorithmusaversion intensiv untersucht worden. Dabei spielen Laborexperimente eine herausgehobene Rolle. In den experimentellen Entscheidungssituationen haben die Probanden in aller Regel die Wahl, ob sie eine Aufgabe selbst bewältigen wollen oder ob sie die Aufgabe von einem Algorithmus bearbeiten lassen wollen. Oder die Probanden haben die Wahl, ob sie die Aufgabe an einen Experten oder an einen Algorithmus delegieren wollen.

Die empirische Konsumforschung zeigt jedoch, dass das Entscheidungsverhalten der Wirtschaftssubjekte unter anderem auch davon beeinflusst wird, wie viele Alternativen zur Verfügung stehen. Wenn Konsumenten zwischen einer großen Zahl von Produktalternativen wählen müssen, fällt die Entscheidung häufig bedeutend schwerer, als wenn nur eine kleine Auswahl zur Verfügung steht. Es zeigt sich sogar, dass eine Kaufentscheidung häufig unterbleibt, wenn die Auswahl zu groß ist. Dieses Phänomen wird als Choice Overload bezeichnet (vgl. Iyengar & Lepper, 2000; Scheibehenne, Greifeneder & Todd, 2010; Park & Jang, 2013; Chernev, Böckenholt & Goodman, 2015; Thai & Yuksel, 2017; Jacob, Thomas & Joseph, 2024).

Sela, Berger & Liu (2009) kommen bei einer Analyse der Choice-Overload-Anomalie zu einer interessanten Einschätzung: Wirtschaftssubjekte, die aus einer großen Zahl von Möglichkeiten eine Alternative auswählen müssen, treffen in der Regel eine Entscheidung, die leicht zu rechtfertigen ist.

Das könnte ein Anknüpfungspunkt sein, um zu einer Abschwächung der Algorithmusaversion beizutragen. Wenn nicht nur ein Algorithmus beziehungsweise ein Experte, sondern mehrere Algorithmen und mehrere Experten zur Auswahl stehen, könnte dies dazu führen, dass häufiger eine leicht zu rechtfertigende Entscheidung getroffen wird. Wenn alle Algorithmen allen Experten hinsichtlich ihrer Erfolgswahrscheinlichkeit erkennbar überlegen sind, wäre die Entscheidung zugunsten eines der Algorithmen leicht zu rechtfertigen. Stehen also mehrere Experten und mehrere überlegene Algorithmen zur Auswahl, könnte dies zu einer Abschwächung der Algorithmusaversion beitragen. Ob dies tatsächlich der Fall ist, wird im Rahmen der vorliegenden Studie untersucht.

2 Literaturüberblick

Algorithmusaversion

Automatisierte Entscheidungswege oder Entscheidungshilfen, so genannte Algorithmen, gewinnen zusehends an Bedeutung. Schon seit Jahrzehnten stehen leistungsfähige Algorithmen zur Verfügung (vgl. bspw. Dawes, Faust & Meehl, 1989). Einige Wirtschaftssubjekte zeigen jedoch eine ablehnende Haltung gegenüber Algorithmen, selbst wenn klar erkennbar ist, dass algorithmusbasierte Entscheidungen mit einer deutlich höheren Erfolgsquote verbunden sind als eigene Entscheidungen oder Expertenentscheidungen. Dieses Phänomen wird als Algorithmusaversion bezeichnet (Dietvorst, Simmons & Massey, 2015). Allerdings gibt es durchaus Unterschiede in den Begriffsdefinitionen. Filiz et al. (2023) sowie Jussupow, Benbasat & Heinzl (2020) heben dabei folgende Aspekte abweichender Begriffsverständnisse hervor: Erfahrbarkeit der Überlegenheit des Algorithmus (nur erahnbar oder klar erkennbar und/oder bezifferbar), Zuverlässigkeit des Algorithmus (perfekt oder unvollkommen), Beobachtung der Zuverlässigkeit (sichtbares Auftreten von Fehlern oder nicht), Häufigkeit der Hervorhebung der Erfolgsquoten (einmalig oder wiederholt), Verwendung historischer Daten (nur zur Erstellung oder auch im Betrieb des Algorithmus), Setting (Algorithmus vs. Experte, Algorithmus vs. Laie oder Algorithmus vs. handelndes Wirtschaftssubjekt) sowie Ausmaß des algorithmischen Eingriffs (Algorithmus fungiert nur als Entscheidungshilfe oder führt die Tätigkeit selbstständig aus).

Bei allen Unterschieden im Detail herrscht Einigkeit darüber, dass die Verhaltensanomalie Algorithmusaversion erhebliche ökonomische Nachteile verursachen kann. Angesichts der enormen praktischen Relevanz hat sich daher in den zurückliegenden zehn Jahren eine lebhaft wissenschaftliche Debatte über Ursachen der Algorithmusaversion (vgl. bspw. Sunstein & Gaffe, 2024), aber auch über Möglichkeiten zur Reduzierung der Algorithmusaversion, entsponnen. So stellen Dietvorst, Simmons & Massey (2018) fest, dass die Neigung zur Algorithmusaversion erheblich verringert werden kann, wenn die handelnden Akteure die Möglichkeit erhalten, das Ergebnis des Algorithmus zu verändern – und sei es auch nur ein wenig (vgl. auch Gubaydullina et al., 2022). Filiz et al. (2021) zeigen, dass eine wachsende Erfahrung mit einem Algorithmus durch regelmäßiges Feedback bezüglich der Erfolgsquoten zu einem allmählichen Rückgang der Algorithmusaversion führen kann (vgl. auch Leffrang, Bösch & Müller, 2023). Watson (2024) hebt die Bedeutung von Informationen über die Zuverlässigkeit des Algorithmus für die Reduzierung der Algorithmusaversion hervor. Berger et al. (2021) machen deutlich, dass Wirtschaftssubjekte, die Lernfortschritte eines Algorithmus zur Kenntnis nehmen, weniger zur Algorithmusaversion neigen (vgl. auch Reich, Kaju & Maglio, 2023). Bogert, Schechter & Watson (2021) stellen fest, dass Vorbehalte gegenüber dem Einsatz von Algorithmen umso geringer ausfallen, je komplexer und schwieriger die Aufgabenstellung

ist. Ein höherer Zeitdruck bei der Aufgabenbewältigung kann ebenso zu einer Reduzierung der Algorithmusaversion beitragen (vgl. Jung & Seiter, 2021). Castelo, Bos & Lehmann (2019) weisen darauf hin, dass eine Erhöhung der wahrgenommenen affektiven (emotionalen) Menschenähnlichkeit von Algorithmen geeignet ist, die Algorithmusaversion insbesondere bei Aufgaben, die offen für Interpretationen sind und auf persönlicher Meinung oder Intuition beruhen, zu reduzieren (vgl. auch Zhao et al., 2024). Judek (2024) stellt fest, dass das Wissen über eine hohe Akzeptanz bei anderen Nutzern zu einer Verringerung der Zurückhaltung bei der Nutzung eines Algorithmus beiträgt. Efendic, Van de Calseyde & Evans (2020) zeigen, dass die Vorbehalte gegenüber Algorithmen geringer ausfallen, wenn sie sehr schnell zu einem Ergebnis kommen. Kim, Giroux & Lee (2021) ermitteln, dass eine Ergebnisdarstellung mit vielen Nachkommastellen das Vertrauen in den Algorithmus stärkt und somit der Algorithmusaversion entgegenwirkt.

Choice Overload

Iyengar & Lepper (2000) zeigen in drei Feld- und Laborexperimenten, dass die Anzahl der Auswahlmöglichkeiten einen erheblichen Einfluss auf das Entscheidungsverhalten von Wirtschaftssubjekten entfalten kann. Erstaunlicherweise sinkt die Entscheidungsfreude ebenso wie die subjektive Zufriedenheit mit einer getroffenen Entscheidung, wenn eine größere Anzahl von Alternativen zur Verfügung steht. Sie prägen dafür den Begriff Choice Overload. Dieser Begriff wird auch von Mogilner, Rudnick & Iyengar (2008), Diehl & Poynor (2010), Scheibehenne, Greifeneder & Todd (2010), Park & Jang (2013), Chernev, Böckenholt & Goodman (2015), Thai & Yuksel (2017) sowie von Jacob, Thomas & Joseph (2024) verwendet. Aber in der Literatur treten auch leicht abweichende Bezeichnungen wie beispielsweise over choice effect (Gourville & Soman, 2005), problem of too much choice (Fasolo, McClelland & Todd, 2007), tyranny of choice (Schwartz, 2000), too-much-choice effect (Scheibehenne, Greifeneder & Todd, 2009), consumer hyperchoice (Mick, Broniarczyk & Haidt, 2004) oder paradox of choice (Schwartz, 2004) auf.

Es liegen inzwischen rund 100 Studien zum Thema Coice Overload vor (vgl. Jacob, Thomas & Joseph, 2024). Dabei wird deutlich, dass die empirischen Befunde erheblich mit den Rahmenbedingungen (bspw. Anzahl der Alternativen, Anzahl der Beurteilungskriterien der Alternativen, Informationsfülle, Zeitbudget für die mögliche Auswahlentscheidung) variieren können. Auch die möglichen Anknüpfungspunkte an schon bestehende psychologische und/oder ökonomische Theorien (bspw. 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) sind hoch umstritten (Jacob, Thomas & Joseph, 2024). Allerdings lässt

sich ein Ergebnis als gesichert hervorheben: Die Anzahl der Alternativen kann einen erheblichen Einfluss auf das Entscheidungsverhalten der Wirtschaftssubjekte haben.

Ein Forschungsergebnis ist von besonderem Interesse. Sela, Berger & Liu (2009) zeigen in fünf experimentellen Erhebungen, dass eine große Zahl von Alternativen tendenziell dazu beiträgt, dass häufiger leicht zu rechtfertigende Alternativen ausgewählt werden. Das erhöht die Häufigkeit, in der tugendhafte und zweckdienliche Alternativen gewählt werden. Rein genussstiftende, hedonistische Alternativen werden hingegen weniger häufig gewählt. Deshalb trägt eine Erhöhung der Anzahl an Alternativen tendenziell dazu bei, dass weniger lasterhafte und mehr tugendhafte Entscheidungen getroffen werden und dass der Zweckdienlichkeit mehr Aufmerksamkeit geschenkt wird als der Genussverheißung.

Darüber hinaus wird diskutiert, ob Choice Overload in unterschiedlicher Weise bei Frauen und Männern zum Tragen kommt. Eine stärkere Detailorientierung bei Frauen (Kempf, Laczniak & Smith, 2006) führt dazu, dass Choice Overload bei Frauen häufiger auftritt und eine größere Rolle spielt als bei Männern (vgl. bspw. Misuraca et al., 2024; Ji, Ha & Sypher, 2014; Williamson & Eaker, 2012).

3 Hypothesen und experimentelles Design

Wir fügen das Experiment in den Kontext einer medizinischen Entscheidung ein. Dafür gibt es zwei Gründe. Zum einen liegen gerade im medizinischen Bereich sehr leistungsstarke Algorithmen vor, denen viele Patienten und auch viele Ärzte überaus skeptisch gegenüberstehen (Longoni, Bonezzi & Morewedge, 2019; Yun, Eun-Ju & Dong, 2021; Robertson et al., 2023; Bankuoru Egala & Liang, 2023; Mellers, Lu & McCoy, 2023; Xue et al., 2024). Zum anderen haben Filiz et al. (2023) sowie Filiz et al. (2024) gezeigt, dass eine Algorithmusaversion besonders häufig auftritt, wenn bei einer Fehlentscheidung schwerwiegende Konsequenzen drohen. Das ist bei Entscheidungen im Gesundheitsbereich häufig der Fall.

Es handelt sich um eine Vignettenstudie, bei der die Probanden als Gesundheitsberater für einen Patienten auftreten. Dem Patienten ist ein malignes Melanom (Hautkrebs mit hoher Metastasierungsrate) entfernt worden. Um auszuschließen, dass von dem Hautkrebs Tochtergeschwülste im Körper gestreut wurden, wurde der gesamte Körper mit einem Magnetresonanztomographen (MRT) gescannt und eine Reihe weiterer Untersuchungen durchgeführt. Nun steht die Frage an, wem die Gesamtbeurteilung aller Untersuchungsergebnisse überlassen werden soll.

Im Treatment 1 stehen zwei Alternativen zur Auswahl:

- Ein erfahrener Arzt, der sowohl Onkologe (Arzt mit dem Schwerpunkt Krebsbehandlung) als auch Radiologe (Arzt mit dem Schwerpunkt bildgebende Verfahren) ist, wertet die Ergebnisse aus (Option A). Der Arzt erzielt in 60% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.
- Ein spezialisiertes Computerprogramm wertet alle Untersuchungsergebnisse aus. Es ist insbesondere auf die Mustererkennung bei MRT-Aufnahmen ausgelegt (Option B). Das Computerprogramm erzielt in 70% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.

Im Treatment 2 stehen sechs Alternativen zur Auswahl:

- Ein erfahrener Arzt, der sowohl Onkologe (Arzt mit dem Schwerpunkt Krebsbehandlung) als auch Radiologe (Arzt mit dem Schwerpunkt bildgebende Verfahren) ist, wertet alle Ergebnisse aus (Option A). Der Arzt erzielt in 60% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.
- Ein erfahrener Arzt, der sowohl Onkologe (Arzt mit dem Schwerpunkt Krebsbehandlung) als auch Radiologe (Arzt mit dem Schwerpunkt bildgebende Verfahren) ist, wertet alle Ergebnisse gemeinsam mit ei-

nem Kollegen aus. Das Vier-Augen-Prinzip soll Flüchtigkeitsfehler vermeiden (Option B). Das Ärzteteam erzielt in 60% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.

- Ein erfahrener Arzt, der sowohl Onkologe (Arzt mit dem Schwerpunkt Krebsbehandlung) als auch Radiologe (Arzt mit dem Schwerpunkt bildgebende Verfahren) ist, wertet alle Ergebnisse aus. Dabei folgt er strikt einem Diagnoseplan, der durch eine internationale Ärzte-Kommission erarbeitet worden ist. Der Diagnoseplan soll verhindern, dass Entscheidungen auf Basis einer unvollständigen Datenlage getroffen werden (Option C). Der Arzt erzielt in 60% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.
- Ein spezialisiertes Computerprogramm wertet alle Untersuchungsergebnisse aus. Es ist insbesondere auf die Mustererkennung bei MRT-Aufnahmen ausgelegt (Option D). Das Computerprogramm erzielt in 70% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.
- Ein spezialisiertes Computerprogramm wertet alle Untersuchungsergebnisse aus, legt dabei allerdings besonderes Gewicht auf die Einbeziehung von Ergebnissen der Elektronen- und Fluoreszenzmikroskopie (Option E). Das Computerprogramm erzielt in 70% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.
- Ein spezialisiertes Computerprogramm wertet alle Untersuchungsergebnisse aus, legt dabei allerdings besonderes Gewicht darauf, den Rohdatenraum (k-Raum) mit Hilfe von neuronalen Netzen besser auszuwerten, um Bildrauschen und Artefakte zu vermeiden (Option F). Das Computerprogramm erzielt in 70% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.

Treatment 1 spiegelt die typische Entscheidungssituation in experimentellen Studien zur Algorithmusaversion wider: Der Proband muss sich zwischen einem Experten und einem Algorithmus entscheiden. Treatment 2 bietet viel mehr Auswahl: Der Proband muss sich für einen von drei Experten oder für einen von drei Algorithmen entscheiden. Es stehen somit sechs Alternativen zur Auswahl. Wenn die Einschätzung von Sela, Berger & Liu (2009) zutrifft, dass Wirtschaftssubjekte sich angesichts einer Vielzahl von Alternativen häufig für eine leicht zu rechtfertigende Alternative entscheiden, müsste im Treatment 2 die Häufigkeit der Algorithmusaversion geringer ausfallen als im Treatment 1. Da alle Algorithmen allen Experten hinsichtlich ihrer Erfolgsquote überlegen sind, würde eine besondere Betonung der Zweckdienlichkeit, die von einer Choice Overload ausgehen kann, im Treatment 2 zu einer geringeren Ausprägung der Algorithmusaversion führen.

Hypothese 1 lautet daher: Die Häufigkeit, mit der ein Algorithmus ausgewählt wird, fällt in Treatment 2 (sechs Alternativen) signifikant größer aus als im Treatment 1 (zwei Alternativen).

Nullhypothese 1 lautet dann: Die Häufigkeit, mit der ein Algorithmus ausgewählt wird, fällt in Treatment 2 (sechs Alternativen) nicht signifikant größer aus als im Treatment 1 (zwei Alternativen).

Wenn die Einschätzungen von Misuraca et al. (2024), von Ji et al. (2014) und Williamson & Eaker (2012) zutreffen, dass Frauen stärker von dem Phänomen der Choice Overload betroffen sind als Männer, und wenn außerdem zutrifft, dass Choice Overload eine dämpfende Wirkung auf die Algorithmusaversion ausübt, dann ist zu erwarten, dass Frauen im Treatment 2 (sechs Alternativen) noch seltener eine Algorithmusaversion zeigen als Männer.

Hypothese 2 lautet daher: Frauen wählen im Treatment 2 (sechs Alternativen) signifikant häufiger einen Algorithmus aus als Männer.

Nullhypothese 2 lautet somit: Frauen wählen im Treatment 2 (sechs Alternativen) nicht signifikant häufiger einen Algorithmus aus als Männer.

Die Probanden erhalten eine Vergütung in Höhe von 4,00 € , sofern eine Diagnose gestellt wird, die sich auch auf mittlere Sicht bestätigt. Entscheidet sich ein Proband für einen der Algorithmen, beträgt der Erwartungswert seiner Vergütung 2,80 € ($0,7 \cdot 4,00 \text{ €}$). Entscheidet sich ein Proband für einen der Experten, beträgt der Erwartungswert der Vergütung 2,40 € ($0,6 \cdot 4,00 \text{ €}$). Der Homo Oeconomicus müsste sich also zwingend für einen der Algorithmen entscheiden. Jede Entscheidung zugunsten eines der Ärzte muss als ein Fall von Algorithmusaversion gedeutet werden.

Die Ermittlung der konkreten Vergütung für jeden einzelnen Probanden erfolgt mit Hilfe eines Zufallsgenerators. Entscheidungen zugunsten eines Arztes führen in 60% aller Fälle zu einer Vergütung in Höhe von 4,00 € und Entscheidungen zugunsten eines Algorithmus führen in 70% aller Fälle zu einer Vergütung in Höhe von 4,00 € .

Das Experiment weist ein Between-Subject-Design auf. Das heißt, jeder Proband spielt nur eines der beiden Treatments. Die Zuordnung der Probanden zu einem der beiden Treatments erfolgt per Zufallsgenerator. Das Experiment ist in z-Tree implementiert (vgl. Fischbacher, 2007).

4 Ergebnisse

Die experimentelle Erhebung wurde im Zeitraum vom 24.04.2024 bis zum 08.05.2024 im Ostfalia Labor für experimentelle Wirtschaftsforschung (OLEW) in Wolfsburg durchgeführt. An der Erhebung haben 200 Studierende der Ostfalia Hochschule für angewandte Wissenschaften teilgenommen. 100 Probanden haben das Treatment 1 und 100 Probanden das Treatment 2 bearbeitet.

87 Probanden (43,5%) sind weiblich und 113 Probanden (56,5%) sind männlich. 134 Probanden (67,0%) gehören der Fakultät Wirtschaft, 55 (27,5%) der Fakultät Fahrzeugtechnik und 11 (5,5%) anderen Fakultäten an. 194 Probanden (97,0%) sind Studierende eines Bachelorstudiums und 6 (3,0%) sind Studierende eines Masterstudiums. Die Probanden haben im Durchschnitt bereits 4,22 Semester studiert und haben ein durchschnittliches Alter von 22,3 Jahren.

38,5% der Probanden entscheiden sich für einen der Ärzte, obwohl dies mit einer geringeren Erfolgswahrscheinlichkeit und mit einem geringeren Erwartungswert der Vergütung verbunden ist. Das heißt, dass 38,5% der Probanden die Verhaltensanomalie der Algorithmusaversion zeigen (Tabelle 1).

Tab. 1: Entscheidungen der 200 Probanden für einen Experten oder einen Algorithmus nach Treatments

	Entscheidungen für		
	Experte	Algorithmus	
Treatment 1 (Zwei Alternativen)	33 (33,0%)	67 (67,0%)	100
Treatment 2 (Sechs Alternativen)	44 (44,0%)	56 (56,0%)	100
Gesamt	77 (38,5%)	123 (61,5%)	200

Tatsächlich zeigt sich bei den beiden Treatments ein Unterschied hinsichtlich der Häufigkeit, in der Algorithmusaversion auftritt. Allerdings stellt sich dieser Unterschied anders dar, als auf Grundlage der Choice-Overload-Überlegungen zu erwarten war. Sela, Berger & Liu (2009) hatten darauf hingewiesen, dass eine größere Anzahl an Alternativen in aller Regel eine Tendenz zur Zweckdienlichkeit nach sich zieht. Die Unübersichtlichkeit, die mit einer großen Zahl von Alternativen einhergeht, führt häufig dazu, dass eine leicht zu rechtfertigende, zweckdienliche Alternative ausgewählt wird.

Angesichts unseres Versuchsaufbaus, in dem alle Algorithmen allen Experten offensichtlich überlegen sind, wäre die leicht zu rechtfertigende, zweckdienliche Alternative, einen der Algorithmen auszuwählen. Deshalb war zu erwarten, dass in Treatment 2 (sechs Alternativen) häufiger ein Algorithmus ausgewählt wird als in Treatment 1 (zwei Alternativen).

In der experimentellen Erhebung ist jedoch das Gegenteil eingetreten: Während in Treatment 1 (zwei Alternativen) nur in 33,0% der Fälle der Experte bevorzugt wird, wird im Treatment 2 (sechs Alternativen) in 44,0% der Fälle ein Experte

bevorzugt (Tabelle 1). Die Häufigkeit, in der Algorithmusaversion zu beobachten ist, geht mit einem Anstieg der Anzahl an Alternativen nicht etwa zurück, sondern sie nimmt sogar um 11,0 Prozentpunkte zu. Dieser Unterschied erweist sich im Chi-Quadrat-Test allerdings als statistisch nicht signifikant (p-Wert = 0,1099).

Die Nullhypothese 1 kann nicht verworfen werden. Der vermutete Zusammenhang zwischen Choice Overload einerseits und Algorithmusaversion andererseits kann im Rahmen unseres Experimentes nicht belegt werden. Es ist somit nicht davon auszugehen, dass durch die Erhöhung der Anzahl der Alternativen eine Reduzierung der Algorithmusaversion erzielt werden kann.

Bei Betrachtung etwaiger Unterschiede zwischen den Geschlechtern tritt ein überraschendes Ergebnis zutage.

Tab. 2: Entscheidungen der 113 männlichen Probanden für einen Experten oder einen Algorithmus nach Treatments

	Entscheidungen für		
	Experte	Algorithmus	
Treatment 1 (Zwei Alternativen)	20 (35,7%)	36 (64,3%)	56
Treatment 2 (Sechs Alternativen)	21 (36,8%)	36 (63,2%)	57
Gesamt	41 (36,3%)	72 (63,7%)	113

Die männlichen Probanden zeigen sich von der unterschiedlichen Anzahl an Alternativen in Treatment 1 (zwei Alternativen) und Treatment 2 (sechs Alternativen) vollkommen unbeeindruckt (siehe Tabelle 2). Gut ein Drittel der männlichen Probanden zeigt eine Algorithmusaversion (im Treatment 1 zu 35,7% und in Treatment 2 zu 36,8%). Es ist kein Zusammenhang zwischen Choice Overload einerseits und Algorithmusaversion andererseits zu erkennen.

Tab. 3: Entscheidungen der 87 weiblichen Probanden für einen Experten oder einen Algorithmus nach Treatments

	Entscheidungen für		
	Experte	Algorithmus	
Treatment 1 (Zwei Alternativen)	13 (29,5%)	31 (70,5%)	44
Treatment 2 (Sechs Alternativen)	23 (53,5%)	20 (46,5%)	43
Gesamt	36 (41,4%)	51 (58,6%)	87

Anders stellt sich die Situation bei den weiblichen Probanden dar (siehe Tabelle 3). Ihr Verhalten wird stark von der Anzahl der Alternativen beeinflusst. Allerdings stellt sich der Zusammenhang anders dar, als zunächst vermutet. Im Treatment 1 (zwei Alternativen) wählen 29,5% der weiblichen Probanden den Arzt

aus und zeigen damit eine Algorithmusaversion. Im Treatment 2 (sechs Alternativen) wählen 53,5% der weiblichen Probanden einen Arzt aus und erliegen somit einer Algorithmusaversion. Der Unterschied zwischen den Treatments 1 und 2 erweist sich bei einer Irrtumswahrscheinlichkeit von <5% als statistisch signifikant (p-Wert = 0,0234).

Die größere Anzahl an Alternativen im Treatment 2 führt bei den weiblichen Probanden nicht zu einem geringeren Ausmaß der Algorithmusaversion, sondern zu einem signifikant größeren Ausmaß der Algorithmusaversion.

Die Nullhypothese 2 kann somit nicht verworfen werden. Die Vermutung, dass Choice Overload gerade bei weiblichen Probanden eine deutlich dämpfende Wirkung auf das Auftreten von Algorithmusaversion habe, hat sich nicht bestätigt. Tatsächlich ist sogar das Gegenteil der Fall. Die größere Zahl an Alternativen in Treatment 2 hat die Neigung zur Algorithmusaversion bei den weiblichen Probanden nicht etwa gedämpft, sondern sogar noch angeheizt.

Angesichts dieser überraschenden Ergebnisse ist es sinnvoll, das Verhalten der Probanden noch etwas näher zu betrachten. Insbesondere ist zu erörtern, ob die sechs Alternativen in Treatments 2 im Vergleich zu den zwei Alternativen im Treatment 1 ausreichen, um eine nennenswerte Erschwernis bei der Entscheidungsfindung hervorzurufen. In dem vorliegenden Experiment wurde die Zeit gemessen, die die Probanden für ihre Entscheidungsfindung benötigt haben. Die Zeiterhebung beginnt erst nach der Lektüre der Spielanleitung, nach der detaillierten Kenntnisnahme der Entscheidungsalternativen und nach Beantwortung der Kontrollfragen. Die Zeiterhebung endet mit dem Abschluss der Entscheidung, das heißt, wenn die Probanden eine Alternative ausgewählt haben und dann auf den Button „o.k.“ drücken. Damit gehen sie zur Erhebung der demografischen Daten über. Die erhobene Zeitspanne bildet somit ausschließlich die Dauer der Entscheidungsfindung ab.

Sowohl Frauen als auch Männer benötigen für die Entscheidungsfindung im Treatment 2 deutlich länger als im Treatment 1 (Tabelle 4).

Tab. 4: Durchschnittliche Dauer der Entscheidungsfindung in Sekunden (Standardabweichung) nach Treatments und nach Geschlecht

	Frauen	Männer
Treatment 1 (Zwei Alternativen)	15,50 (12,06)	12,29 (7,51)
Treatment 2 (Sechs Alternativen)	33,07 (21,24)	23,19 (14,49)

Frauen brauchen für die Entscheidung zwischen zwei Alternativen durchschnittlich 15,50 Sekunden und für die Entscheidung zwischen sechs Alternativen durchschnittlich 33,07 Sekunden. Der Unterschied erweist sich im Mann-Whitney-U-Test als hochsignifikant (p-Wert < 0,001). Männer benötigen für die Entscheidung zwischen zwei Alternativen durchschnittlich 12,29 Sekunden und für

die Entscheidung zwischen sechs Alternativen durchschnittlich 23,19 Sekunden. Auch dieser Unterschied erweist sich im Mann-Whitney-U-Test als hochsignifikant (p -Wert $< 0,001$). Es kann also geschlossen werden, dass die Entscheidung im Treatment 2 (sechs Alternativen) sowohl für Frauen als auch für Männer herausfordernder war. Dies steht im Einklang mit den Ergebnissen der bisherigen Choice-Overload-Forschung.

Von besonderem Interesse sind jedoch die Unterschiede zwischen Frauen und Männern in den beiden Treatments. In Treatment 1 erweist sich der Unterschied in der durchschnittlichen Dauer der Entscheidungsfindung (15,50 Sekunden bei den Frauen und 12,29 Sekunden bei den Männern) im Mann-Whitney-U-Test nicht als signifikant (p -Wert = 0,179). Das stellt sich in Treatment 2 anders dar. Frauen benötigen durchschnittlich 33,07 Sekunden für die Auswahl aus den sechs Alternativen. Männer brauchen für dieselbe Entscheidung durchschnittlich nur 23,19 Sekunden. Dieser Unterschied erweist sich im Mann-Whitney-U-Test bei einer Irrtumswahrscheinlichkeit von $< 5\%$ als signifikant (p -Wert = 0,012). Auch das steht im Einklang mit der bisherigen Choice-Overload-Forschung, die von einer stärkeren Betroffenheit von Frauen bei einer Erhöhung der Zahl an Alternativen ausgeht (vgl. bspw. Misuraca et al., 2024; Ji, Ha & Sypher, 2014; Williamson & Eaker, 2012).

Das erklärt allerdings nicht, warum ausschließlich die Gruppe der weiblichen Probanden im Treatment 2 (sechs Alternativen) deutlich häufiger zur Algorithmusaversion neigt als im Treatment 1 (zwei Alternativen). Es bleibt weiteren Forschungsanstrengungen im Grenzbereich von Algorithmusaversion und Choice Overload vorbehalten, den Hintergrund dieses überraschenden empirischen Befundes aufzuklären.

Abschließend kann festgestellt werden, dass eine Erhöhung der Anzahl an Entscheidungsalternativen nicht zu einer Verringerung der Algorithmusaversion beiträgt. Bei Männern bringt eine Erhöhung der Anzahl an Entscheidungsalternativen keinen erkennbaren Einfluss auf ihre Neigung zur Algorithmusaversion mit sich. Bei Frauen führt eine Erhöhung der Anzahl der Entscheidungsalternativen sogar zu einem Anwachsen der Neigung zur Algorithmusaversion.

5 Zusammenfassung

Algorithmusaversion ist eine schädliche Verhaltensanomalie, weil sie zu einer Verringerung der Erfolgswahrscheinlichkeit beiträgt. Deshalb werden intensive Forschungsanstrengungen unternommen, um herauszufinden, wie die Algorithmusaversion reduziert werden kann. In einem kontrollierten Laborexperiment wird untersucht, ob eine Erhöhung der Anzahl an Alternativen zu einer Milderung der Algorithmusaversion beitragen kann. Forschungsanstrengungen zum Thema Choice Overload haben nämlich gezeigt, dass Wirtschaftssubjekte, sofern sie aus einer großen Zahl von Alternativen auswählen müssen, eher dazu neigen, leicht zu rechtfertigende Entscheidungen zu treffen. Bei großer Auswahl wird daher eher eine Alternative gewählt, die offensichtlich zweckdienlich ist.

In der vorliegenden Vignettenstudie nehmen die Probanden die Rolle eines Gesundheitsberaters ein, der einem Patienten eine Empfehlung geben soll. Im Treatment 1 stehen zur Diagnoseerstellung ein Arzt und ein Algorithmus zur Auswahl, wobei die Erfolgsaussichten bei der Wahl des Algorithmus offensichtlich um zehn Prozentpunkte höher liegen als bei der Wahl des Arztes. Im Treatment 2 stehen zur Diagnoseerstellung drei Ärzte mit unterschiedlichen Herangehensweisen und drei Algorithmen mit jeweils unterschiedlicher Schwerpunktsetzung zur Auswahl. Dabei ist die Erfolgsquote aller Algorithmen offensichtlich zehn Prozentpunkte höher als die Erfolgsquoten aller Ärzte.

Wenn eine größere Anzahl an Alternativen tatsächlich dazu beiträgt, dass eher eine leicht zu rechtfertigende Entscheidung getroffen wird und dass die Zweckdienlichkeit in den Mittelpunkt der Entscheidung rückt, dann müsste im Treatment 2 ein geringeres Maß an Algorithmusaversion auftreten als im Treatment 1.

Tatsächlich zeigt sich jedoch das Gegenteil. Im Treatment 2 (sechs Alternativen) wird seltener auf einen der überlegenen und deshalb zweckdienlichen Algorithmen gesetzt als im Treatment 1 (zwei Alternativen). Allerdings erweist sich der Unterschied zwischen den beiden Treatments nicht als statistisch signifikant.

Etwas anders stellt sich das Bild dar, wenn die Analyse auf die weiblichen Probanden beschränkt wird. Es gibt Hinweise darauf, dass Frauen noch stärker als Männer von Choice Overload betroffen sind. Wenn Choice Overload eine dämpfende Wirkung auf die Algorithmusaversion hat, müsste dieser Effekt besonders bei den weiblichen Probanden auftreten. Das Gegenteil ist jedoch der Fall. Die weiblichen Probanden neigen im Treatment 2 (sechs Alternativen) statistisch signifikant häufiger zur Algorithmusaversion als im Treatment 1 (zwei Alternativen).

Abschließend ist festzuhalten, dass von einer Erhöhung der Anzahl an Entscheidungsalternativen (mehr Experten beziehungsweise mehr Algorithmen) keine Milderung der Algorithmusaversion zu erwarten ist.

Literatur

- 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., Schechter, 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, 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., Balsler, 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.

Anhänge

Anhang 1: Aufgabenbeschreibung, Kontrollfragen und Entscheidungssituation im Treatment 1

Aufgabenbeschreibung

Sie beraten einen Patienten, dem ein malignes Melanom (Hautkrebs mit hoher Metastasierungsrate) entfernt werden musste. Um auszuschließen, dass der Hautkrebs im Körper Tochtergeschwülste gestreut hat, wurden zahlreiche Untersuchungen durchgeführt. Insbesondere wurde der gesamte Körper mit einem Magnetresonanztomographen (MRT) durchleuchtet.

Bei der Auswertung der MRT-Aufnahmen und weiteren Untersuchungsergebnissen gibt es zwei Alternativen zur Auswahl:

- a) Ein erfahrener Arzt, der sowohl Onkologe (Arzt mit dem Schwerpunkt Krebsbehandlung) als auch Radiologe (Arzt mit dem Schwerpunkt bildgebende Verfahren) ist, wertet alle Ergebnisse aus. Der Arzt erzielt in 60% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.
- b) Ein spezialisiertes Computerprogramm wertet alle Untersuchungsergebnisse aus. Es ist insbesondere auf die Mustererkennung bei MRT-Aufnahmen ausgelegt. Das Computerprogramm erzielt in 70% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.

Ihre Aufgabe besteht nun darin, dem Patienten eine dieser beiden Alternativen zu empfehlen. Dabei wird davon ausgegangen, dass der Patient ihrer Empfehlung folgen wird.

Zum Ablauf: Nach dem Lesen dieser Instruktionen und der Beantwortung der Kontrollfragen wird Ihnen die Entscheidungssituation präsentiert. In dieser müssen Sie eine der beiden Entscheidungsoptionen auswählen.

Zur Auszahlung: Für die Teilnahme an dieser Aufgabe erhalten Sie eine Vergütung in Abhängigkeit der von Ihnen getroffenen Entscheidung und eines Zufallsprinzips, welches sich nach den oben genannten Eintrittswahrscheinlichkeiten bemisst. Sollte es gelingen, eine Diagnose zu erstellen, die sich auch auf mittlere Sicht bestätigt, erhalten Sie 4,00 €. Sollte dies nicht gelingen, erhalten Sie keine Vergütung.

Kontrollfragen

1. Was ist ein malignes Melanom?

- Ein Hautkrebs mit hoher Metastasierungsrate (korrekt)
- Ein radiologisches diagnostisches Verfahren
- Eine krankhafte Neigung zur Melancholie

2. Welche Alternativen stehen hier zur Verfügung, um ein Untersuchungsergebnis zu erzielen, das sich auch auf mittlere Sicht bestätigt?

- Zum einen ein erfahrener Arzt, der sowohl Onkologe als auch Radiologe ist, und zum anderen ein spezialisiertes Computerprogramm. (korrekt)
- Zwei erfahrene Ärzte einerseits und vier verschiedene spezialisierte Computerprogramme andererseits.
- Ein spezialisiertes Computerprogramm einerseits und ein Wahrsager andererseits.

3. Wie häufig ist es in der Vergangenheit gelungen, mit Hilfe des spezialisierten Computerprogramms ein Untersuchungsergebnis zu erzielen, das sich auch auf mittlere Sicht bestätigt?

- In 60% der Fälle
- In 70% der Fälle (korrekt)
- In 80% der Fälle

4. Wie hoch fällt Ihre Vergütung aus, wenn es gelingt, eine Diagnose zu erstellen, die sich auf mittlere Sicht bestätigt?

- 2,00
- 3,00
- 4,00 (korrekt)

Entscheidungssituation

Treffen Sie nun Ihre Auswahl, welche Option Sie dem Patienten empfehlen!

- Ich empfehle, die Diagnose von einem erfahrenen Arzt, der sowohl Onkologe als auch Radiologe ist, erstellen zu lassen.

Ich empfehle, die Diagnose von dem spezialisierten Computerprogramm erstellen zu lassen, das insbesondere auf die Mustererkennung bei MRT-Aufnahmen ausgelegt ist.

Anhang 2: Aufgabenbeschreibung, Kontrollfragen und Entscheidungssituation im Treatment 2

Aufgabenbeschreibung

Sie beraten einen Patienten, dem ein malignes Melanom (Hautkrebs mit hoher Metastasierungsrate) entfernt werden musste. Um auszuschließen, dass der Hautkrebs im Körper Tochtergeschwülste gestreut hat, wurden zahlreiche Untersuchungen durchgeführt. Insbesondere wurde der gesamte Körper mit einem Magnetresonanztomographen (MRT) durchleuchtet.

Bei der Auswertung der MRT-Aufnahmen und weiteren Untersuchungsergebnissen gibt es sechs Alternativen zur Auswahl:

- a) Ein erfahrener Arzt, der sowohl Onkologe (Arzt mit dem Schwerpunkt Krebsbehandlung) als auch Radiologe (Arzt mit dem Schwerpunkt bildgebende Verfahren) ist, wertet alle Ergebnisse aus. Der Arzt erzielt in 60% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.
- b) Ein erfahrener Arzt, der sowohl Onkologe (Arzt mit dem Schwerpunkt Krebsbehandlung) als auch Radiologe (Arzt mit dem Schwerpunkt bildgebende Verfahren) ist, wertet alle Ergebnisse gemeinsam mit einem Kollegen aus. Das Vier-Augen-Prinzip soll Flüchtigkeitsfehler vermeiden. Das Ärzteteam erzielt in 60% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.
- c) Ein erfahrener Arzt, der sowohl Onkologe (Arzt mit dem Schwerpunkt Krebsbehandlung) als auch Radiologe (Arzt mit dem Schwerpunkt bildgebende Verfahren) ist, wertet alle Ergebnisse aus. Dabei folgt er strikt einem Diagnoseplan, der durch eine internationale Ärzte-Kommission erarbeitet worden ist. Der Diagnoseplan soll verhindern, dass Entscheidungen auf Basis einer unvollständigen Datenlage getroffen werden. Der Arzt erzielt in 60% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.
- d) Ein spezialisiertes Computerprogramm wertet alle Untersuchungsergebnisse aus. Es ist insbesondere auf die Mustererkennung bei MRT-Aufnahmen ausgelegt. Das Computerprogramm erzielt in 70% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.
- e) Ein spezialisiertes Computerprogramm wertet alle Untersuchungsergebnisse aus, legt dabei allerdings besonderes Gewicht auf die Einbeziehung von Ergebnissen der Elektronen- und Fluoreszenzmikroskopie. Das Computerprogramm erzielt in 70% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.
- f) Ein spezialisiertes Computerprogramm wertet alle Untersuchungsergebnisse aus, legt dabei allerdings besonderes Gewicht darauf, den

Rohdatenraum (k-Raum) mit Hilfe von neuronalen Netzen besser auszuwerten, um Bildrauschen und Artefakte zu vermeiden. Das Computerprogramm erzielt in 70% der Fälle ein Untersuchungsergebnis, das sich auch auf mittlere Sicht bestätigt.

Ihre Aufgabe besteht nun darin, dem Patienten eine dieser sechs Alternativen zu empfehlen. Dabei wird davon ausgegangen, dass der Patient ihrer Empfehlung folgen wird.

Zum Ablauf: Nach dem Lesen dieser Instruktionen und der Beantwortung der Kontrollfragen wird Ihnen die Entscheidungssituation präsentiert. In dieser müssen Sie eine der sechs Entscheidungsoptionen auswählen.

Zur Auszahlung: Für die Teilnahme an dieser Aufgabe erhalten Sie eine Vergütung in Abhängigkeit der von Ihnen getroffenen Entscheidung und eines Zufallsprinzips, welches sich nach den oben genannten Eintrittswahrscheinlichkeiten bemisst. Sollte es gelingen, eine Diagnose zu erstellen, die sich auch auf mittlere Sicht bestätigt, erhalten Sie 4,00 €. Sollte dies nicht gelingen, erhalten Sie keine Vergütung.

Kontrollfragen

1. Was ist ein malignes Melanom?
 - Ein Hautkrebs mit hoher Metastasierungsrate (korrekt)
 - Ein radiologisches diagnostisches Verfahren
 - Eine krankhafte Neigung zur Melancholie
2. Welche Alternativen stehen hier zur Verfügung, um ein Untersuchungsergebnis zu erzielen, das sich auch auf mittlere Sicht bestätigt?
 - Es stehen drei verschiedene Ärzte bzw. Ärzteteams mit jeweils unterschiedlichen Ausrichtungen sowie drei verschiedene spezialisierte Computerprogramme mit unterschiedlichen Schwerpunkten zur Auswahl. (korrekt)
 - Es steht ein erfahrener Arzt und fünf verschiedene spezialisierte Computerprogramme mit unterschiedlichen Schwerpunkten zur Auswahl.
 - Es stehen zwei erfahrene Ärzte mit jeweils unterschiedlichen Ausrichtungen zur Auswahl.
3. Wie häufig ist es in der Vergangenheit gelungen, mit Hilfe eines der spezialisierten Computerprogramme ein Untersuchungsergebnis zu erzielen, das sich auch auf mittlere Sicht bestätigt?
 - In 60% der Fälle
 - In 70% der Fälle (korrekt)
 - In 80% der Fälle
4. Wie hoch fällt Ihre Vergütung aus, wenn es gelingt, eine Diagnose zu erstellen, die sich auf mittlere Sicht bestätigt?

- 2,00 €
- 3,00 €
- 4,00 € (korrekt)

Entscheidungssituation

Treffen Sie nun Ihre Auswahl, welche Option Sie dem Patienten empfehlen!

- Ich empfehle, die Diagnose von einem erfahrenen Arzt, der sowohl Onkologe als auch Radiologe ist, erstellen zu lassen.
- Ich empfehle, die Diagnose von einem erfahrenen Arzt, der sowohl Onkologe als auch Radiologe ist und der mit Hilfe eines Kollegen auf das Vier-Augen-Prinzip setzt, erstellen zu lassen.
- Ich empfehle, die Diagnose von einem erfahrenen Arzt, der sowohl Onkologe als auch Radiologe ist und der strikt einem Diagnoseplan folgt, erstellen zu lassen.
- Ich empfehle, die Diagnose von dem spezialisierten Computerprogramm erstellen zu lassen, das insbesondere auf die Mustererkennung bei MRT-Aufnahmen ausgelegt ist.
- Ich empfehle, die Diagnose von dem spezialisierten Computerprogramm erstellen zu lassen, das besonderes Gewicht auf die Einbeziehung von Ergebnissen der Elektronen- und Fluoreszenzmikroskopie legt.
- Ich empfehle, die Diagnose von dem spezialisierten Computerprogramm erstellen zu lassen, das besonderes Gewicht darauf legt, den Rohdatenraum (k-Raum) mit Hilfe von neuronalen Netzen besser auszuwerten.