

Thomas Spengler · Wolf Fichtner
Martin Josef Geiger · Heinrich Rommelfanger
Olga Metzger *Hrsg.*

Entscheidungs- unterstützung in Theorie und Praxis

Tagungsband zum Workshop FEU
2016 der Gesellschaft für Operations
Research e.V.



Springer Gabler

Entscheidungsunterstützung in Theorie und Praxis

Thomas Spengler · Wolf Fichtner
Martin Josef Geiger
Heinrich Rommelfanger · Olga Metzger
(Hrsg.)

Entscheidungs- unterstützung in Theorie und Praxis

Tagungsband zum Workshop FEU
2016 der Gesellschaft für Operations
Research e.V.

Herausgeber

Thomas Spengler
Magdeburg, Deutschland

Heinrich Rommelfanger
Frankfurt am Main, Deutschland

Wolf Fichtner
Karlsruhe, Deutschland

Olga Metzger
Magdeburg, Deutschland

Martin Josef Geiger
Hamburg, Deutschland

Wir danken dem Forschungszentrum für Sparkassenentwicklung e.V. für die großzügige Unterstützung der Publikation.

ISBN 978-3-658-17579-5

ISBN 978-3-658-17580-1 (eBook)

DOI 10.1007/978-3-658-17580-1

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer Gabler

© Springer Fachmedien Wiesbaden GmbH 2017

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Springer Gabler ist Teil von Springer Nature

Die eingetragene Gesellschaft ist Springer Fachmedien Wiesbaden GmbH

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Vorwort

Im März 2016 fand erstmalig ein gemeinsamer Workshop der GOR-Arbeitsgruppen „OR im Umweltschutz“, „Entscheidungstheorie und –praxis“ sowie „Fuzzy Systeme, Neuronale Netze und Künstliche Intelligenz“ an der Otto-von-Guericke-Universität in Magdeburg statt. Der Workshop wurde von Prof. Dr. Thomas Spengler (Otto-von-Guericke-Universität Magdeburg) und seinem Lehrstuhl-Team in Zusammenarbeit mit den Leitern der Arbeitsgruppen Prof. Dr. Wolf Fichtner (Karlsruher Institut für Technologie), Prof. Dr. Martin Josef Geiger (Helmuth-Schmidt-Universität Hamburg) sowie Prof. Dr. Heinrich Rommelfanger (Emeritus der Universität Frankfurt a. M.) organisiert und vom Forschungszentrum für Sparkassenentwicklung e.V. sowie von der Gesellschaft für Operations Research e.V. (GOR) unterstützt.

Der Schwerpunkt des gemeinsamen Workshops lag auf dem Rahmenthema „Entscheidungsunterstützung in Theorie und Praxis“ und bot Raum für die Präsentation und Diskussion aktueller Forschungsansätze, die zur Unterstützung von allgemein betriebswirtschaftlichen sowie umweltorientierten Entscheidungsproblemen entwickelt und mit Methoden des Operations Research gelöst werden. In insgesamt 19 konzeptionellen, empirischen und experimentellen Beiträgen wurden im Einzelnen Entscheidungen bei Risiko und Ambiguität, Decision Support Systemen, Anwendungen von Fuzzy-Systemen und Neuronalen Netzen, Methoden zur Analyse umweltorientierter Entscheidungsprobleme, Probleme der Asset-Allokation sowie Komplexitätsfragen thematisiert. Im vorliegenden Tagungsband werden acht von Dritten begutachtete Beiträge in der Langfassung vorgestellt.

Im ersten Aufsatz „Error rates’ non-monotone dependence on problem difficulty with regard to human mathematical problem solving and their relation to individual characteristics“ untersuchen Lirije N. von Petersdorff, Stephan Schosser und Bodo Vogt das Reaktionsverhalten realer, menschlicher Entscheidungsträger auf unterschiedliche mathematische Fragestellungen. Im Detail konfrontie-

ren sie Testpersonen mit unterschiedlich komplexen Aufgabenstellungen und stellen die Problemlösungskompetenz der Probanden der Schwierigkeit der Fragestellung gegenüber. Hierbei arbeiten die Autoren nichtmonotone Zusammenhänge heraus. Es wird deutlich, dass die wahrgenommene Schwierigkeit der Problemstellung in den Kontext neuerer Erkenntnisse der Motivationstheorie zu setzen ist.

Der Beitrag “Using Objectives to Improve Decision-Making in Manufacturing Companies” von Julia Käppler, Johannes Siebert, Tom Drews und Paul Molenda beschäftigt sich mit der Entwicklung eines webbasierten Entscheidungsunterstützungssystems zur Steigerung der Effizienz der Produktionslogistik in Unternehmen. Da Entscheidungsträger typischerweise aus einer Vielzahl an Alternativen bei begrenzten Zeit-, Personal- und Budget-Ressourcen auswählen müssen, bedarf es der Entscheidungsunterstützung. Dabei kommt dem Zielsystem eine entscheidende Rolle zu, es sollte folgende Anforderungen erfüllen: Es sollte vollständig, einfach, redundanzfrei, exakt messbar und zerlegbar sein. Der Beitrag beschreibt die Entwicklung eines solchen Zielsystems.

Im darauffolgenden Beitrag mit dem Titel „Zur Bestimmung von unteren Schranken im bikriteriellen k -unähnlichen Tourenplanungsproblem“ stellt Sandra Zajac einen heuristischen Ansatz zur Bestimmung von unteren Schranken für k -unähnliche Tourenplanungsprobleme vor. Die Autorin thematisiert dabei den zentralen Konflikt solcher Probleme, die eine Minimierung der Distanz des längsten Tourenplans, bei gleichzeitiger Maximierung der minimalen Unähnlichkeit zwischen jeweils zwei Tourenplänen fordern. Illustriert wird die Anwendbarkeit des vorgestellten heuristischen Ansatzes, der sich der lokalen Suche bedient, im Rahmen eines anschließenden Rechenexperiments.

Martin J. Geiger behandelt in seinem Beitrag „Optimale Torpedo-Einsatzplanung – Analyse und Lösung eines Ablaufplanungsproblems der Stahlindustrie“ das Problem der Ablaufplanung von Torpedowagen innerhalb eines Stahlwerks. Die Problemstellung entstammt einem Implementierungswettbewerb der Association for Constraint Programming (ACP). Der korrespondierende Optimierungsansatz wird unter Berücksichtigung multipler Zielsetzungen und Einbezug von Ressourcen- und Zeitbeschränkungen modelliert. Eine Baumsuche dient als Basis

für das vom Autor vorgeschlagene Lösungsverfahren, welches auf die Instanzen des ACP Challenge angewendet wird. Die Leistungsstärke des Ansatzes wird im Zuge dessen in einer experimentellen Untersuchung demonstriert.

Der Aufsatz „Ein gemischt-ganzzahliges lineares Optimierungsmodell zur Ermittlung des minimal notwendigen Bedarfes an Seefahrern zur Besetzung eines Schiffes“ von Anisa Rizvanolli und Ole John behandelt ein aktuelles Planungsproblem im Kontext der maritimen Logistik. Ganz generell stellt die Einsatzplanung der einem Schiff zugeordneten Besatzungsmitglieder ein schwieriges Optimierungsproblem dar, welches wiederkehrend bestmöglich zu lösen ist. Offensichtlich ergeben sich in diesem Spannungsfeld Konsequenzen für die Betriebskosten der Schifffahrtsunternehmen, wobei weitere Aspekte aus Arbeitnehmersicht hinzutreten. Für diese Fragestellung stellen Rizvanolli und John ein gemischt-ganzzahliges Optimierungsmodell vor. Durch eine konkrete Implementierung des Formalmodells in OPL bzw. dessen Lösung mit CPLEX wird die Leistungsfähigkeit des Ansatzes für praxisnahe Datensätze herausgearbeitet.

In ihrem Beitrag „Subjektiver Erwartungsnutzen und intuitionistische Fuzzy-Werte“ präsentieren Olga Metzger und Thomas Spengler einen innovativen Lösungsansatz für Entscheidungssituationen, bei denen der Entscheider sowohl die Eintrittswahrscheinlichkeiten der Umweltzustände als auch die Konsequenzen der einzelnen Alternativen lediglich in Intervallform beschreiben kann. Das Verfahren basiert auf der von Atanassov entwickelten Theorie „Intuitionistischer Fuzzy Mengen“, bei der neben Zugehörigkeitswerten auch Nicht-Zugehörigkeitswerte und Indeterminiertheitsgrade die Lösung beeinflussen. Zusätzlich zu den formulierten Modellansätzen wird auch auf die Verbindung der intuitionistischen Fuzzy-Set-Theorie mit weiteren modallogischen Konstrukten und Konzepten eingegangen.

In dem Beitrag „Individualisation of a Fuzzy System for Asset Allocation“ zeigt Reiner North anhand ausgewählter Beispiele auf, wie das von ihm entwickelte Fuzzy-Expertensystem an die individuellen Beratungsphilosophien einzelner Finanzinstitute angepasst werden kann. Auch wenn sein Software-Modell für

die Asset Allokation in Anlehnung an die Vorgehensweise einer international führenden Kapitalanlagegesellschaft entwickelt wurde, ist davon auszugehen, dass Anlageberater anderer Finanzinstitute Anpassungs- und Erweiterungswünsche an die Software haben. Im Einzelnen beschreibt Reiner North, wie das Pilotmodell sich mit begrenztem Aufwand an die Aufnahme zusätzlicher Eingangsgrößen, zusätzlicher Assetklassen und weiterer Anlagestrategien anpassen lässt.

Im letzten Beitrag “Status quo bias and consumers’ willingness to pay for green electricity: A discrete choice experiment with real economic incentives” beschreiben Fabian Grabicki und Roland Menges ein Discrete Choice Experiment zur Analyse des Status Quo Bias. An dem Experiment nahmen über 100 Studenten der Technischen Universität Clausthal teil. Neben dem Status Quo Bias sollte analysiert werden, ob Unterschiede in hypothetischen und nicht-hypothetischen Szenarien beobachtet werden können. Die Ergebnisse dieser ersten Untersuchung zeigen keine signifikanten Unterschiede zwischen neutralem Framing und Status Quo Framing. Das gleiche Ergebnis konnte für die Untersuchung der hypothetischen und nicht-hypothetischen Zahlungsbereitschaften festgestellt werden.

Im Anschluss an die Langfassungen der acht ausgewählten Beiträge wird eine Übersicht über alle Vorträge der Tagung zur Verfügung gestellt.

Abschließend möchten wir uns bei den Teilnehmern, den Referenten und den Diskutanten des Workshops bedanken. Ein großer Dank geht ebenfalls an die anonymen Gutachter für ihre Unterstützung bei der Vorbereitung dieser Veröffentlichung sowie an alle Helfer, die an der redaktionellen Arbeit für den vorliegenden Tagungsband beteiligt waren.

Magdeburg, im Dezember 2016

Die Herausgeber

Inhaltsverzeichnis

Error rates' non-monotone dependence on problem difficulty with regard to human mathematical problem solving and their relation to individual characteristics	1
Lirije N. v. Petersdorff, Stephan Schosser and Bodo Vogt	
Using Objectives to Improve Decision-Making in Manufacturing Companies	15
Julia K��ppler, Johannes Siebert, Tom Drews and Paul Molenda	
Zur Bestimmung von unteren Schranken im bikriteriellen <i>k</i>-un��hnlichen Tourenplanungsproblem.....	37
Sandra Zajac	
Optimale Torpedo-Einsatzplanung – Analyse und L��sung eines Ablaufplanungsproblems der Stahlindustrie	63
Martin Josef Geiger	
Ein gemischt-ganzzahliges lineares Optimierungsmodell zur Ermittlung des minimal notwendigen Bedarfes an Seefahrern zur Besetzung eines Schiffes	87
Anisa Rizvanolli und Ole John	
Subjektiver Erwartungsnutzen und intuitionistische Fuzzy-Werte	109
Olga Metzger und Thomas Spengler	
Individualisation of a Fuzzy System for Asset Allocation	139
Reiner North	
Status quo bias and consumers' willingness to pay for green electricity: A discrete choice experiment with real economic incentives.....	169
Fabian Grabicki and Roland Menges	
Vortrags��bersicht der Tagung	191

Error rates' non-monotone dependence on problem difficulty with regard to human mathematical problem solving and their relation to individual characteristics

Lirije N. von Petersdorff, Stephan Schosser and Bodo Vogt

Abstract

We use mathematical problems with varying levels of difficulty (easy, intermediate, difficult) to test whether the error rates of human problem solvers depend monotonically on the problems' level of difficulty. In an experiment with 45 students, we observe a non-linear dependence. The hypotheses tested relate to whether the function type of the error rate depends on individual characteristics and/or the IQ of a subject. Our data shows that the different types of non-linearity (U-shaped and inverse U-shaped) depend on openness in the NEO-FFI but not on the subjects' IQ. We discuss our results in the context of problem solving in computer science, especially complexity theory, and motivation theory.

L. N. v. Petersdorff (✉) · S. Schosser · B. Vogt
Empirical Economics,
Otto-von-Guericke-University Magdeburg, 39106 Magdeburg, Germany
e-mail: lirije.petersdorff@ovgu.de

Content

- 1 Introduction
- 2 The Tasks and Tests
 - 2.1 The Tasks
 - 2.2 The Tests: IQ test
 - 2.3 The Tests: NEO-FFI
- 3 Hypotheses
- 4 Experiment
- 5 Results
- 6 Conclusion

1 Introduction

One approach to model problem solving is to describe how computer and computer programs solve problems of varying difficulty. In computer science, this is a very important topic on which a lot of literature can be found (for an example of an overview, see Fortnow and Holmer, 2003). Different limitations are considered in this literature, but all models have one thing in common: The more difficult a problem is, the more resources are required to solve it. If the resources (such as time, as in the present study) are fixed, fewer problems are solved correctly.

Different automaton models have been used to explain human problem solving. Neyman (1985) and Rubinstein (1986) were the first to deal with decision-making behavior on computational complexity. Complexity can be formally defined, which helps to distinguish deviations from utility maximization. Piccione and Rubinstein (1993), Neyman (1998), Chatterjee and Sabourian (2000), and Salant (2011) worked with and extended computational models. Salant (2011) applies complexity theory to individual decision theory instead of games with two or more players. He finds that complexity theory can justify the use of several satisficing strategies.

What all these models have in common is that, for a given amount of time, fewer problems are solved as the difficulty increases. This way of reasoning also seems to guide the main thrust of economic thinking regarding the difficulty of problem solving.

We regard complexity as one major part of the difficulty of a problem, although other factors like a subject's context, ability or motivation also appear to be important in human decision making.

Besides this literature, work on motivation theory (Atkinson, 1957) states that at least the motivation depends on individual characteristics. In this theory, tasks' relative attractiveness depends on the motive to achieve success and on the motive to avoid non-success. Fig. 1 shows the curvatures of the relative attractiveness depending on the difficulty of a problem, i.e. the objective probability of success in this case.

The resulting curves are U-shaped or inverse U-shaped. If one assumes that motivation is related to this relative attractiveness, motivation is not a monotone function of difficulty. If one further assumes that motivation determines the number of correctly solved problems in a fixed time or the error rate in a monotone way, neither the number of correctly solved problems nor the error rate depends on the difficulty in a monotone way.

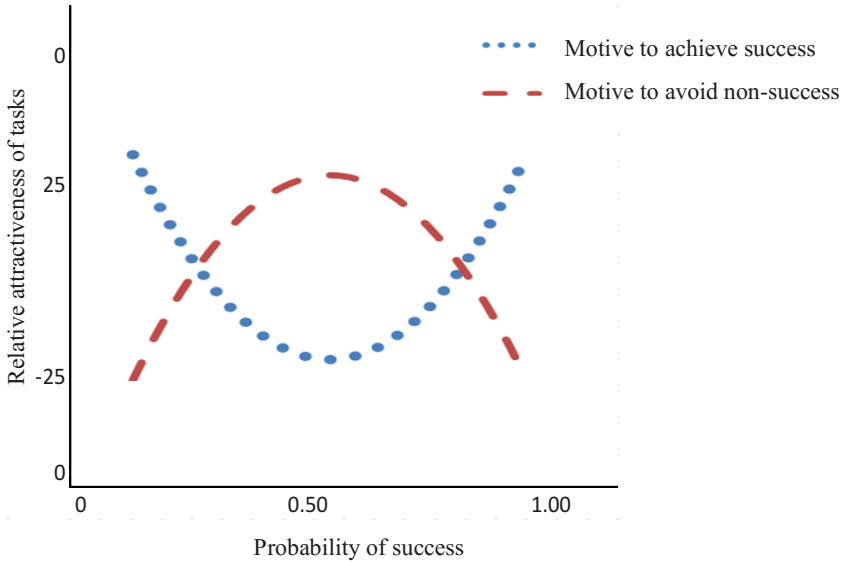


Fig. 1: Relative attractiveness of tasks (adopted from Fig. 2 in Atkinson, 1957)

In this paper, we will first test the idea of complexity theory that the error rates monotonically depend on the difficulty of a problem. Our results show that this is not the case. Then we follow the idea that the error rate in problem solving, especially the non-monotonicity of the functional relation, depends on individual characteristics. We will not use the characteristics that Atkinson proposes but will employ the five forces measured by the NEO-FFI (Borkenau and Ostendorf, 2008, following Costa and McCrae, 1989), since these characteristics are assumed to be more general and stable. Besides personal characteristics, we will also measure IQ in order to have a measure of the subject's objective capability to solve problems.

We restrict ourselves to mathematical problem solving because the difficulty of a problem seems to be more easily controllable than it is in other problems. Therefore, when we write problem solving, we mean mathematical problem solving, and our results are restricted to these problems. We will have three different tasks of equations with varying difficulty: A, B, and C, with the difficulty increasing from A to C. These tasks have to be solved in a fixed amount of time (10 minutes for each task). According to the predictions of computer science, the error rate should increase monotonically with difficulty. Following our interpretation of human problem solving, different curvatures might be expected depending on individual characteristics.

2 The Tasks and Tests

2.1 The Tasks

To disentangle the differences in human decision making as a function of task difficulty, we created three tasks of 50 mathematical problems (single mathematical equations) each, analogous to an IQ test. These are analogous to the I-S-T 2000R test, which consists of three subtests: analogies, numerical series, and matrices calculation (Liepmann, Beauducel, Brocke and Amthauer, 2007). Each problem consists of different arithmetic operations, including addition, subtraction, multiplication, division, root extraction, and exponentiation, as well as the use of logarithms and fractions. A task consists of 50 problems of equal difficulty. The three tasks differed with regard to the level of difficulty of the mathematical problems they contained: Task A was easy, Task B was intermediate, and Task C was difficult.

All three tasks contained mathematical problems requiring the use of multiplication, addition, subtraction, and division. All operands were integers lower than 300. We increased the difficulty from one task to the next by increasing the number of operands per mathematical problem: In Task A (see lines 1-3 of Table 1 for examples), each mathematical problem consisted of a maximum of one operator (e.g. subtraction, multiplication or addition). Corresponding problems are solvable

by pupils in Grade 3 of a typical German elementary school. In Task B (lines 4 and 5 in Table 1), we increased the number of operators to two (e.g. multiplication and division, division and addition) and designed mathematical problems so that the corresponding equation had to be rearranged in order to find the solution. In the most difficult task, Task C (lines 6 and 7 in Table 1), the mathematical problems consisted of more than three different operands (e.g. fraction, addition, brackets, and multiplication).

Table 1: Mathematical operations

(1)	$135 - 18 =$	Simple
(2)	$8 * 16 =$	
(3)	$98 + 63 =$	
(4)	$42 * 9 \div 5 =$	Intermediate
(5)	$30 \div 6 + 12 =$	
(6)	$(7/10 + 3/10) * 4 =$	Difficult
(7)	$5 * 3 * [3 + (2*3)] =$	

2.2 The Tests: IQ test

An IQ test is used to determine the intelligence of a human decision maker and requires the participant to solve several different types of tasks. For this paper, we chose the CT-20R, an extension of the CFT-20 in terms of “general fluid ability” pursuant to the work of Cattell (1960). It is not susceptible to cultural differences because it consists primarily of graphic and numerical tasks. The CFT-20R yields valuable results pupils at the age of 8,5 and 19 and adults between 20 and 60 years of age (Weiß, 2006). This IQ test helps to differentiate among the upper ability range. We expect it to be particularly suitable for testing the intelligence of university students, who, according to our hypothesis, are more intelligent than the average decision maker. Furthermore, stress reactions, slower work, and anxiousness have no real impact on the test. The CFT-20R we used was the German version in the form of a paper-and-pencil questionnaire (Weiß, 2006).

2.3 The Tests: NEO-FFI

To derive personality traits, we conducted the NEO five-factor inventory (NEO-FFI), which attempts to reflect a converging general consensus in differential psychology (Borkenau and Ostendorf, 2008, following Costa and McCrae, 1989). The Big Five model can incorporate a number of other personality theories. As such, it serves as a useful general framework to analyze human behavior (Costa and McCrae, 1995). This test measures five personality traits: neuroticism (N), extraversion (E), openness (O), agreeableness (A), and conscientiousness (C). Neuroticism (N) describes the extent to which humans experience psychological strain.

Decision makers with high neuroticism tend to be anxious and nervous; they also demonstrate depressive behavior because they experience the concomitant feelings for a longer period of time and more intensely than do the decision makers with low neuroticism levels. Decision makers with low levels of neuroticism tend to be calm and self-confident. Extraversion (E) correlates with activity and sociability. Extroverted decision makers are more active, joyful, and talkative than introverted decision makers, who tend to be reserved. Openness (O) focuses on the openness to new experiences. Open decision makers are imaginative, inquisitive, and interested in art. They are critical of fixed values and standards. Decision makers with low scores in openness are more likely to be critical and distrustful. Agreeableness (A) measures decision makers' interpersonal skills. A high score in agreeableness is an indicator of confidence, empathy, and cooperativeness. By contrast, decision makers with low agreeableness are cynical, cold-hearted, and hostile. Conscientiousness (C) assesses self-control. Conscientious decision makers are careful, organized, and diligent, and they exhibit reliable behavior (Costa and McCrae, 1989).

3 Hypotheses

According to the computer science literature, we expect the following:

Hypothesis 1a (Monotonicity): The more difficult a problem is, the higher the error rate will be.

According to our interpretation of Atkinson (1957), or at least analogously to the non-linear attractiveness of problems, as described in the introduction, we expect:

Hypothesis 1b (Non-monotonicity): The error rate does not depend on the increased difficulty of a problem in a linear way: U-shaped and inverse U-shaped error rate functions can be observed.

In line with our discussion of Atkinson (1957) in the introduction, we also expect individual characteristics to determine the error rate:

Hypothesis 2 (Individual characteristics): The function type of the error rates (U-shaped, inverse U-shaped) depends on individual characteristics.

Hypothesis 3 is about objective abilities to solve problems, i.e. IQ. The IQ should determine the number of correctly solved problems but not the shape of the error rate function.

Hypothesis 3 (IQ): The IQ does not determine the relative error rate function.

4 Experiment

Forty-five master's degree students (males: $n=17$, age: 22–32 years, mean=25.51, $SD=2.10$) from the Otto-von-Guericke University in Magdeburg participated in this study¹. At the end of the experiment, the participants received an average pay-out of €6.35 each for their participation in the experiment. All participants were assigned to isolated semi-cubicles to prevent them from seeing each other's answers. We also ensured that the participants could not see the others' results after the experiment had ended. Therefore, it is possible to assume that the students were not afraid their IQ or other results would be revealed to other students, and they undertook all the tasks and tests on their own. The study, including all the

¹ We are aware that the sample determines our results.

tasks and the test, lasted 180 minutes in total. After assigning the participants to the semi-cubicles, we conducted an IQ test. An experimenter explained all the tasks in the IQ test to the participants, and by having all the participants solve a few test questions, we made sure they understood every task. After the IQ test had been completed, the participants were given 10 minutes to solve 50 mathematical problems per task (Task A = 50 simple mathematical problems, Task B = 50 intermediate mathematical problems, Task C = 50 difficult mathematical problems).

No participant solved all of the problems in any of the tasks. In total, they had 30 minutes for 150 mathematical problems with different difficulty levels. After this part, the participants filled in the psychological questionnaire.

5 Results

Hypothesis 1 (Monotonicity, non-monotonicity)

We classified subjects as monotone if the relative error rate increased monotonically with the difficulty of Task A, B or C. The U-shaped type consists of people for whom the relative error rate shows a U shape in its dependence on the difficulty. The last type was the inverse U-shaped class comprising all subjects with this functional dependence on the relative error rate. Fig. 2 shows the average relative error rates of the types (with at least 10 subjects per type).

Our result is a decreasing trend from Task A to Task B and an increasing trend to Task C.

With regard to the p-value of the U-type decision-making behavior, it is clear that the relative error in both the intermediate and the difficult mathematical tasks is highly significant ($p = 0.015$). Using the Wilcoxon test, we show a significant relationship between Task A and Task B ($p = 0.007$). The same is true of the connection between Task B and Task C ($p = 0.007$). Our findings show that the participants made more mistakes in the simple mathematical (Task A) tasks than in the intermediate (Task B) ones. However, they made fewer errors in the intermediate mathematical tasks (Task B) than in the difficult ones (Task C).

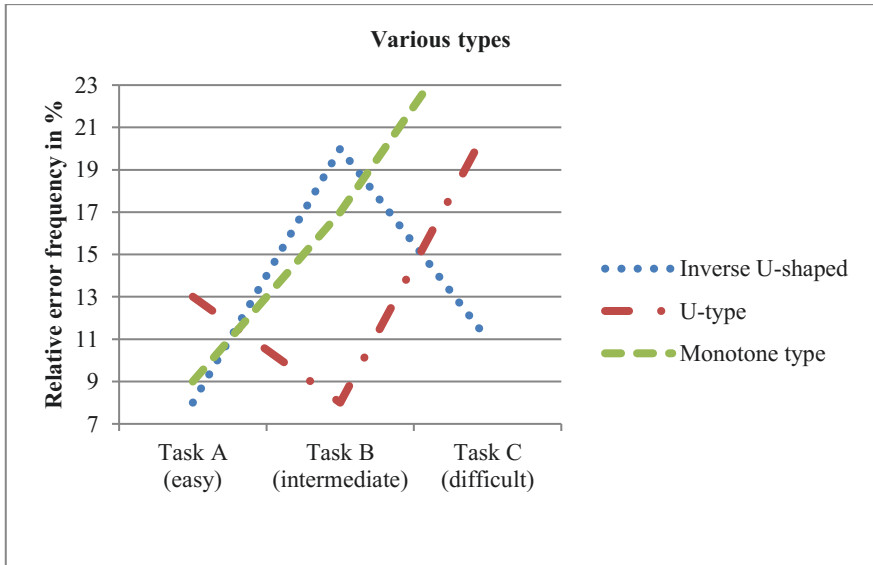


Fig. 2: The average relative error rate for the different types

U-type error frequency

There is a positive correlation between the relative error and the difficulty level of the mathematical tasks. This means that the participants made more relative errors in simple mathematical tasks and that the relative error increases in the mathematical tasks of an intermediate difficulty (Spearman correlation test between simple and intermediate mathematical tasks, $r = 0.865$; between simple and difficult, $r = 0.736$; between intermediate and difficult, $r = 0.569$).

Table 2: Spearman correlation matrix for the inverse U-shaped type of the mathematical tasks

Level of difficulty	Simple	Intermediate	Difficult
Simple	1	0.856	0.736
Intermediate	0.856	1	0.569
Difficult	0.736	0.569	1

Furthermore, the correlation between simple mathematical tasks and intermediate/difficult ones is significant ($p = 0.001$ and $p = 0.013$).

Inverse U-shaped relative error rate

In this type of decision-making behavior, the participants show a rising error rate in the simple tasks. There is a falling relative error rate from the intermediate to the difficult level. The relative error mean values show that inverse U-shaped decision-making behavior follows a course that contrasts with that of the previous types. With regard to the p-value of the inverse U-shaped decision-making behavior, the relative error in the intermediate mathematical task and in the difficult one is significantly different ($p = 0.015$). Overall, the existence of the U-shaped type and the inverse U-shaped type shows that Hypothesis 1a (monotonicity) is rejected but that Hypothesis 1b (non-monotonicity) cannot be rejected.

Hypothesis 2 (Individual characteristics)

The analysis of the results of the NEO-FFI and the correlation with the tasks shows no correlation for four forces. But there is a correlation between the openness (O) personality trait and people's decision making (chi-square test, $p = 0.010$). Fig. 3 shows that the values for openness are different for the different types. The difference between U-shaped and inverse U-shaped is significantly different on the 5% level. This correlation supports Hypothesis 2 in that the error rate in mathematical problem solving depends on individual characteristics.

Hypothesis 3 (IQ)

On average, the inverse U-shaped-type decision maker has the highest IQ (114), followed by the U-type (110) and the monotone-type (109) decision makers. A test shows that there is no correlation between the intelligence quotient and the decision-making types (according to the Mann–Whitney U test, linear type–U type, $p = 0.739$; linear type–bell type, $p = 0.883$; U type–bell type, $p = 0.805$). However, there is a strong positive correlation between the relative error rates and the IQ.

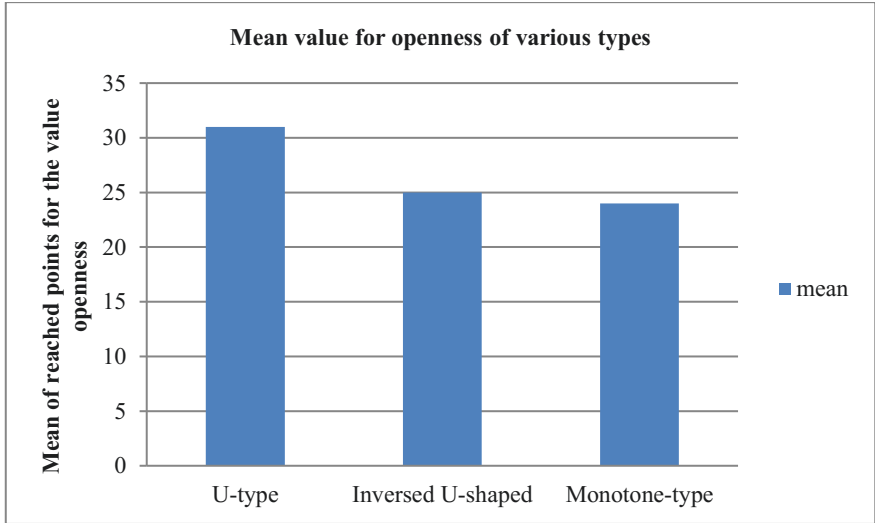


Fig. 3: The mean value of openness depending on the different types

This result supports Hypothesis 3 that the types do not depend on an objective ability to solve mathematical problems.

6 Conclusion

We used mathematical problem solving in fixed time as an experimental paradigm for problem solving. The difficulty of the tasks varied. We selected mathematical problems because, unlike other problems, the difficulty of these particular ones seems to be apparent.

The main result of our experiment is that the error rates in solving mathematical problems do not necessarily increase with difficulty. We observed non-linear dependences. This non-linear dependence correlates with personal characteristics (openness in the NEO-FFI) but not with the IQ.

Given our experimental results, the modeling of human problem solving should contain more factors than the ones used to describe problem solving by computers. One possible candidate is motivation theory, but there are also other