HELMUT-SCHMIDT-UNIVERSITÄT / UNIVERSITÄT DER BUNDESWEHR HAMBURG

MAKING STRATEGIC DECISIONS UNDER TIME PRESSURE – A Process-based Analysis Approach

DISSERTATION

zur Erlangung des akademischen Grades eines Doktors
der Wirtschafts- und Sozialwissenschaften
des Fachbereiches Wirtschafts- und Sozialwissenschaften
der Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg

vorgelegt von
Marco Kremer aus Berlin

Hamburg, 2019

Danksagungen

Die Entstehung der vorliegenden Dissertation ist ein echtes Langzeitprojekt und hat mich von der Annahme als Doktorand durch Herrn Professor Tüshaus im Juli des Jahres 2009 bis zum heutigen Tage stetig über einen Zeitraum von mehr als neun Jahren beschäftigt. Der Fortschritt des Projekts, insbesondere in den kritischen Phasen Anfang, Mittelteil und Schluss wäre nie möglich gewesen ohne die strategische und inhaltliche Begleitung durch Herrn Professor Tüshaus und Herrn Professor Beckmann. Beide haben mich durch ihre immense Erfahrung auf den richtigen Weg gebracht und den Fokus für das Wesentliche in der Arbeit an solch einem Projekt geschärft. Sie ermöglichten mir darüber hinaus, meine Arbeit an verschiedenen Anlässen zu präsentieren und das ambitionierte Experiment mit den notwendigen Mitteln auszustatten.

Meiner Familie und meinen Freunden möchte ich danken für ihr reges Interesse am Fortschritt und den Ergebnissen der Arbeit. Es zeugt von einer ehrlichen Anteilnahme am Projekt und an meiner Person. Auf ihre Wertschätzung bin ich sehr stolz.

Ganz besonderer Dank gebührt meiner Frau Anja und meinen beiden Kindern Ben und Ida. Sie haben mir während der gesamten Zeit die größtmögliche Unterstützung zuteilwerden lassen.

Abschließend möchte ich meinen Großeltern für ihre bedingungslose Unterstützung danken.

Berlin, im Februar 2019

Marco Kremer

Contents

List of Figures

List of Tables

List of Equations

List of Abbreviations

Abbr.	Meaning (English Translation)
AOI	Area of Interest – gazing spot
CH	Cognitive Hierarchy
CI	Confidence Interval
eHS	erweiterter Hauptschulabschluss (Extended General Certificate of Secondary Education)
EIPs	Elementary Information Processes
EMMAs	Elementary Manual Motor Acts
fMRI	functional Magnetic Resonance Imaging
GNU	recursive acronym meaning "Gnu's not Unix" – name of a prominent free-software-project that provides general public licenses
GUI	Graphical User Interface
H	Hypothesis
HS	Hauptschulabschluss (General Certificate of Secondary Education)
HTML	Hypertext Markup Language
I/O	Input / Output
JS	JavaScript
LS	Method of Least Squares
mR	mittlere Reife (Intermediate Level Certificate of Secondary Education)
PHP	Hypertext Preprocessor
PSM	Problem Solving Method
PTM/PTT	Process Tracing Method/Techniques
SQL	Structured Query Language
URL	Uniform Resource Locator – web source
VBA	Visual Basic for Applications
VHB	Verband der Hochschullehrer für Betriebswirtschaft e.V.

List of Symbols

Symbol	Meaning
$\propto$	Confidence level
ε	Value for Cauchy's convergence test in simulation
μ	Population mean
$\mu*$	Population mean based on data-set-limitations according to Hypothesis IV
$\mu^{\backslash}*$	Population mean based on data-set-limitations according to Hypothesis V
$\mu_{nonsoph}$	Population mean of non-sophisticated dataset
μ_{soph}	Population mean of sophisticated dataset
σ	Population standard deviation
a	Alternative of choice
c_i	Complexity-level of task i
CI	Confidence interval
d_i	Ratio of time per EIP
f	Degrees of freedom in t-Test
g_i	Goal i of evaluation concept
H_0	Null-hypothesis
H_1	Alternative hypothesis
k	Level of cognitive hierarchy, $k \in \mathbb{N}$
n	Sample size of dataset
n_{Eq}	Sample size of dataset with characteristic "Equilibrium Choice"
nxm	Matrix size in normal-form games with n rows and m columns according to the players' alternatives
p	p-Value of hypothesis test
p_0	Population proportion reference value
$p_{DR,Ve}$	Population dropping rate in experiment version V_e
p_{Eq}	Proportion of equilibrium choices in dataset
p_i	Population share of task i
$\hat{p}$	Sample share
$\hat{p}_{12}$	Pooled sample proportion of sample numbers 1 and 2
$\hat{p}_i$	Sample share of task i
r_i	Ratio of focussing on own payoff
s_x	Sample standard deviation of reference dataset x
s_y	Sample Standard deviation of comparing dataset y
T	Test statistic
$t_{finish,i}$	Time a player needed to finish task i
t_i	Time limit of task i
$t_{own\ focus,i}$	Time focussing on own payoff in task i

Symbol	Meaning
t_p	Processing time for a single EIP
$t_{total\ focus,i}$	Time focussing on payoff in task *i*
u_i	Player's absolute payoff in task *i*
$u_{max,i}$	Maximal achievable payoff for a player in task *i*
$u_{min,i}$	Minimal achievable payoff for a player in task *i*
$u_{rel,i}$	Player's relative payoff in task *i*
V_e	Experimental setup V in version e with $e \in \{1,2,3\}$
x_i	Value from dataset *i*
$\bar{x}$	Sample Mean of dataset *x*
$\bar{y}$	Sample Mean of dataset *y*
$z*$	Value from the standard normal distribution corresponding to the sought confidence level

1 Introduction

"sedulo curavi, humanas actiones non ridere, non lugere, neque detestari, sed intelligere;"[1]

Spinoza – Tractatus politicus, 1 § 4

1.1 Study's Objectives

Making (strategic) decisions under time pressure is part of everyday experience of many people (Ariely and Zakay 2001, p. 204). People reportedly feel stressed or even overwhelmed under such conditions, negatively influencing their decisions in general (Edland and Svenson 1993, p. 30). Psychological studies of Miller (1960, p. 697), Thomas and Weaver (1975, p. 366), Ben Zur and Breznitz (1981, p. 102), and Zakay (1993, p. 60) report a systematic change of behavior under growing time pressure. Findings in an economic decision-making context by Payne et al. (1988, p. 576) and Reutskaja et al. (2011, pp. 922 f.) widely confirm that change. However, decision-making processes under time pressure, especially with strategic goals, are not well understood yet (Ariely and Zakay 2001, p. 204). Further research is required to understand the impact of time pressure on the cognitive decision-making process that leads to the observed change of behavior (Ordóñez et al. 2015, p. 535).

In 2001, Costa-Gomes et al. introduce a process-based analysis approach to study strategic decision-making in game theory's two-person noncooperative normal-form games (Costa-Gomes et al. 2001). The authors employ mouse tracking as process tracing method (PTM) to acquire data that allows inferring on the underlying cognitive process of decision-making. Searching for behavioral patterns, they compare observations to common heuristics and consequently classify subjects. This approach is adopted from earlier works of Johnson and Payne (1985), Payne et al. (1988) and Payne et al. (1992) who developed a similar approach to study nonstrategic decision-making. However, the authors' basic assumption that people apply complete heuristics in their decision-making without learning has already been criticized earlier by Bettman (1979). It is more likely that rather parts of a heuristic and not necessarily only one per choice situation are employed and constructed while processing, he argues (Bettman 1979, p. 33). His suggestion gains support from experimental findings of Johnson et al. (2008, pp. 269 f.).

[1] Latin for "*I have labored carefully, not to mock, lament, or execrate, but to understand human actions;"* (de Spinoza 1883).

At the time Costa-Gomes et al. (2001) published their paper, mouse tracking was not capable of acquiring data necessary to study this hypothesis properly. Nowadays, techniques have improved considerably, offering new functionalities that help to study this issue in more detail. Especially the process model can be improved by integrating more means of observation.

Costa-Gomes et al. (2001) moreover limited their study to game complexities of two by two to two by four choice alternatives for the players. This approach failed to elucidate the impact of complexity on the decision-making process. Time pressure was not studied either, even though duration times of specific observed actions play an essential role in their econometric framework to classify behavior.

In this paper, decision-making patterns in strategic, noncooperative two-person normal-form games are examined. The patterns' sensitivity to task complexity and to time limitation is further determined. The process-theoretical research approach follows the one of Costa-Gomes et al. (2001). However, the process tracing technique 'mouse tracking' is updated. In this way, more components of the underlying cognitive processes can be identified, allowing more detailed models. This in turn helps to classify the behavior in such problem situations under different conditions and reasonably reflect on Bettman's critique. Potentially successful behavioral patterns in the challenging task of "good" decision-making under time pressure can be identified. At the same time, the understanding of cognitive processes themselves can be improved.

1.2 Considered Research Questions

The main research intention is to study cognitive processes in strategic decision-making under time pressure to improve understanding of human behavior under such conditions. Research questions provided in the following concretize that intention. They are identified by Roman numerals, which are referred to in the following parts of this treatise.

(I) How can one describe, explain and determine the influence of time pressure on the cognitive decision-making process?

(II) How can one identify the cognitive process and its components in strategic decision-making?

The answer to the second question enables describing and modeling the process and its components as well as identifying them in observed decision-making behavior.

(III) What behavioral patterns in the cognitive processes can be distinguished and how can they be classified?

(IV) How does time pressure affect behavior on a process-component level and concerning patterns?

Following Costa-Gomes et al. (2001) the present research interest is limited to initial behavior, omitting learning effects. This perspective is chosen to merely reduce the complexity of the research approach so far, even though learning surely matters in this context. To answer the four research questions considered, the author chose the procedure described below.

1.3 Research Approach

The research approach follows a deductive path, whereby first a model is derived from the literature, which describes, explains and predicts cognitive processes under time pressure. The implications of the model are elaborated by a simulation. The predictions thus derived are finally verified by an experiment that observes and analyzes the behavior of decision-makers.

First of all, the theoretical results and empirical evidence from the literature dealing with the objectives of this treatise are examined. For this purpose, a literature review is conducted.[2]

The approach of studying the cognitive process in strategic decision-making under time pressure is to the knowledge of the author not present in the literature yet. However, similar research intentions that concentrate on either (strategic) decision-making processes or time pressure in (strategic) decision-making exist. They are expected to offer valuable references in observing and modeling processes and their components as well as characterizing, interpreting and classifying observed behavior. The focus is laid on concepts from decision-making theory and, to a smaller extent, on psychological aspects. Neurological research that sets cognitive processes in a physiological framework is beyond the author’s point of view and thus widely not taken into account.

[2] The review followed some specifications: the search is carried out in the primary scientific online search engines SCOPUS, Science Direct, IEEE Xplore, Springer Link and Google Scholar. Reviewed articles in selected journals, as well as topic-relevant monographs and collected works, are taken into account. Literature in English and German language is considered. The journal selection is based on the recommendation of the VHB Teilranking Allgemeine Betriebswirtschaftslehre with rating A and better Verband der Hochschullehrer für Betriebswirtschaft e.V. 2018. The following keywords and keyword combinations, as well as their German equivalents, form the search terms: strategic decision-making, game theory, information processing theory, process tracing, decision rules, decision heuristics, elementary information processes, mouse tracking, time pressure, rationality and experiments in behavioral economics. Publication biographies of articles identified and selected are scanned for further relevant works based on their titles. Search results in journals are limited to recent publications (01.01.2015 - 31.12.2017) to record state of the art. For monographs and collected works titles with a higher citation index take precedence over those with a lower citation index. No time limit was considered here.

In this treatise, the structure of cognitive processes is studied, which are executed in strategic decision-making under time pressure conditions from Elementary Information Processes (EIPs). William G. Chase introduces EIPs as basic cognitive operations that can be modularly combined to form more complex operations, such as solving decision problems (Chase 1978, p. 19). The resulting sequence of EIPs can be interpreted as a cognitive process. However, the set of EIPs depends on the problem task observed.

On the basis of EIPs, a process model of decision-making in strategic tasks that considers time pressure conditions is described in this treatise. In the following, it is denoted as 'preparation time model'. The model's implications help to predict behavior under time pressure and formulate hypotheses for identified aspects of the cognitive process. It is at the same time the answer to the first research question.

To determine the set of relevant EIPs here, the author follows the process-theoretical approach intensively studied by Johnson, Payne, and Bettman (Johnson and Payne 1985; Payne et al. 1988; Bettman et al. 1990; Payne et al. 1992). They focus on common heuristics to infer on the EIPs in use. Heuristics represent solution concepts to specific problem tasks. Their normative character allows for using heuristics as benchmarks, as far as their performance in a particular problem task is known. In this treatise a set of common heuristics applicable to normal-form games is selected. The heuristics are formulated as a set of condition-action rules – a form proposed by Simon and Newell who called it "production systems" (Simon and Newell 1971, p. 156). Applying such a production system to a problem task results in a characteristic sequence of EIPs. This sequence represents the cognitive process of a decision maker who completely follows the instructions of such a heuristic. By analyzing all resulting sequences, the author infers on the basic set of EIPs in use when people solve problem tasks of this type. Also, the sequences of EIPs work as patterns to which the author compares human decision-making behavior and thus helps to classify observed behavior. This framework contributes to answer the second research question.

To develop benchmarks from the heuristics, it is essential to acquire information about their performance in strategic decision-making tasks under time pressure. For that purpose, the author simulates the behavior of human decision makers who strictly follow a heuristic's solution concept to solve the studied problem tasks under various time pressure conditions. The structure of the games is varied for simulation. To evaluate this dataset, the author defines performance indicators that consider effectivity and efficiency of process and choice. Based on this evaluation the author ranks the heuristics and derives reasonable recommendations towards their applicability under certain time pressure conditions. At the same time, the results from simulation help to formulate more precise hypotheses regarding decision-making behavior under time pressure. This contributes to answering the second research question.

Studying real decision-making by the processual approach requires identifying EIPs. Since EIPs are mental operations, they are not directly observable. Process tracing methods (PTM) need to be applied to infer on the cognitive processes. Following the approach of Costa-Gomes et al. (2001), the author selects mouse tracking. This PTM relies on the connection between mind and motor system (Freeman et al. 2011, p. 227). People use a mouse pointer while working on a problem. Mouse movement and actions are recorded, building a data stream of what is subsequently denoted as Elementary Manual Motor Acts (EMMAs).[3] EMMAs represent the motor system's actions. Those actions need to be interpreted in the light of the underlying problem task and its concrete visualization (Kühberger et al. 2011, p. 13). To link motor system actions to the cognitive system, the author develops a problem-centered interpretation concept. This concept allows for interpreting EMMAs as EIPs.

In addition, the author conducts an experiment to obtain a dataset from real decision makers. Finally, by applying the interpretation concept to this data set, one can infer the cognitive processes of the decision makers. The dataset is analyzed considering aspects which follow the preparation time model and results from simulation of heuristic-induced behavior under time pressure. The predictions of this model regarding identified aspects are tested with proper hypotheses. In the next step, the set is scanned for types of decision makers. The classification considers similarities in the behavioral aspects. These results contribute to answer the third research question. Finally, the author concentrates on significant changes of patterns under varying time pressure conditions. Analyzing these changes allows to describe and evaluate the effect of time pressure on decision-making processes and to answer the fourth research question in conclusion.

[3]. This notion follows Amosov (1967, p. 109) who used the term ‘elementary motor acts’ to refer to all human motor acts, not limited to the manual ones.

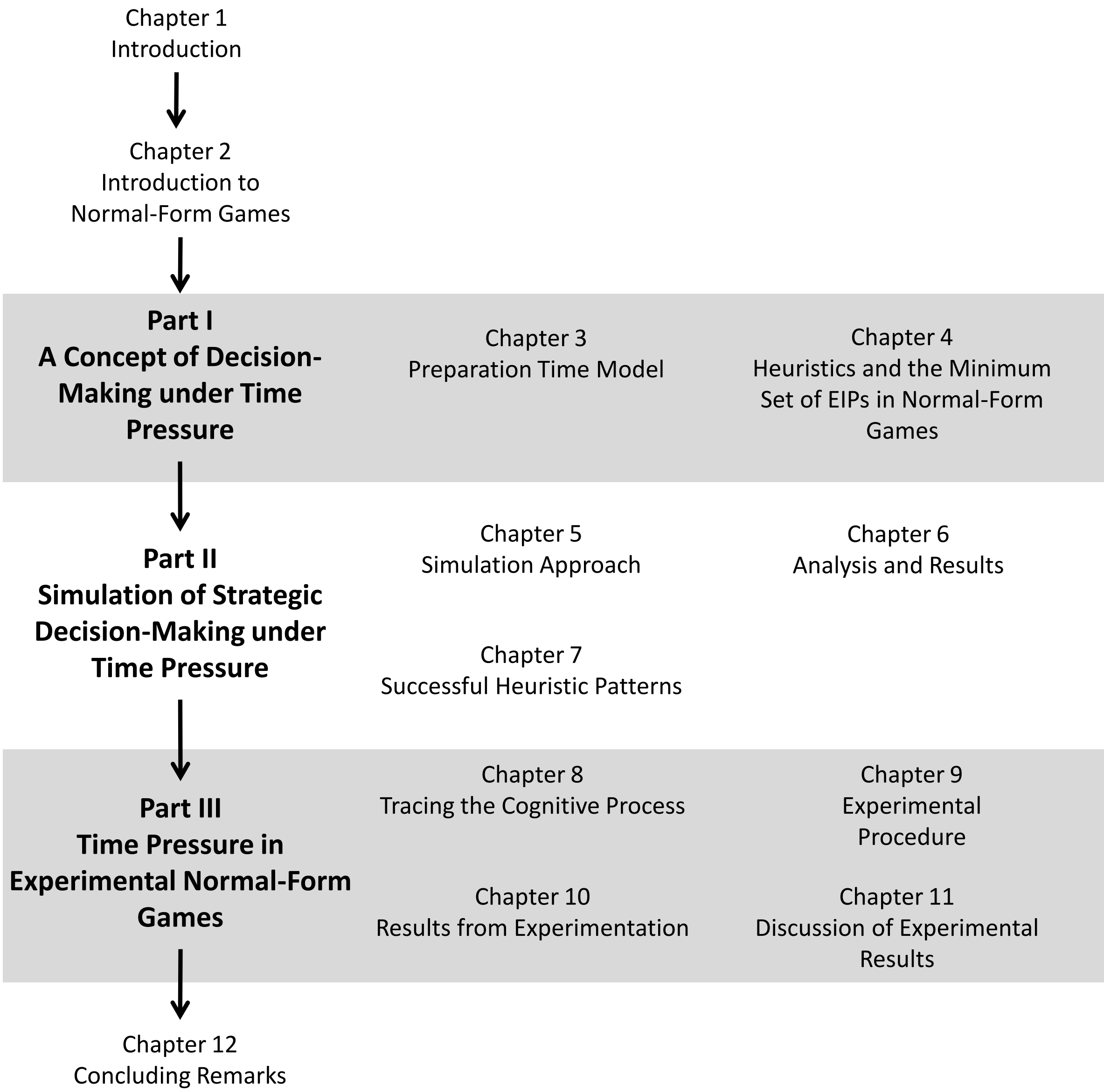
Overview of research approach:

1. Scan literature for approaches to explain strategic decision-making under time pressure on a cognitive process-based view.
2. Develop process model of strategic decision-making under time pressure from EIPs.
3. Select a set of common heuristics and define the corresponding process as well as production systems from EIPs.
4. Derive a minimum set of EIPs in use when solving strategic tasks from a set of heuristics.
5. Simulate application of heuristics to various strategic tasks under time pressure.
6. Develop a performance evaluation concept for processes.
7. Rank heuristics, recommend application scenarios and formulate hypotheses regarding behavior under time pressure.
8. Identify EIPs in real decision-making processes: develop an interpretation concept to link EMMAs to EIPs.
9. Experiment and evaluate interpretation concept with the acquired dataset.
10. Identify and classify patterns of the cognitive process to answer hypotheses of behavior.
11. Identify significant changes in patterns under time pressure.
12. Conclude by answering initial research questions.

1.4 Structure of Work

The structure of this treatise follows the proceeding described above. It comprises three major parts with interdependent content, as shown in Figure 1. The preliminary notes of Chapter 1 follows a short introduction to the decision-making tasks of game theory studied in this treatise (Chapter 2). In the subsequent Part I, a process-model of decision-making in strategic task environments under time pressure conditions is proposed, introducing the concept of preparation time (Chapter 3). This model intends to combine the resource-based approach of decision theory with elements of game theory, discussing relevant literature. Cognitive processes are modeled employing Elementary Information Processes (EIP). EIPs in use in the context of normal-form game tasks are identified and described in Chapter 4. Production systems of a set of commonly used decision-making heuristics in analogy to Simon and Newell (1971) are developed and analyzed to derive the minimum set of EIPs in use.

FIGURE 1 – STRUCTURE OF RESEARCH



Part II presents a simulation of heuristics' performances in strategic problem tasks under varying time pressure conditions. The set of heuristics considered was first analyzed by Costa-Gomes et al. (2001). Chapter 5 contains a description of the simulation, employing the production systems, introduced in Part I. Also, a performance evaluation concept the author developed for this case is presented. In Chapter 6 results of the simulation are presented and analyzed concerning the concept of preparation time. Successful heuristic information acquisition and processing patterns are discussed along with resulting elaborated predictions of the preparation time model in Chapter 7.

Part III comprises an experiment to evaluate the predictions from the preparation time model and from the simulation regarding the cognitive processes of decision-making. Since EIPs cannot be observed directly, the application of process tracing techniques is necessary. Chapter 8 gives an introduction to process tracing methods, comparing applicability and performance in strategic decision task environments. Here, the selected PTM mouse tracking is discussed in more detail. Consequently, a fitting interpretation concept for mouse tracking is developed and presented that enables interpretation of the observable data. It provides an econometric framework to convert mouse pointer movement information into EIPs and thus links observable behavior to the cognitive process.

In Chapter 9 the experimental procedure is presented. The experimental setup follows an axiomatic structure. The design is consequently derived from the requirements of the fundamental research questions and determines the technical implementation. Outcomes from experiment together with a comprehensive data analysis, especially hypothesis testing and classification of behavior, follow in Chapter 10. Results from the analysis are discussed in Chapter 11.

The final Chapter 12 is dedicated to a summary of essential conclusions drawn in this study, reflecting on the initial research intentions. It ends with an outline of questions that remained unanswered and further ideas regarding this study's objectives which are recommended for future research.

2 Introduction to Normal-Form Games

The fundamental idea behind game theory is to describe conflict situations among two or more parties in a mathematical model, which allows for a mathematical solution. Conflicts are understood as situations where the concerned parties pursue different (often opposite) interests (Vorob′ev 1972, pp. 13 f.). Each of the parties influences the result of the conflict through their respective decision. In game theory, such conflicts can be modeled as ‘strategic games’. A game is defined by three elements (Amann 1999, pp. 7 f.):

- the players representing the identified conflict parties,
- the players’ (finite) set of discrete alternatives representing their possible pure strategies of action, and
- the payoff function that determines the outcome for each combination of alternatives for each player.

Such a triple is called the normal-form of a game (Gintis 2009, pp. 41 f.). In case of all players make a decision simultaneously, such that no one can see (and react on) the action of any other player, the game can be displayed in matrix form.

An example is given in Figure 2 with two players, each having two pure strategies to choose from. In the example, the two players are denoted as row player and column player respectively. The row player has the two strategies ‘high’ and ‘low’, while the column player has the strategies ‘left’ and ‘right’. For each strategy combination of the two players’ sets of alternatives, their payoff is given in a matrix cell. From the semicolon-separated values, the row player owns the first and the column player the second as payoff, if this particular combination of strategies is realized. If the row player and the column player can interchange their identities without an effect on the game, it is called symmetric, otherwise asymmetric.

FIGURE 2 – TWO-PERSON NORMAL-FORM GAME

		column player	
		left	right
row player	high	3 ; 3	1 ; 4
	low	4 ; 1	2 ; 2

In games with complete information, players know all combinations of own and opponent's strategies as well as payoff functions from the beginning (Matsumoto and Szidarovszky 2016, p. 2). In case of denying negotiations and mediations among the players, a game is noncooperative (Matsumoto and Szidarovszky 2016, p. 2). Given such a game, game theory provides solution concepts to determine the best alternative for each player to maximize their payoff under the assumption that both players act rational.

Besides this fundamental type, many other games are modeled and studied in game theory. These differ not only regarding the number of players, the number of strategies or payoff functions. Furthermore, the basic assumptions are altered: the rationality of players, number and sequence of choices, payoff structures, and payoff information given in advance. For each of those games, normative solution concepts are developed and their applicability, as well as their application (from a descriptive point of view) in real life decision-making situations, are studied.

This brief introduction to the research field of game theory is considered to be sufficient for the purpose of this treatise. The tasks studied in the following are all of the same kind: strategic, non-cooperative two-person normal-form games with complete information. However, the numbers of strategies as well as the payoff information will differ among the games.

Part I: A Concept of Decision-Making under Time Pressure

This part comprises two chapters that are dedicated to studying processual aspects of decision-making. In Chapter 3, the concept of ‘preparation time’ is introduced. It offers an alternative interpretation to explain the decision-making in two-person noncooperative normal-form games under time pressure as a process. Furthermore, theoretical basics are developed that are necessary to identify and model cognitive processes that are applied in such decision-making tasks. The preparation time model, as well as the theoretical basis, are employed in the following two parts of this treatise. In Part II, they are used for simulating heuristics’ behavior and in Part III for developing a framework to interpret observable behavior as cognitive operations.

3 Preparation Time Model

This chapter introduces the concept of ‘preparation time’ for a game-theoretic decision-making environment. Two questions are in the center of discussion: Why is another decision-making model necessary to introduce the process-theoretical view to game theory? What advantageous implementations provides the new model? The answers to these two questions are crucial for Chapter 4 where the components to model and describe cognitive processes are defined.

3.1 Why a New Model?

In real life, there are many strategic decision-making situations where time to develop a decision (referred to as decision-making time in the following) is a scarce good (MacGregor 1993, p. 73; Rastegary and Landy 1993, p. 220). A decision maker is forced to lower the mental effort of thinking through the apparent problem, trading it against accuracy in choice (Johnson and Payne 1985, p. 395; Förster et al. 2003, p. 148). To develop a ‘good’ decision that also considers the opponent’s potential choices under such conditions could become quite a challenging task.

The body of literature concerning decision-making under time pressure is successively growing. However, comparatively little is published in the field of strategic decision-making. The label ‘strategic’ refers to its game-theoretically coined meaning, where the results of choice are determined by an interaction of at least two persons. Strategic choice refers to the result of a decision that reflects choosing the optimal strategic option. Strategic thinking (in literature sometimes referred to as strategic sophistication) is understood as a sophisticated analysis of the problem environment including interaction prior a decision.

Thinking is not always strategic, frequently resulting in a poor decision. This fact is common knowledge to all researchers that challenge the rationality paradigm of classic game theory.[4] However, how is inadequacy explained? Most behavioral studies presented in this context focus on determining the degree of rationality under certain conditions. They seldom provide in-depth analysis of potential reasons for the examined degree. Generally, excepted explanations are thus merely repeated. Camerer (2003), for example, offers the following behavioral components as reasons for bounded rationality, without adding a discussion or sources: understanding of the problem situation, social utility, limited iterated reasoning (which is more a description than a reason), learning (which could be cause and effect at the same time here) (Camerer 2003, pp. 23 ff), belief and expectations about contester and finally working memory (Camerer 2003, p. 252). Limited information processing capacities and deficits in problem-solving capability are some other frequent arguments (not stated from Camerer), which are in line with resource-based models of the economic and psychological theory.[5] Those reasons indeed seem elaborated and adequate. Nonetheless, the bottom line is: reasons for bounded rationality are seldom in focus of such studies since the imperfectness of the human decision-making capability is a too clear argument for explaining examined behavior rather than to address the reasons explicitly. Human's imperfectness is one apparent reason why time pressure is considered less frequently to explain non-rational behavior. Consequently, the quantitative effect of such reasons in these approaches is generally not investigated, let alone operationalized.

Studies of behavioral game theory thus often limit their focus to decision information. Strategic choices are assumed to prove strategic thinking. This conclusion is fallacious. To which extent choice reflects analysis is still subject to "many unresolved questions", as Costa-Gomes et al. (2001, p. 1193) put it. Essential questions dealing with the cognitive mechanisms of decision behavior remain unanswered: Lindner and Sutter (2013), for example, might be the only ones yet

[4] Rationality in this context classifies behavior that leads a subject to choose its best alternative (among all possible and available alternatives) concerning its preference structure. By definition, this ‘best alternative’ offers the highest expected utility ('Von Neumann–Morgenstern utility theorem'; discussed in Amann (1999, p. 6) for example). Irrational then is to call a choice that is formally based on all information available, yet differs from the nominally optimal strategy. Irrationality usually involves violating the laws of logic, for instance, transitivity of preferences. Between these two characteristics, numerous sub-categories are examined in human decision behavior.

[5] Psychologic studies present a more elaborated set of individual influence parameters on decision-making, such as mood and emotion. Some of those also include approaches to operationalize and quantify the impact of such motives.

who investigate personal sophistication in terms of cognitive hierarchy[6] in time-limited, strategic decision tasks. Relying on choice information only, they offer interesting experimental results. Still, some counterintuitive behavior cannot be explained satisfactorily by the employed model of decision-making and the corresponding classification of subjects.[7] Their study does not focus on the cognitive process. Consequently, Lindner and Sutter neither give a process-based explanation for the observed effects on behavior nor come forward with an own comprehensive concept to implement time pressure in strategic decision-making. However, investigating the influence of time pressure in the environment of game-theoretic tasks supposes that decision-making is a time-consuming procedure which could be affected by time limitation and occurs in a pre-stage of choice.

The influence of time pressure on decision-making cannot be denied. Among others, Johnson et al. (1993, p. 115) report in their study of behavioral adaptation under time pressure conditions. From observations in experiments, the authors identified two different types of decision makers. Each of them shows the typical system of reactions first reported by Miller (1960, p. 697) when adapting to time limitation in choice: 1. acceleration of processing and 2. information filtering. One of the two types showed a third reaction: 3. change of the information acquisition and processing strategy. The experimental findings implicate that one type is mentally more flexible than the other and able to adjust its problem-solving strategy in between. This qualitative response scheme finds confirmation for different decision tasks (e.g., Wright (1974, p. 560), Thomas and Weaver (1975, p. 366), Ben Zur and Breznitz (1981, p. 102), Payne et al. (1988, p. 579), Edland and Svenson (1993, p. 36), Zakay (1993, p. 60), Ariely and Zakay (2001, p. 197)). The hypothesis that claims the existence of different behavior types in strategic decision situations gains strong support through these findings.

[6] The Cognitive Hierarchy Model (CH-model) proposed by Camerer et al. (2004, pp. 862 ff.) explains decision behavior with individual levels of reasoning. In dominance solvable games those levels are equivalent to the number of rounds/iterations executed before choosing a (remaining) strategy. The chosen strategy is said to be directly linked to the level of reasoning. Principally, several game forms exist to investigate such levels. Besides those already mentioned, games in normal-form can be applied in this context (see for instance Costa-Gomes et al. (2001)). Such games can be designed flexibly according to the number of iteration cycles required. For an implementation example, see Gintis (2009, pp. 99 ff). This game form is moreover qualified for application in process-related research and thus offers a broader scope for behavioral studies in general.

[7] They observed a growing number of equilibrium decisions under decreasing time limitation and explained it by chance (Lindner and Sutter 2013, p. 544). The employed CH-model of Camerer et al. (2004, pp. 862 ff.) provides no answer to this phenomenon. It is further only partly able to capture heuristic-based behavior and related classifying, as Devetag et al. (2016, p. 199) determine in their experimental study.

Kahneman (2012) proposes a more general model of cognition, based on findings and observable effects of various psychological experiments. He postulates two mental operation systems that deal with problem-solving, and especially with decision making (pp. 33 ff.). Both systems are part of every (healthy) human brain. "System I" is affectional and fast working, driven by emotions or other than rational motives (in a game theoretic sense). "System II" is analytical, accurate and more or less slowly working. This model gives an adequate explanation of a variety of behavioral characteristics. However, the underlying processes of the systems working mechanisms remain unclear.

In real life, many observable decisions seem ill-conceived. Those decisions are probably the result of Kahneman's System I which by definition comes to a solution effortlessly (Kahneman 2012, p. 33). Strategic decision-making by analyzing or at least reading normal-form games' payoff matrices usually takes time and hence effort – properties also associated with processes.[8] Following Kahneman's concept, one can find the equivalence of System II in this information processing which is said to be the more effort-intensive way of thinking.

This model is helpful to explain the effort-dependent speed of decision-making and also the accuracy of choice. However, it does not offer a process perspective to study mental operations in more detail. Also, the time pressure conditions under which the two systems work remain unclear. It thus cannot be applied as a framework in this study. However, it serves as a reference when determining the effort of elementary cognitive operations in the context of developing a proper identification framework (Section 8.3).

Game theory provides very few approaches that question the period in which players develop their choices right before a game starts. The decision time is usually implicitly and seldom explicitly neglected.

Cooperative game theory, for example, knows the concept of "pre-play communication" to model players' collusion ahead of a game and to coordinate the equilibrium decision (Matthews and Postlewaite 1989, p. 239). A rather old concept from the field of learning which explicitly tackles the time frame prior to a decision is Brown's 'fictitious play'. It is fictively played for an infinite number of times to determine the optimal strategy in a game iteratively. Brown (1951, p. 375) neither addresses the time frame nor the duration for this learning in more detail.

[8] Although that seems to be true for game theorists, there might be people dealing with such a problem in no time by playing randomly and without taking further notice of the information presented in a payoff matrix.

Regarding preparation time, Lambertini (1998, p. 1) formulates in his paper more explicitly: "The choice […] occurs in a preplay stage [of the game] which does not take place in real time." Unfortunately, he gives no further explanation for this assumption. In fact, for real decision-making, the opposite must be true for the ‘preplay stage' of the game.

From this consideration of literature, two questions arise regarding decisions under time pressure which cannot be adequately explained by existing models: why does time pressure affect strategic decision-making? How does time pressure affect strategic decision-making?

3.2 Model Description

In this concept, preparation time is interpreted as the available time that players in strategic decision situations use to convert a given set of information into a decision. This period begins with getting aware of a decision situation and ends with a decision.[9] In cases of sufficient preparation time, this decision is similar to predictions of other game-theoretic decision models without time pressure conditions (all other things being equal). The other case, substantial time limitation and its effects on decision-making behavior, depicts the focus of the following consideration. In that context, time pressure is defined as effective time limitation.[10]

Probably the first paper that discusses decision behavior from a process perspective is Simon and Newell (1971). The authors studied human problem-solving methods, postulating the existence of "primitive information processes" (Simon and Newell 1971, p. 151) of which such a method is composed. Chase coined the name ‘elementary information processes' (EIPs) for these components and gave a more detailed characterization of certain applications (Chase 1978). The term ‘process’ could be misleading in this context since in the understanding of this treatise ‘process’ refers to the complete act of decision-making. Nonetheless, a decision-making process consists of primitive (or elementary) information processing operations that these authors designated as single processes.

[9] Note that in classic game theory this would be the optimal decision since players are regarded as utterly rational.
[10] Ordóñez et al. (2015, p. 520) use the terms 'time pressure' and 'time limit' in a different context, hence giving a different definition: in their notion time pressure describes the "subjective feeling of having less time than is required”. Time limits are “internally or externally imposed deadlines" (ibidem, p. 520). They use this distinction to study the perception of time limitation. It is thus essential to be aware of the definition of use in the context of this treatise. In Part II and III of this treatise, the definition is extended by adding task complexity. Complexity has a similar effect. The explanation is straightforward: To put pressure on the process, one could either limit the time available for a given task or one could increase the complexity of the task for a given time. In both settings, time per task is reduced. This reduction needs to be effective to meet the definition.

With their ‘production systems’ Newell and Simon provided a technique to develop a general definition of a problem-solving method based on EIPs (Simon and Newell 1971, p. 146). This way, decision-making processes could be modeled according to the problem task, and the stream of components could be predicted.

However, a process model of decision-making is not applicable to every cognitive task (Newell et al. 1958). As Bettman (1979) remarks, the information processing approach is a powerful and successful tool in modeling problem-solving behavior. It still requires well-structured problems “in which […] information available is fairly unambiguous and precise” (Bettman 1979, p. 8). Its applicability is thus limited to certain classes of problem tasks. This point is indeed important when discussing the decision task environment in the context of this treatise’s intent.

A simple and general model of decision-making processes with a focus on timely aspects is depicted in Figure 3. Not referring to a particular problem task, it provides all relevant properties to explain the concept of preparation time.

The figure has a horizontal time axis in the middle with two ticks labeled ‘start' and ‘decision/time limit' (black). The time in between is labeled ‘preparation time (= decision time available)'. Above the horizontal axis, two processes are depicted, each in a light-gray box. The above process belongs to a (fictional) heuristic i and the lower to a (fictional) mental process of a human decision maker.[11]

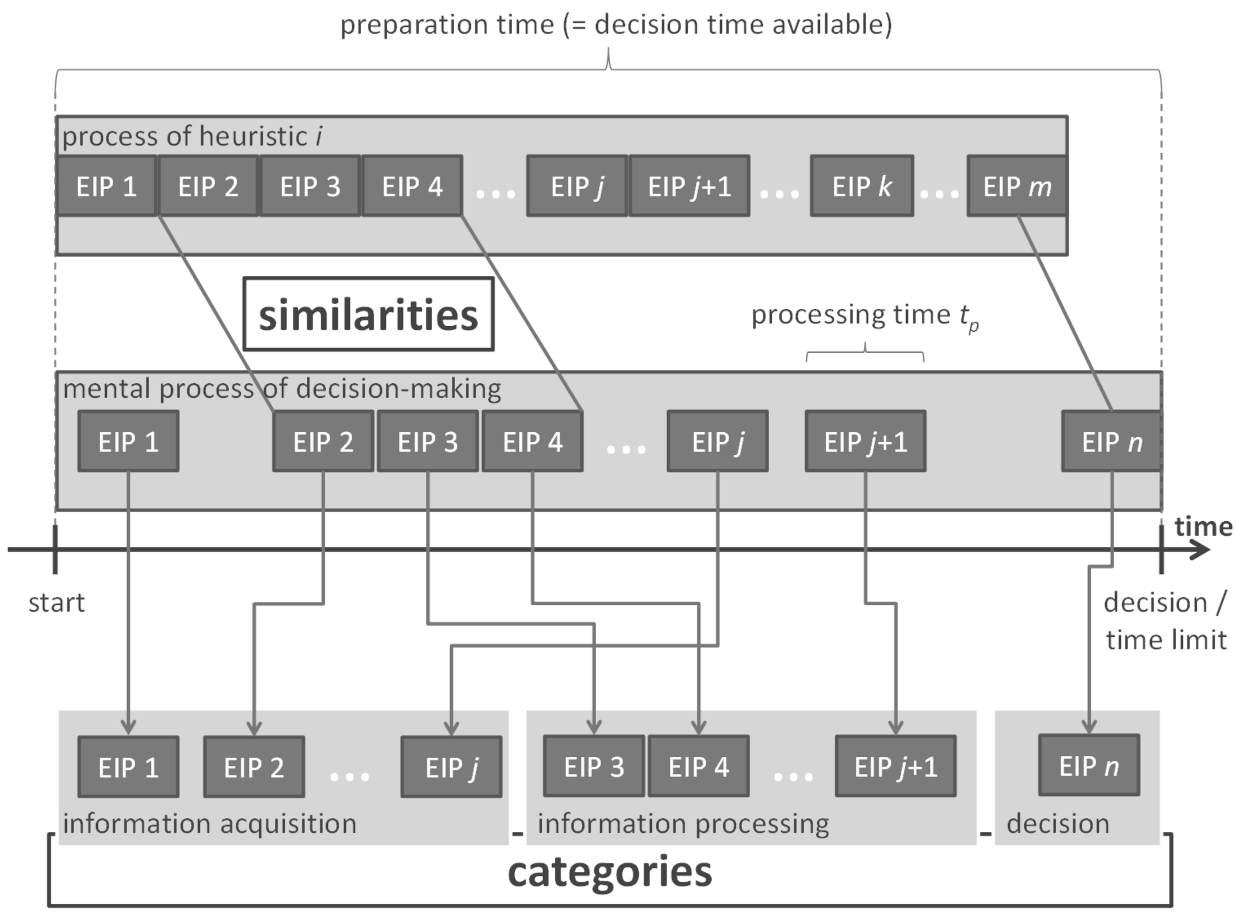
Each process is composed of process components, labeled ‘EIP' in a determined order and number (dark gray rectangles). The components symbolize specific mental activities (distinctions not explicitly depicted), each requiring a specific amount of time t_p larger zero in execution.[12] In the process of the decision maker, space is noted between components. This space represents time that cannot be matched with a mental activity. In the (normative) process of the heuristic, no such space exists: mental activities succeed without a pause until the time limit is exceeded or a decision is made. The duration of the processes (equal to decision time; depicted by the width of the box) usually depends on the number and types of components. The duration time could either exceed the time limit (decision maker) or be less than that (heuristic i). If the time limit is ex-

[11] Interpreting a heuristic as Simon and Newell (1971) primarily propose and Johnson and Payne (1985), Payne et al. (1988) and Costa-Gomes et al. (2001), among others, later adopt.

[12] Measuring time aspects of the process is a standard way to determine mental effort proposed in decision-making literature (Newell and Simon 1972). Time is regarded as an accurate and easy-to-measure substitute to estimate resource consumptions of mental processes (Johnson and Payne 1985, p. 396).

ceeded, the process stops without a decision or with a forced decision, according to the selected definition. Components can contribute to different aspects of the process. In this simple process model of decision-making, those aspects are categorized as 'information acquisition', 'information processing' and 'decision (/choosing an alternative)' following Hoffrage (1999, p. 142) and Kühberger et al. (2011, p. 3). When comparing both processes, similar parts might occur that contain some similar components in same order, which is often referred to as 'patterns' in the literature (for instance Payne et al. (1988, p. 569) and Costa-Gomes et al. (2001, p. 1196)).

FIGURE 3 – PROCESS MODEL OF DECISION-MAKING TIME



With this model on hand, the two questions that concluded Section 3.1 can be answered:[13]

1. Why does time pressure affect strategic decision-making?

[13] In the following, time pressure is assumed to be effective. I.e., the time limit is smaller than the duration of the decision process without time pressure.

Lowering the time limit is one option to increase time pressure on a decision maker. Hereby, less time is available to execute the process. Thus, the process could either not be finished or is altered. That in turn expectably affects the resulting choice.

2. How does time pressure affect strategic decision-making?

   If a decision maker is aware of a time-limited task, the process is likely to be adjusted, according to findings of Payne et al. (1988, p. 43). Differences are likely to occur regarding component structure (possibly changing problem-solving strategies and thus altering type and order of components) as well as processing time of components (potential acceleration).

If a decision maker is unaware of a time limit when processing a task, the behavior is not necessarily adjusted. However, since time pressure is effective (see footnote 13 on page 17), preparation time (available) is smaller than decision time (necessary). Thus, the process terminates before it regularly finishes. The resulting choice is determined according to a value that is defined by the termination criteria.

This model can be customized for a variety of applications by specifying components, heuristics or termination criteria. Indeed, the process structure is expected to depend on the given problem task (to which the decision behavior is adjusted). Even the assumed singularity in processing is changeable by allowing concurrency. Its illustrative presentation of the processes' time consumption as well as its component structure supports understanding of the process.

When interpreting the available decision time as preparation time, in which decision-making takes place, the influence of time pressure is evident. Also, by interpreting the preparation as a cognitive process, where information is acquired and processed until a choice is selected, time is consumed. Time pressure thus has an interpretable effect on any of the stages of the process. Following this, time pressure consequently affects choice, which is the result of the process. That, in turn, allows for discussing time pressure as a reason for poor choices and thus as an alternative, at least supplemental explanation for limited rationality. Also, as long as the cognitive process can be determined, the impact of time pressure on decision-making can be operationalized.

3.3 Model's Implications

In this section, the preparation time model is used to develop hypotheses about cognitive processes in time pressure decision tasks. These hypotheses are further tested in simulation (Part II) and experiment (Part III).

Two necessary and reasonable implications follow from the process model. First, cognition can be interpreted as a sequence of information processing steps. This interpretation supports the heuristic approach, which suggests an algorithm-based problem-solving. Observed behavior can thus be compared sequence-wise or even component-wise to heuristics. Second, numbers, types, and duration of the single steps relate to time consumption.

Ben Zur and Breznitz (1981, p. 43) and Edland and Svenson (1993, p. 38) name four steps people sequentially adopt in decision tasks under growing time pressure. The two implications help to explain the behavior from a process perspective:

1. *Accelerating execution* is the result of decreasing time per EIP up to a specific – most likely physically induced – minimum.
2. *Information filtering* means omitting single EIPs or sequences of EIPs that deal with information acquisition. Those can either be redundant, hence unnecessary, or crucial for the applied process – with particular consequences in choice.
3. *Changing problem-solving strategies* means altering decision-making methods during the process. Therefore, the process is interrupted. The process resumes after a certain amount of time with a different problem-solving method. Patterns of EIP-sequences are expected to differ, compared to the ones observed before the interruption. The probability of a change directly links to the remaining task time.
4. *Canceling task processing* is equivalent to an unfinished execution of the underlying problem-solving procedure. Canceling is in principle possible at any time within the process. Usually, however, it can be observed more often with narrower time limits. If the time limit is known and monitored during the cognitive procedure, a choice is either made as the ultimate action before the process terminates. In this case, the decision relies on the information basis already processed. Alternatively, the choice is omitted, and the sequence of EIPs terminates abruptly. The choice rate would decrease in that case.

Generally, the number of EIPs should be reduced in cases of time limitation compared to processes without such restrictions. However, acceleration can add to compensate for this issue.

Such a purely qualitative consideration needs to be specified quantitatively by other means. The apparent difficulty remaining here is to develop a concept of identifying EIPs in a decision behavior data stream. Such identification requires a preliminary step of specifying the EIPs that are most likely applied in such cases. The development of a minimum set of EIPs in use under such conditions is subject to the following chapter.

4 Heuristics and the Minimum Set of EIPs in Normal-Form Games

Understanding the components of a process in strategic decision-making is necessary to develop means for identifying them from observed behavior. It also represents the basis for modeling and studying such processes. In this chapter, the concept of EIPs is presented in more detail. It is further described, how a possible set of EIPs, applied in normal-form games, can be derived from heuristics. In the understanding of this work, heuristics can be defined as simple decision rules that help to identify an adequate alternative quick and with little effort. They usually reduce the complexity of a given problem task by only considering a subset of the task information. In doing so, the decision-making speed increases, but the prediction accuracy is limited (Kaplan et al. 1993, p. 256).

The selection of heuristics in this treatise is mainly identical to that of Costa-Gomes et al. (2001), who identified a set of heuristics commonly in use when solving strategic problem tasks (Costa-Gomes et al. 2001, p. 1205). Production systems are developed from the normative description of these heuristics, using the method introduced by Newell and Simon (1972). This form of representation allows direct access to EIPs applied in heuristic processes. Finally, a minimum set of EIPs is identified from the production systems of the set of heuristics. This minimum set finds application in the following parts of simulation and experimentation.

4.1 Elementary Information Processes

Elementary information processes (EIPs) are understood as basic cognitive operations humans mentally employ when thinking (Chase 1978, p. 19). The set of employed EIPs is related to the problem task with which the process of thinking is concerned. Complex operations, such as solving decision problems, can be formed by a modular combination of EIPs from the set (Huber 1980, p. 188; Johnson and Payne 1985, p. 398). The resulting sequence of EIPs is interpreted as a cognitive process. EIPs cannot be observed directly. However, they can be studied by employing process tracing techniques (PTTs). PTTs record observable behavior data. This data is interpreted to infer on the EIPs in use. Interpretation requires a framework that considers the problem task.

To generate a framework that supports the interpretation of behavioral data with respect to the cognitive process, a closed set of EIPs is needed. This set shall comprise EIPs that are part of a representative variety of strategies regarded as typical in human decision-making behavior.[14] Johnson and Payne (1985) developed their set of operations ('J&P'85 set') by following suggestions of Chase (1978) and Huber (1980, p. 398). Chase (1978) proposed more abstract EIPs, for instance, mental rotation, visual search and memory search, which better serve as categories of EIPs than real mental processes. Huber, a psychologist who supported the idea of cognitive processing in specific decision environments, searched for a common theoretical basis for simple decision rules. He found several behavioral problems unresolved regarding their general applicability, motivating his studies. His goal was to develop a "psychological theory of decision-making" (Huber 1980, p. 187). Employing the process tracing technique "Thinking Aloud Protocols"[15], he assumed the existence of two kinds of EIPs. He described general EIPs, such as SELECTING (a piece of information) and CHECKING (a condition). In addition to that, he defined decision process-specific EIPs, such as CONCATENATE (meta information), MAX, MIN, DIFF(erence) or EVALUATE (numerical variables), and CRITERION (of acceptance). Johnson and Payne (1985) neither followed that classification nor adopted much of that set of EIPs. They reported no reasons, but one could argue that especially EVALUATE and CRITERION (excluded from the J&P'85 set) are technically incompatible with the later applied mouse tracking approach. Thus, their resulting set consists of EIPs as shown in Table 1.

TABLE 1 – DEFINITION OF EIPS BY JOHNSON AND PAYNE (1985)

EIP	**Definition**
READ	Read an alternative's value of an attribute into Short Term Memory (STM)[16].
COMPARE	Compare two alternatives for an attribute.
DIFFERENCE	Calculate the size of the difference between two alternatives for an attribute.
ADD	Add the values of an attribute in STM.
PRODUCT	Weight one value by another (i.e., multiplying).
ELIMINATE	Remove an alternative from consideration.
MOVE	Go to next element of the external environment.
CHOOSE	Announce preferred alternative and stop the process.

[14] To the knowledge of the author, there is no such an attempt for a strategic environment so far.
[15] The idea of Thinking Aloud Protocols is to ask a subject what it is thinking while fulfilling a task, hoping that it verbalizes every conscious step of its inner process. A detailed introduction to this PTT will be given in Chapter 8.
[16] Even if it primarily remains a model, the notion of addressing human's ability to store information for different durations is widely accepted (Ericsson and Kintsch 1995, p. 234) and hence is followed here.

4.2 Heuristics and Corresponding Production Systems

Following Costa-Gomes et al. (2001), the present study deals with decision-making in the strategic environment of two-person normal-form games of complete information. The focus is laid on the decision-making process as an initial answer to a so far unknown problem task. Costa-Gomes et al. (2001) present a set of heuristics they assume to be most common, identified in several behavioral studies or at least suggested by other researchers. With their selection, they aim for a proper diversity to describe "decision [making] and information searches without overly constraining the data analysis, yet small enough to avoid overfitting" to gain a maximum of explanatory power (Costa-Gomes et al. 2001, p. 1205). Several papers are suggesting other behavioral patterns. However, those are either differing just slightly regarding information processing and hence apply a similar set of EIPs (e.g., Stahl and Wilson (1995)). Alternatively, they simply not suit this kind of decision tasks (e.g., Tversky (1969), Tversky (1972), Bettman (1979), Thorngate (1980)).

The following heuristics and their descriptions are in reasonable compliance with the suggestions of Costa-Gomes et al. (2001). Table 2 gives an overview of the heuristics considered and of selected aspects. The heuristics are classified as either strategic or nonstrategic. Nonstrategic heuristics stated are *Random*, *Altruism*, *Optimism*, *Pessimism*, and *Naïve*. On the side of strategic heuristics, one can find *Equilibrium* and *Sophisticated* expanded by the bounded rational strategic heuristics *D1*, *D2*, and *L2*. A total of ten heuristics serve as analysis pool to identify the underlying set of EIPs. The strategic heuristics are in line with the level-k-strategy approach of the cognitive hierarchy model proposed by Camerer et al. (2004, pp. 862 ff.). Several experimental studies so far employ them (Sutter et al. 2003; Camerer et al. 2004; Reutskaja et al. 2011; Arad and Rubinstein 2012; Lindner and Sutter 2013; Lindner 2014; Chen et al. 2014).

In the following the heuristics are described regarding objective and procedural aspects, resulting EIPs included. The procedure of information acquisition, processing and choice recommendation when applied to normal-form games is presented. Even though not necessary for the goals of identifying the underlying set of EIPs, their description implies the concept of time limitation behavior. It includes the behavior that follows a termination through time limitation. By suggesting a more risk-averse behavior, especially nonstrategic heuristics rather rely on information which has already been gathered than on chance. That means, in case of exceeding a time limit, those heuristics choose from alternatives which are at least partly observed, neglecting the unobserved ones. However, it is questionable whether this leads to an effective decision in terms of maximizing payoffs. Besides that, the procedure is a simplification, since other attitudes towards risk are thinkable.

TABLE 2 – SET OF HEURISTICS AND SELECTED CHARACTERISTICS

Heuristic	**Classification**	**Goal**	**Use of information**	**Level of reasoning**
Random	Nonstrategic	Choose	None	0
Altruism	Nonstrategic	Max social payoff	All	0
Optimism	Nonstrategic	Max own payoff	Own payoff	0
Pessimism	Nonstrategic	Max minimal own payoff	Own payoff	0
Naïve	Nonstrategic	Max own payoff	Own payoff	1
L2	Strategic	Max own payoff	All	2
D1	Strategic	Max own payoff	All	1
D2	Strategic	Max own payoff	All	2
Equilibrium	Strategic	Max own payoff	All	∞
Sophisticated	Strategic	Max own payoff	All	∞

Nonetheless, it is helpful, since it reduces the behavioral degrees of freedom and hence the amount of investigation. Finally, this simplification can be addressed in forthcoming studies. The implementation of an ultimate action taken at the moment the time expires is useful for later investigations regarding heuristics' performance under time pressure conditions.

Along with each heuristic of the given set, a production system is presented. The development process of those production systems follows suggestions of Newell et al. (1972), who verbalize the execution of a heuristic as a condition/action set in the following form:

“If condition X is met, do action Y.”

The sequence of conditions within the set is the actual order of testing the conditions. Actions shall be interpreted as fundamental as possible (i.e., basic steps). A claim limiting this proposal says that a single step should build a unit of meaning for the heuristic and should substantially serve the execution of the heuristic's procedure. After one condition is met, and an action took place, the set is virtually scanned again from the beginning. The set is complete in the sense that undefined conditions within the production system are invalid. The heuristics' intention and commonly used description depict the starting points of the production systems' development. The heuristics are applied to a normal-form game task, recording the steps taken for each condition. Finishing the task means finishing the execution of the production system at the same time. As mentioned above, the production systems consider the time limit condition for future purpose in the context of the present work. This condition requires a random decision in some cases. In this treatise all random decisions that are designed in the heuristics' production systems are based on discrete uniform probability distributions: each of the m (remaining) alternatives is chosen with a probability of $1/m$.

The production systems of the corresponding set of heuristics are presented in two-column tables. The interpretation is as follows: the satisfaction of the condition (left column) triggers a corresponding action (right column).

The *Random* heuristic is the only one of the ten mentioned that does not necessarily require knowledge of the payoff sets. Alternatives available are chosen with the same probability, without reducing the set of alternatives. In game theory, such choice behavior is also referred to as nature-like play, implying that the outcome of a decision is driven by chance. Assume that a player designated as 'Row' has n and a player 'Column' has m pure strategies in a normal-form game. In this case, there are $n * m$ cells in the payoff matrix and a combined $2 * n * m$ payoff information set. In its extreme version, the *Random* heuristic considers some information subset sized between zero and $2 * n * m - 1$ before choosing a strategy. In fact, applying this heuristic there is no need to look up all given payoff information. Regardless of the later defined production set when analyzing experimental data, not using all presented information is always a proper hint of a *Random* choice. Of course, it is possible to process all information, but that – by definition – does not influence the decision itself. To minimize the effort of the cognitive process, no information should be processed. It is hence expectable that the amount of applied EIPs in processing the production system remains constant at a shallow level, regardless of changing time pressure parameters. Table 3 shows the production system of *Random*.

TABLE 3 – PRODUCTION SYSTEM OF HEURISTIC *RANDOM*

If an alternative is chosen,	**END game.**
If the game starts,	CHOOSE an alternative randomly.

The heuristic *Altruism* aims at the maximization of the two persons' combined payoff, also referred to as social payoff. It is pseudo strategic since it takes the opponent's payoff into account (in exclusively noncooperative games).[17] The heuristic schedules to check any payoff information given. The order of checking the cells of the payoff matrix could be arbitrary. A schematic procedure starting in any of the matrix' corners and checking consecutive cells of either row or column direction should minimize the effort. If it comes to an early termination of the procedure through effective time limitation, the production system produces in its final step a random decision between all alternatives. Alternatively, if already determined, the alternative which contains the current maximum is chosen. Table 4 presents the design of the production system following this idea.

TABLE 4 – PRODUCTION SYSTEM OF HEURISTIC *ALTRUISM*

If an alternative is chosen,	**END game.**
If time limit is reached and current max = empty,	CHOOSE an alternative randomly.
If no new alternative or time limit reached,	CHOOSE alternative containing current max.
If no new cell in alternative,	MOVE to next new alternative.
If COMPARE finished,	MOVE to next new

[17] This heuristic is regarded as non-strategic since no reasoning for the behavior of the other player in a strategic sense is scheduled. An altruistic player implicitly assumes that the opponent behaves altruistic, too (Costa-Gomes et al. 2001, p. 1195, footnote 5).

	cell.
If payoff sum cell > current max,	READ (Store) payoff sum cell as new current max. COMPARE finished.
If payoff sum cell < current max,	ELIMINATE cell. COMPARE finished.
If new cell READ & ADD,	COMPARE to current max.
If new cell open,	READ and ADD payoff pair.
If new cell closed,	OPEN new cell.
If the game starts,	MOVE to the first alternative, first cell. Current max = empty.

The *Optimism* heuristic focusses on the payoff of one player exclusively, since it looks for the maximum in this payoff set. That means every payoff cell is at least visited once before a decision can be chosen. An also common name for this heuristic is the *Maximax* decision rule (Costa-Gomes et al. 2001, p. 1195). In case of an early termination through time limitation, a random decision is made or, if at least one is already determined, the strategy which contains the current maximum is chosen. The corresponding production system is presented in Table 5.

TABLE 5 – PRODUCTION SYSTEM OF HEURISTIC *OPTIMISM*

If an alternative is chosen,	**END game.**
If time limit is reached and current max = empty,	CHOOSE an alternative randomly.
If no new alternative or time limit reached,	CHOOSE alternative containing current max.
If no new cell in alternative,	MOVE to next new alternative, new cell.
If COMPARE finished,	MOVE to next new

	cell.
If own payoff cell > current max,	READ (Store) own payoff cell as new current max. COMPARE finished.
If own payoff cell < current max,	ELIMINATE cell. COMPARE finished.
If new cell READ,	COMPARE to current max.
If new cell open,	READ own payoff.
If new cell closed,	OPEN new cell.
If the game starts,	MOVE to the first alternative, first cell. Current max = empty.

Same as the *Optimism* heuristic, the *Pessimism* heuristic only focus on one player's payoff aiming to maximize the minimal strategic outcome. That is optimally done through a strategy-wise search of every payoff cell, working out the minimal outcome of each strategy and choosing the strategy where the minimum is maximal. An also common name for this heuristic is the *Maximin* decision rule (Costa-Gomes et al. 2001, p. 1195). The behavior in case of early termination is equivalent to the one of *Optimism*. The corresponding production system is presented in Table 6.

TABLE 6 – PRODUCTION SYSTEM OF HEURISTIC *PESSIMISM*

If an alternative is chosen,	**END game.**
If time limit is reached and current max = empty,	CHOOSE an alternative randomly.
If no new alternative or time limit reached,	CHOOSE alternative containing current max.
If COMPARE finished,	MOVE to next new alternative, new cell.
If current min < current max,	ELIMINATE current alternative. COM-

	PARE finished.
If current min > current max,	READ current min as current max. (Store containing alternative). COMPARE finished.
If no new cell in alternative,	COMPARE to current max.
If COMPARE finished,	MOVE to next new cell.
If payoff cell < current min,	READ (Store) payoff cell as new current min. COMPARE finished.
If payoff cell > current min,	ELIMINATE cell. COMPARE finished.
If new cell READ,	COMPARE to current min.
If new cell open,	READ payoff.
If new cell closed,	OPEN new cell.
If the game starts,	MOVE to the first alternative, first cell. Current max = empty. Current min = empty. CHOOSE focus=own payoff.

The *Naïve* heuristic is pseudo-strategic since it takes the opponent's behavior into account. However, this behavior is simply expected to be a random choice with an underlying uniform probability distribution. That means the heuristic optimizes a player's outcome against a nature-like behaving player. Thus, it is not necessary to look up the opponent's payoff using this heuristic. The easiest way to proceed is summing up the own strategies' payoffs and choosing the strategy with the maximum sum. If the process is interrupted by time limitation without having evaluated a single alternative completely, a random decision considers all alternatives available with same probability. Else, the alternative which contains the current maximum is chosen. Following this idea, the production system for this heuristic is as shown in Table 7.

TABLE 7 – PRODUCTION SYSTEM OF HEURISTIC *NAÏVE*

If an alternative is chosen,	**END game.**
If time limit reached AND alternative containing current max exists,	CHOOSE alternative containing current max.
If time limit reached,	CHOOSE an alternative from remaining alternative set randomly.
If no new alternative,	CHOOSE alternative containing current max.
If COMPARE finished,	MOVE to next new alternative, new cell.
If current sum < current max,	ELIMINATE current alternative. Current sum = 0. COMPARE finished.
If current sum > current max,	READ current sum as current max. (Store containing alternative.) Current sum = 0. COMPARE finished.
If no new cell in alternative,	COMPARE to current max.
If ADD finished,	MOVE to next new cell.
If new cell READ,	ADD to current sum.
If new cell open,	READ own payoff.
If new cell closed,	OPEN new cell.
If the game starts,	MOVE to the first strategy, first cell. Current sum = 0. Current max = empty.

Naïve concludes the set of nonstrategic heuristics discussed here. In the following paragraphs, the focus shifts to the set of so-called strategic heuristics. In opposition to nonstrategic heuristics, all of them take the opponent's action into account to find the optimal strategy. These heuristics are presented in their order of increasing sophistication.

The heuristic *L2* is the best response to *Naïve*. Thus, it is of bounded rationality. In the beginning, the production system (presented in Table 8) is similar to the one of *Naïve* to detect the strategy of a potential *Naïve*-playing opponent. In the next step, it develops the best answer towards that strategy. If it comes to an early termination of the process through time limitation, a random decision between the remaining set of alternatives is made. The set, in fact, will not be reduced until the opponent's alternative is determined and at least two of the own alternatives are compared towards their outcome against the opponent's choice. The reduction seems remarkable late compared to all other heuristics, which aim at a reduction of alternatives and will be subject to further studies addressing the performance of heuristics under time pressure conditions.

TABLE 8 – PRODUCTION SYSTEM OF HEURISTIC *L2*

If strategy is chosen,	**END game.**
If time limit reached and current max = empty,	CHOOSE an alternative from remaining alternative set randomly.
If no new cell in max_strat or time limit reached,	CHOOSE own strategy containing current max.
If own payoff < current max,	ELIMINATE cell containing own payoff.
If own payoff > current max,	own payoff = current max, MOVE to next cell in max_strat.
If cell of max_strat READ,	COMPARE own payoff to current max.
If max_strat stored,	MOVE to first cell of max_strat.
If no new strategy,	STORE opponent

	strategy containing current max as max_strat, all cells of max_strat = new cells, CHOOSE focus = own payoff, current max = empty.
If COMPARE finished,	MOVE to next new strategy, new cell.
If current sum < current max,	ELIMINATE current strategy. Current sum = 0. COMPARE finished.
If current sum > current max,	READ current sum as current max. (Store containing strategy.) Current sum = 0. COMPARE finished.
If no new cell in strategy,	COMPARE to current max.
If ADD finished,	MOVE to next new cell.
If new cell READ,	ADD to current sum.
If new cell open,	READ focussed payoff.
If new cell closed,	OPEN new cell.
If the game starts,	MOVE to first opponent strategy, first cell. Current sum = 0. Current max = empty. CHOOSE focus= opponent payoff.

Heuristic *D1* symbolizes bounded rationality with a player having a cognitive hierarchy equivalent to a reasoning level of $k = 1$. Hence, one round of iterated elimination of dominated strategies is played. Out of the remaining strategies the one is chosen that reflects the best answer to a decision an opponent takes who plays his non-eliminated strategies with equal probability. That procedure, in turn, is equivalent to *Naïve's* choice determination process. The production system for the case of starting to search for dominated strategies in one's own strategy set is shown in Table 9.

TABLE 9 – PRODUCTION SYSTEM OF HEURISTIC *D1*

If strategy is chosen,	**END game.**
If time limit reached and alternative containing current max exists,	CHOOSE alternative containing current max.
If CHANGE_count = 2 (= if focus changed twice)[18] or time limit reached,	CHOOSE a strategy from remaining strategy set randomly.
If CHANGE_count = 1, (= if focus changed once)[19]	Current max = empty. MOVE to first (nearest) strategy, first cell. First strategy = current strategy. First cell = current cell. Next MOVE = MOVE 1.
If all strategies compared to each other,	CHANGE focus. CHANGE_count = CHANGE_count + 1.
If all payoffs in strategy compared,	ELIMINATE strategy not containing max. Next MOVE = MOVE to next new strategy, new cell.
If no new cells in strategy,	all payoffs in strategy compared.
If MOVE 3 + MOVE 2 not possible,	no new cells in strategy.
If COMPARE finished,	do next MOVE.
COMPARE_strategies(a,b):	
case a = b:	next MOVE = MOVE 3 + MOVE 2. COMPARE_strategies finished.
case else:	ELIMINATE payoff and strategies from STM. Next MOVE = MOVE to next new strategy, new cell.

[18] If focus changed two times, elimination is finished and a strategy is to be chosen from the remaining set.
[19] After the first round, variables are initialized for round two.

If two max containing strategies in STM,	COMPARE_strategies(strat1, strat2).
COMPARE_payoff(a,b):	
case a>b:	max = a. COMPARE_payoff finished.
case else:	max = b. COMPARE_payoff finished.
If two payoffs in STM,	COMPARE_payoffs(payoff_1, payoff_2). STORE max containing strategy in STM. ELIMINATE payoffs from STM.
If one payoff in STM,	next MOVE = MOVE 1. Do next MOVE.
If cell READ,	STORE payoff in STM.
If cell open,	READ payoff.
If payoff focus = opponent,	FOCUS on opponent payoff data. MOVE 1 = horizontal to new cell in next strategy, MOVE 2 = vertical to new cell in current strategy, MOVE 3 = −(MOVE 1).
If payoff focus = own,	FOCUS on own payoff data. MOVE 1 = vertical to new cell in next strategy, MOVE 2 = horizontal to new cell in current strategy, MOVE 3 = −(MOVE 1).
If cell closed,	OPEN cell.
If the game starts,	CHOOSE payoff focus. Current max = empty. MOVE to first (nearest) strategy, first cell. First strategy = current strategy. First cell = current cell. CHANGE_count = 0.

The algorithm of *D2* works analogously to *D1*, except having two rounds of iterated elimination of dominated strategies. For games with one dominated strategy, *D1* and *D2* lead to the same result. One round of iteratively eliminating dominated strategies implies the following procedure: for both players each, eliminate the strategies that are dominated by at least one strategy of the same set. That is done by pairwise comparison of strategy payoffs, for two strategies each. The second round begins with restructuring the remaining strategy set for both players. That is a restriction and implies normative behavior since it excludes the possibility that one player is eliminating just one strategy (the first recognized) and immediately begins the second round by building up the remaining strategy set. That could be effective if no other dominated strategy is present in that current round since no more strategies need to be inspected. Nonetheless, this possible behavior is not affecting the minimum set of EIPs and hence could be neglected.[20] In case of an early termination through effective time limitation, a random choice is made between the remaining alternatives. The production system, in case of starting to search for dominated strategies in one's own strategy set, is shown in Table 10.

TABLE 10 – PRODUCTION SYSTEM OF HEURISTIC *D2*

If strategy is chosen,	**END game.**
If time limit reached and alternative containing current max exists,	CHOOSE alternative containing current max.
If CHANGE_count = 4 or time limit reached,[21]	CHOOSE a strategy from remaining strategy set randomly.
If CHANGE_count < 4,	Current max = empty. MOVE to first (nearest) strategy, first cell. First strategy = current strategy. First cell = current cell. Next MOVE = MOVE 1.[22]
If all strategies compared to each other,	CHANGE focus. CHANGE_count = CHANGE_count + 1.
If all payoffs in strategy compared,	ELIMINATE strategy not containing max. Next MOVE = MOVE to next new strategy,

[20] With a proper set of EIPs, this behavior is likely to be detected in a potential experiment database.
[21] This is true if focus changed twice. In that case, elimination is finished, and an alternative is selected from the remaining set.
[22] Initialization for round two.

	new cell.
If no new cells in strategy,	all payoffs in strategy compared.
If MOVE 3 + MOVE 2 not possible,	no new cells in strategy.
If COMPARE finished,	do next MOVE.
COMPARE_strategies(a,b):	
case a = b:	next MOVE = MOVE 3 + MOVE 2. COMPARE_strategies finished.
case else:	ELIMINATE payoff and strategies from STM. Next MOVE = MOVE to next new strategy, new cell.
If two max containing strategies in STM,	COMPARE_strategies(strat1, strat2).
COMPARE_payoff(a,b):	
case a>b:	max = a. COMPARE_payoff finished.
case else:	max = b. COMPARE_payoff finished.
If two payoffs in STM,	COMPARE_payoffs(payoff_1, payoff_2). STORE max containing strategy in STM. ELIMINATE payoffs from STM.
If one payoff in STM,	Next MOVE = MOVE 1. Do next MOVE.
If cell READ,	STORE payoff in STM.
If cell open,	READ payoff.
If payoff focus = opponent,	FOCUS on opponent payoff data. MOVE 1 = horizontal to new cell in next strategy, MOVE 2 = vertical to new cell in current strategy, MOVE 3 = −(MOVE 1).
If payoff focus = own,	FOCUS on own payoff data. MOVE 1 = vertical to new cell in next strategy, MOVE 2 =

	horizontal to new cell in current strategy, MOVE 3 = −(MOVE 1).
If cell closed,	OPEN cell.
If the game starts,	CHOOSE payoff focus. Current max = empty. MOVE to first (nearest) strategy, first cell. First strategy = current strategy. First cell = current cell. CHANGE_count = 0.

The production system of the heuristic *Equilibrium* is quite similar to the one of *D1* or *D2*. Only the abort criterion differs, since the algorithm does not end after one or two elimination rounds, but if elimination is no longer possible. In case of effective time limitation, the termination behavior is equivalent to the one of *D1* and *D2*. The production system is shown in Table 11.

TABLE 11 – PRODUCTION SYSTEM OF HEURISTIC *EQUILIBRIUM*

If strategy is chosen,	**END game.**
If time limit reached and alternative containing current max exists,	CHOOSE alternative containing current max.
If no more ELIMINATION after last two changes of focus,	CHOOSE a strategy from remaining strategy set *Naïve*ly.
If all strategies compared to each other,	CHANGE focus. CHANGE_count = CHANGE_count + 1.
If all payoffs in strategy compared,	ELIMINATE strategy not containing max. Next MOVE = MOVE to next new strategy, new cell.
If no new cells in strategy,	all payoffs in strategy compared.
If MOVE 3 + MOVE 2 not possible,	no new cells in strategy.
If COMPARE finished,	do next MOVE.
COMPARE_strategies(a,b):	

case a = b:	next MOVE = MOVE 3 + MOVE 2. COMPARE_strategies finished.
case else:	ELIMINATE payoff and strategies from STM. Next MOVE = MOVE to next new strategy, new cell.
If two max containing strategies in STM,	COMPARE_strategies(strat1, strat2).
COMPARE_payoff(a,b):	
case a>b:	max = a. COMPARE_payoff finished.
case else:	max = b. COMPARE_payoff finished.
If two payoffs in STM,	COMPARE_payoffs(payoff_1, payoff_2). STORE max containing strategy in STM. ELIMINATE payoffs from STM.
If one payoff in STM,	next MOVE = MOVE 1. Do next MOVE.
If cell read,	STORE payoff in STM.
If cell open,	READ payoff.
If payoff focus = opponent,	FOCUS on opponent payoff data. MOVE 1 = horizontal to new cell in next strategy, MOVE 2 = vertical to new cell in current strategy, MOVE 3 = −(MOVE 1).
If payoff focus = own,	FOCUS on own payoff data. MOVE 1 = vertical to new cell in next strategy, MOVE 2 = horizontal to new cell in current strategy, MOVE 3 = −(MOVE 1).
If cell closed,	OPEN cell.
If the game starts,	CHOOSE payoff focus. Current max = empty. MOVE to first (nearest) strategy, first cell. First strategy = current

	strategy. First cell = current cell. CHANGE_count = 0.

The most complex heuristic is entitled *Sophisticated*. It is more of a theoretical construct than a heuristic. Costa-Gomes et al. (2001, p. 1195) describe this heuristic as perfectly rational. It perfectly predicts the opponents' behavior and is thus able to generate the perfect answer regarding strategy choice. How this optimal assumption about the distribution of types within potential opponents is determined ex-ante, remains unclear. Hence, it needs to be developed ex-post. What could be argued is that if no heuristic can be excluded ex-ante, *Sophisticated* would need to determine the choice of all other players. Suppose those players apply (at least parts of) heuristics to elaborate their decision. In that case, the production system of *Sophisticated* necessarily has to include the production systems of all those heuristics applied. The size of the production system would thus be by far the largest, even without considering the determination of the heuristics distribution among the players. When applied under effective time limitation this heuristic seemingly needs to rely mainly on *Random* choice, since it takes so long to conduct. Following the arguments stated, there might be no proper production system with which a human decision-maker could execute this heuristic effectively while problem-solving, especially under time pressure. For those reasons, a production system cannot be presented here.

4.3 Minimum Set of EIPs

The specific EIPs in use in the production systems mentioned above form the set stated in Table 12. Both, constitution of the set and definition of the EIPs, are differing slightly in comparison to Johnson et al. (1985). This fact is not surprising since the decision tasks studied by Johnson et al. (1985) are different from the ones of Costa-Gomes et al. (2001) regarding fundamental aspects. The set of EIPs is task-specific, making it a distinguishable property of a set. Note that the set of EIPs is aimed to be as small as possible in the distinct types of EIPs it comprises, but still big enough to describe the mental process accurately. This limitation has practical reasons: tracing EIPs in experimental data of human behavior is easier as fewer aspects need to be considered. Also, given the assumption that a minimal set of EIPs is complete, it is easier to distinguish among EIPs when their definitions are disjunctive. Thus, the interpretation of behavior regarding EIPs is much more comfortable.

TABLE 12 – (RE)DEFINITION OF EIPS AND SET OF EIPS

EIP	**Definition**
READ	Read (focussed) cell content into Short Term Memory (STM).
COMPARE	Compare two strategies (COMPARE I) or two payoffs (COMPARE II) regarding max/min/equivalence.
OPEN	Uncover payoff information of a cell.
ADD	Add a payoff to a sum in STM.
END	Stop process.
ELIMINATE	Remove cell/ strategy from consideration.
MOVE	Go to (next) cell/strategy.
CHOOSE	Announce preferred alternative.
FOCUS	Focus view/mind on own/opponent/both payoff(s).
STORE	Save date/score/result in STM.

The heuristics differ in the size of their production systems. Since the nonstrategic heuristics only evaluate one player's payoff, one visit per cell is sufficient. The strategic heuristics, which additionally take into account the opponent's payoff, need to revisit the cells. Generally, that increases the total number of EIPs applied within the task processing. However, this fact alone is not a sufficient indicator for the expectable total number of EIPs or the variety of the set of EIPs.[23] Nonstrategic heuristics also require fewer types of different EIPs in their production systems than the strategic heuristics. In general, strategic heuristics use a greater number of different types of EIPs and more EIPs compared to nonstrategic ones. They are thus more effort-intensive per se.

However, this statement is only valid for the heuristics' production systems. A human player might use none of the heuristics mentioned above or only one or more than one heuristic when solving a decision-making problem.[24] That would result in an individual requirement of different types of EIPs as well as an individual total number of EIPs applied within the task.

[23] This number seems to be more a matter of the number of pure strategies the players face in a game.
[24] This issue is stated by Devetag et al. (2016, p. 198) after analyzing relevant studies in strategic decision-making.

Looking at the derived set of EIPs as presented in Table 12, one finds some EIPs causing specific motions or at least motoric effects: READ, OPEN, END, MOVE, and FOCUS. That makes them easier to trace. Others are seemingly limited to cognition: COMPARE, ADD, ELIMINATE, and STORE. In that case, the EIPs must be traced indirectly, for instance through analysis of the behavioral context. CHOOSE could be both. For example, choosing a strategy by clicking on a button is a motoric issue. However, choosing a payoff focus is limited to a mental process.

Nonetheless, FOCUS is a perfect indicator, providing the result of CHOOSE in the latter case. What is more interesting is the fact that some heuristics use certain EIPs exclusively. Only the strategic heuristics *D1*, *D2*, and *Equilibrium* apply COMPARE to compare two strategies. It is thus further referred to as strategic EIP. In Table 12, this EIP is designated as COMPARE I. From that the case is to be distinguished when comparing two payoffs (COMPARE II).

In the first part of this treatise, the author developed a concept that includes preparation time in a process model of strategic decision-making. This seemed necessary, since studies of strategic decision-making under time pressure, which omit the process perspective and rely on decision information only, have difficulties in interpreting some of the observed effects. The author could show the relevance of process components, named ‘EIPs', for understanding, modeling and studying such processes and finally identifying patterns of the cognitive process. The set of EIPs potentially in use in strategic decision-making was derived from a set of heuristics commonly adopted in such tasks. Their descriptions were transformed into production systems which offer direct access to EIPs used in the process. This representation can be transformed into computer code and hence easily applied to different problem tasks and various time pressure conditions.

The implications of the preparation time model need to be studied in more detail to elaborate the predictions of behavior. In this context the heuristics’ performance under time pressure conditions is of particular interest from a normative perspective. A simulation approach can acquire performance benchmarks and characteristic EIP-patterns which potentially show up in real decision-making, too. Those results will help formulating hypotheses of observable decision behavior with deeper understanding. The simulation of heuristics' behavior under time pressure is subject of the following Part II of the treatise.

Part II: Simulation of Strategic Decision-Making under Time Pressure

Part II of this treatise provides a description of the simulation approach that is developed to elaborate the predictions of the preparation time model. Two-person normal-form games with complete information in various forms serve as decision-making tasks that are solved under varying time limit conditions. The players' behavior is represented by a set of commonly used heuristics.

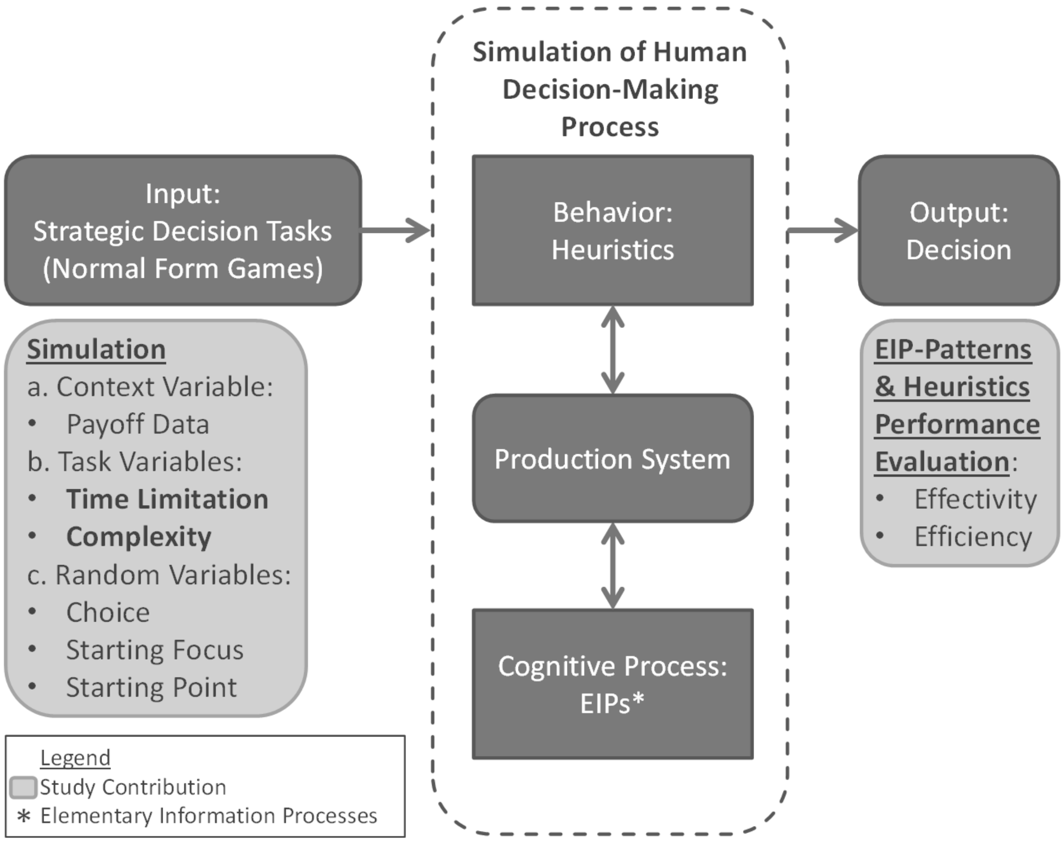
The approach is described in Chapter 5. Data analysis and results are presented in the subsequent Chapter 6. A discussion of successful behavioral patterns in Chapter 7 concludes Part II. The elaborated implications of the preparation time model are evaluated in the following Part III of this treatise.

5 Simulation Approach

In this chapter an approach is presented to simulate behavior in a variety of two-person normal-form games of complete information under time pressure conditions. The players' behavior in the decision-making tasks is represented by heuristics from a given set. This approach allows studying the performance of the heuristics under time pressure. Benchmarks can be defined from the results and characteristic EIP-patterns can be determined. The results help to elaborate the predictions from the preparation time model. Finally, hypotheses of cognitive processes under time pressure can be formulated with deeper understanding of decision-making. Figure 4 illustrates the approach with study contributions depicted in light gray boxes. The approach follows suggestions of Kelton and Barton (2003).

In the simulation, heuristics' production systems represent the players' information acquisition and processing schemes. Heuristics are matched in two-person-games in changing combinations to study their interactive potential concerning payoff generation. Time pressure is modeled as a function of task complexity and time limitation (see p. 15, footnote 10). Effectiveness and efficiency determine performance. A concept to measure performance is derived from the preparation time model and its implications. The following sections present this concept along with the simulation procedure and the technical implementation of simulation and data analysis with *Matlab* (in version R2014b). Chapter 6 deals with data analysis and results, while in Chapter 7 successful behavioral patterns are discussed. The results contribute to partly answering the second questions of this treatise formulated in Section 1.2.

FIGURE 4 – SIMULATION CONCEPT



5.1 Measuring Performance

In the context of this research, effectiveness and efficiency determine the performance of heuristics. The Oxford Dictionary describes effectiveness as "the degree to which something is successful in producing a desired result" (Oxford Dictionaries 2018a). Effectiveness refers to the fact of achieving a goal ("result"), with the goal often being linked to some optimality criteria ("desired"). Bettman (1979) discusses the role of goals in the context of choice motivation in more detail. He defines goals as desired end states which are to be achieved from initial states, characterized by the amount of information and resources available (Bettman 1979, p. 19). This definition fits perfectly in a process approach and is thus followed here.

The Oxford Dictionary further defines efficiency as "the state or quality of being efficient" and efficient as "achieving maximum productivity with minimum wasted effort or expense" (Oxford

Dictionaries 2018b). The term efficiency addresses the relation between input resources and realized output of a goal, referring to quantitative aspects. Its applicability for evaluating processes thus is evident.

Before these two aspects can be further characterized concerning its use in the evaluation of the heuristics' performance, it is crucial to give an idea of the term 'optimality' in this context. Zakay and Wooler (1984) intensively discuss the multiple meaning of optimality, especially in the fields of psychological studies and in comparison to economic studies (pp. 273 f.). As a result of this discussion, they point out the importance to concretize the term regarding the objectives of the analysis.

The purpose of this simulation approach is to compare the performance of a set of heuristics under time-limited conditions. Thus, the understanding of the term 'performance' need be defined in this context. This understanding is given with the following four goals: the primary goal of the heuristics discussed here is to provide a recommendation for a decision. This adds to the aspect of effectiveness. It is accomplished if the heuristic provides a recommendation and not accomplished otherwise. For constructional reasons, every heuristic's production system offers an ultimate choice recommendation, even when developing a decision under severe time pressure. The first goal of the performance measurement approach has thus no discriminatory power in the examined set of heuristics and needs to be supplemented by other criteria. In the evolving goal hierarchy, the primary goal represents a necessary condition, triggering the consideration of the subsequently implemented goals. If the primary goal remains unachieved, the following goals would not be either, making further considerations pointless. However, as single goal it is not sufficient to define optimality of a heuristic's performance.

Johnson et al. (1985) follow a comparable path in evaluating the performance of their heuristics in use. They rely on two aspects: the effort of and the accuracy in choice. Whereas the number of involved EIPs estimates effort, the accuracy is linked to consistency in choice and in the degree of following the (normative) rule of "Maximization of Expected Values (EV)" (Johnson and Payne 1985, pp. 396 f.). They expect the heuristics to differ markedly in the amount of information processed until a solution (i.e., decision) can be provided.

Measuring the effort and evaluating its amount is both reasonable and technically easy to implement in this simulation. Accuracy in the understanding of Johnson et al. (1985) aims at evaluating the choice regarding rational aspects, relating it to the most rational decision rule they found. In the context of the present study, this would mean to relate choice to the *Equilibrium* decision. However, decisions of the *Equilibrium* heuristic might not be the most sophisticated under time restricted conditions. Instead, linking it to the *Equilibrium* decision made without time pressure would go beyond the context, comparing choices under totally different conditions. That would hence set the wrong framework for interpretation. It is thus necessary to develop an adequate and applicable evaluation system.

The three supplementing goals presented in the following are related to the primary one of generating a choice recommendation and formulate principles to evaluate and discriminate the heuristics' performance.

The first of those three is the reduction of the set of alternatives, referring to the aspect of effectiveness. Apparently, a reduction to one remaining alternative is equivalent to a decision, and hence fulfilling the primary goal. The reduction procedure of the examined set of heuristics principally follows reasonable criteria in the sense of rational choice. An arbitrary reduction, having the same effect as a random choice, would serve that goal adequately. Hence, this goal needs to be complemented by a quality condition of choice to interpret the performance comprehensively. The second supplementing goal implements such a quality criterion when demanding the maximization of a certain payoff[25] in a game. It refers to the aspect of effectiveness. The last goal proposed in this context is the reduction of effort, or formulating equivalently, the minimization of decision time. That issue refers to the aspect of efficiency, describing measurable means of the level of goal achievement.

Those four goals together determine the term optimality in use here and hence form the basis of the evaluation concept of this study. Table 13 summarizes this evaluation concept. The degree of fulfilling these goals describes the optimality of a heuristic. The four goals can be further categorized regarding their contribution to the aspects of effectiveness and efficiency to characterize the performance of the heuristics. So, when examining a heuristic's effectiveness, the degree to which it fulfills the Goals 1 to 3 is reflected. Efficiency is determined by the results achieved at Goal 4. It is further possible to compare the heuristics concerning optimality, classified in goals (1 – 4) and category (effectiveness/efficiency).

[25] Usually it is the payoff of one player. The exceptions are the heuristics *Altruism*, which aims to maximize payoff of both players, and *Pessimism*, which aims to maximize a player's minimum payoff.

TABLE 13 – EVALUATION CONCEPT

Goal		Effectiveness	Efficiency	Measured variable [dimension]
1st	Propose a decision.	X		Fulfillment [Boolean][26] (const.)
2nd	Reduce set of alternatives.	X		Sum of remaining alternatives [Integer]
3rd	Maximize certain payoff.	X		Payoff [Rational]
4th	Minimize effort.		X	Sum of EIPs [Integer]

Generally, the term effectiveness describes the degree to which a result achieves a predefined goal. The underlying methodology henceforth must interpret this degree. Zakay and Wooler (1984, p. 274) define effectiveness as the "decision's optimality in the narrow meaning of the logical handling of existing inputs and values", obviously referring to a process view of decision-making. In the same spirit, several authors use this approach when comparing choice behavior with a (normatively) optimal heuristic.[27] However, the conditions of optimality in case of selecting a benchmark heuristic do not necessarily match the ones of the study's evaluation concept aspect. There might be others, being at least as good as the benchmark rule (see Zakay and Wooler (1984, p. 274) for example).[28] As much as it seems evident to define one heuristic as a benchmark, this approach is not followed in the following examination.

For the goals of this research, effectiveness is related to a heuristic's ability to propose a decision (see Table 13, 1st goal). Since the random choice of remaining alternatives as 'last action' is inherent in any heuristic per definition, coming to a solution is no discriminatory aspect. It is hence

[26] Whether or not the primary objective is achieved can be principally measured by employing a Boolean variable. However, since the heuristics' production systems force a decision under any time limits by their design, this goal is achieved for every simulated time limit condition.
[27] For the same purpose, other authors introduce the concept of "accuracy in choice" (Johnson and Payne 1985; Payne et al. 1988) or "quality of choice" (Sutter et al. 2003; Kocher and Sutter 2006) in their studies. Hence, the meaning of the terms effectiveness, accuracy, and quality regarding choice behavior should mostly correspond, with variations in the concrete implementation.
[28] They employed the (normative) multi-attribute utility rule as a benchmark for optimal choice behavior and name its easy-to-handle ability as the predominantly reason for their choice.

extended by the ability to reduce the set of alternatives (see Table 13, 2nd goal). The reduction can be represented by the number of remaining alternatives in the moment of choice – the smaller the number, the more effective is the heuristic. Reducing the set of alternatives is no guaranty of a reasonable choice regarding the goal of maximizing payoff (see Table 13, 3rd goal). For this, an evaluation concept has to consider a criterion assessing the chosen alternative and the resulting payoff needs. Three potential ways to implement such a criterion presents the following consideration. All of them support the interpretation of the data.

Following the commonly proposed concept of effectiveness, benchmark heuristics are frequently used to explicitly point out the performance of a heuristic in relation to the benchmarks. The identification of those benchmark heuristics, in turn, depends on the anticipated perspective. Usually, the assumingly best, worst or a combination of both regarding one or more criteria is applied to set the results of the heuristics in context (e.g. Johnson and Payne 1985; Payne et al. 1988; Bettman et al. 1990; Sutter et al. 2003; Kocher and Sutter 2006; Lindner and Sutter 2013; Lindner 2014). In case of this study and similar to Costa-Gomes et al. (2001, p. 1195), the heuristic *Equilibrium* represents the most sophisticated one of the set of examined heuristics regarding rationality.[29] Of course, it can only provide the equilibrium strategy, if at least one exists and if a subject can fully process the heuristic. The opposite represents the heuristic *Random*, with supposedly no rationality in choice according to maximize payoffs (3rd goal). The overall results of all other heuristics could then be set in context to those two benchmark heuristics for a particular game.

Alternatively, the payoff results could be transferred into a ranking. Since all heuristics face the same payoff set, this transformation preserves the context. The absolute distances are neglected. It is further independent of benchmarks and also offers the possibility of a comparison of the heuristics' performance overall goals that is easy to execute.

---

[29] Note that Costa-Gomes et al. (2001, p. 1205) investigated another heuristic named *Sophisticated*. That theoretical construct surpasses the decision quality of *Equilibrium* since it correctly estimates the distribution of heuristics applied within a given population of subjects. Nevertheless, two of their central assumptions conflict with the approach taken here, making it impossible for *Sophisticated* to be considered: 1. Costa-Gomes et al. (2001) assume that the distribution of heuristics is fixed for their set of tasks, ruling out adaptation. They need this for stochastic reasons when estimating the distribution of heuristics within the population of subjects. 2. They explicitly examine the initial behavior of subjects with no learning and no signaling. This intention implements that the distribution of heuristics applied within a population cannot be determined (by either experimenter or subject) prior to a decision. Since this would be necessary for applying the *Sophisticated* heuristic, this type is limited to theoretical considerations, where retro-perspective information can be implemented in choice. In the current context, this is not possible.

The third method presented here transforms the resulting payoff of the heuristics in accordance with the maximum and minimum payoff in a payoff set. Herewith, especially the impact of the payoff values, as important context variable, could be reduced.

All three methods seem adequate, each highlighting different aspects, and are hence considered for presentation and discussion in this paper. Of course, the absolute values of the heuristics' performance remain crucial for further analysis and thus will be presented, too.

The possible decisions of both players' perspective need to be taken into account to determine a heuristic's choice payoff in normal-form games. The evaluation of heuristics in normal-form games regarding the quality of choice, measured in payoff numbers, is to the knowledge of the author not examined in any study so far. Some nonstrategic decision-making studies apply the expected value of a particular heuristic in a related approach (see Johnson and Payne (1985) and Payne et al. (1988) for instance). Transferring that procedure to the current strategic concept implicates basic assumptions of opponent's choice behavior to the effect that generality of the findings could principally be limited. Alternatively, the performance of a heuristic regarding the 3rd goal can be estimated by consequently matching each heuristic with any other from the set. This approach supports examining the impact of the heuristics' interactivity properly.

Efficiency is often addressed in context with effectiveness, evaluating means and quantity of means used to achieve a goal. Models in decision-making theory that consider the consumption of resources often examine the efficiency of decisions (e.g., the cost-benefit approach). The information processing approach is to number among those models since this approach provides immediate access for evaluating resource consumption of decision-making (as a process characteristic). This consumption is usually linked to the effort of a heuristic. Thus, efficiency can be related directly to effort. As elaborated earlier in this text, the concept of EIPs is undoubtedly useful to measure effort, even if determining the effort of a single EIP by measuring its processing time still raises methodological concerns.

The time span between starting a task and accomplishing it is comfortable to determine if the measurement design is appropriate. Whoever wants to determine the duration of types of EIPs faces the problem of generalizing the findings. That is because processing times can be widely linked to individual abilities and the concrete task (Dansereau and Gregg (1966, p. 71), Kahneman (1973, 24 f.), Russo (1978, p. 94), Wickens (1981, p. 37), Bull and Johnston (1997, p. 17), VanRullen and Thorpe (2001b, p. 458)). Further, there remain several unresolved problems with the identification of basic cognitive tasks that occur while decision-making.[30] This identification also complicates measuring the duration of the basic cognitive tasks. Together with existing criticism concerning the method of determining the duration of EIPs the use of time as an appropriate means to estimate effort needs to be discussed in more detail.

One interesting result in this context, presented by Johnson and Payne (1985), eases the measurement of effort markedly. In their study, the authors simulate two models of estimating effort through time duration of EIPs. The first model applies empirically derived durations for every type of EIP as reported by Dansereau and Gregg (1966, p. 71), Chase (1978, p. 85), and Russo (1978, pp. 94 ff.). The effort is calculated by summing up all EIPs, with each summand weighted by its (supposedly) type-specific duration. The second model weighs each EIP equally, even though assuming that all EIPs consume the same amount of time might be a substantial simplification. Nonetheless, when comparing the models regarding their estimated effort, the results show no significant difference (Johnson and Payne 1985, p. 406). In consequence, a single EIP can be used as a measuring unit independent of its concrete type.

However, a notable disadvantage of this approach in studying the influence of time pressure is the scale of results, which differs between the number of EIPs and time. Nonetheless, modeling time pressure by the number of EIPs and hence replacing time limitation with the limitation of the number of EIPs is a natural extension of the idea of Simon and Newell (1971). It is already successively implemented in studies of Johnson and Payne (1985), Payne et al. (1988) and Costa-Gomes et al. (2001). The authors' proposal is partly based on findings of Just and Carpenter (1976, p. 477) and Card et al. (1980, pp. 61 ff.) who could show a significant relationship between the number of EIPs and response times for several cognitive tasks. Following the concept of information processing and interpreting EIPs as processes, each EIP consumes resources (including time), by definition. Therefore, the relationship between the number of EIPs and response time is stringent.

The resulting numbers of EIPs can further be used to evaluate the efficiency of heuristics – isolated or in comparison. It should be noted that the comparison of effort between two heuristics

[30] Process tracing methods, even though evolving constantly, still rely heavily on the interpretation of observable behavioral data to draw conclusions with respect to the underlying cognitive process. Hence, EIPs are usually traced indirectly.

only makes sense if both heuristics meet the required condition (achieve 1st goal, Table 13,) and are applied to the same decision task. Following the presented evaluation concept, the performance of a heuristic is more efficient when less EIPs are applied to reach the same goal (for the corresponding goal see Table 13, 4th goal).

In order to set the resulting effort in context, Johnson et al. (1985) again prefer a benchmark. They use *Random* as a lower bound for effort since this heuristic needs by far the fewest number of EIPs to make a decision, independent of any task variable (Johnson and Payne 1985, p. 397). This bound holds in case of strategic decision tasks, too. In fact, it is stable under all task varying conditions.[31] Employing *Random* as a benchmark would hence have the same effect as using an absolute scale with the *Random* value as an additive constant. Absolute numbers (which are naturally measured on an absolute scale) and rankings are therefore regarded as sufficient here to measure efficiency.

Three variables are identified that link simulation data to the two aspects effectiveness and efficiency. The first one is the sum of EIPs used by a heuristic until a choice is made from the set of alternatives. This aspect is regularly used to measure effort and to compare heuristics' efficiency (see, for instance, Johnson and Payne (1985, p. 398) and Payne et al. (1988, p. 542)). The second aspect considers the choice and resulting payoff when exposed to alternatives identified by other heuristics. Hereby, a more detailed evaluation of the heuristics' performance regarding quality can be achieved as several other studies propose (Johnson and Payne (1985, 396 f.), Payne et al. (1988, p. 551), Costa-Gomes et al. (2001, p. 1194)).

Finally, the number of remaining alternatives before the ultimate choice occurs is presented as the third variable. It is directly linked to the effectiveness of the heuristics, with *Random* choice serving as an upper bound for comparison. Table 13 summarizes the evaluation concept, including the categorization of the goals identified and corresponding measured variables.

5.2 Heuristics

Costa-Gomes et al. (2001, p. 1195) propose the set of heuristics which serves as simulation basis here. They cite the commonness of use in the strategic decision-making environment of normal-form games as the reason for selecting this set. Since the heuristics face broad distribution within

[31] See the time limit section below in this chapter for a detailed description of the proposed performance of *Random* under time limitation.

relevant studies, their involvement offers a reasonable possibility of linking the results obtained here to those of numerous other studies.

Costa-Gomes et al. (2001) categorize their set in strategic heuristics, implying the consideration of the opponent's payoff information, and nonstrategic heuristics. The latter include *Random*, *Altruism*, *Optimism*, *Pessimism*, and *Naïve* whereas the strategic heuristics are *D1*, *D2*, *L2*, *Equilibrium*, and *Sophisticated*. The set of heuristics is identical to the one applied for deriving the minimum set of EIPs in Chapter 3, which also describes heuristics. Section 4.2 provides a detailed description of the Corresponding production systems for this set of heuristics. This form is transferable into flow charts which support implementation in various programming languages without further changes.

In comparison to other solution concepts of classic game theory, the choice of mixed equilibria is ruled out. None of the heuristics offer this possibility in their production system's implementation. The reason is its missing representation in human choice behavior in initial decision-making (i.e., one round of play).

Whenever the execution of a heuristic is close to being interrupted by exceeding the time limit, the ultimate action is per definition a random choice between remaining alternatives.[32] Hence, any heuristic presented here finishes with selecting one of the pure strategies.

5.3 Task Variables in Simulation

Task variables and context variables specify decision tasks. Payne et al. (1988, p. 536) define task variables as "[…] general characteristics of a decision problem […]" – their impact on decision-making is independent of the "[…] particular values of the alternatives […]". Context variables are particular values of alternatives. In normal-form games, task variables are game type, number of alternatives and time limitation, whereas the payoff values count as context variables. It seems relevant to emphasize the relationship between game type and payoff value to clarify the categorization of both variables. The game type provides a specific payoff structure with a fixed status for the alternatives. This status, for example, describes whether an alternative is dominated or not by another alternative. The structure thus determines the performance of solution concepts, such as the iterated elimination of dominated strategies, when applied to the problem task. Payoff val-

[32] Note that the trivial case of making no decision or failing to make a decision is not of interest here. No decision implicates that the game is not played. A punishment as consequence of not choosing is not part of the simulation.

ues vary within certain bounds without changing the structure of the game. The payoff value can be regarded as a context variable as long as staying within these bounds. Exceeding these bounds would change the structure of the game, and the classification would change to task variable. Nonetheless, same as task variables, context variables can principally influence the selection of problem-solving strategies and thus the decision-making process (Payne 1982, p. 538). In the following, the implication of the three task variables game type, complexity and time limitation is described in more detail. The determination of payoff values especially gains importance in the context of experimentation and is thus discussed in a later section of this treatise (Section 9.3).

5.3.1 Game Types

The investigated set of games consists of six traditional game concepts (Hawk-Dove, Chicken, Prisoners' Dilemma, Battle of Sexes, Throwing Fingers, Stackelberg's Leadership) and four games proposed by Costa-Gomes et al. (2001, p. 1203). These normal-form games are designated as Costa 2A, Costa 2B, Costa 3A and Costa 3B. Appendix B shows the payoff matrices of all games. The selection contains a variety of games which are established concepts in game theory, well described in the literature and thus offer a link to other studies. Following this argumentation, it is henceforth appropriate to examine games especially proposed by Costa-Gomes et al. (2001) since their approach is related to the one used here. Beside the comparability to other studies, this concept could mainly be extended to any other game in normal-form, independent of the number of alternatives without further assumptions.

5.3.2 Task Complexity

The complexity of a decision task can be related to the number of alternatives as Johnson and Payne (1985, p. 401) for a nonstrategic and Costa-Gomes et al. (2001, pp. 1202 ff.) for a strategic decision-making environment report. Costa-Gomes et al. (2001, p. 1203) limit the complexity of their games, with a maximum game size of 2 by 4 and 4 by 2 alternatives respectively. Their selection of games focussed on distinguishability between sophisticated and non-sophisticated information search rather than on varying complexity (Costa-Gomes et al. 2001, p. 1202). However, varying the task complexity seems crucial to examine as many aspects of problem solving as possible (Woods 1993, p. 229).

Usually, a payoff matrix shows the payoff of both players at a time. Each cell of the matrix stores the own payoff and the one of the opponent. Assuming that the payoff of one player in one cell is equivalent to one unit of information, a payoff matrix of 2 by 4 would contain 16 units of information. Costa-Gomes et al. (2001, p. 1201) preferred a not very common alternative presentation form, where the payoff matrix is split into two. Each matrix contains only one player's payoff. In their experiment subjects face two matrices with a maximum of eight cells each and one unit of information per cell, giving a total of 16 units of information and hence 16 stimuli at maximum per task. Why is the number of stimuli relevant here? Considering Miller's law on human information processing capacity that predicts an upper limit of 7 ± 2 stimuli, both ways of presentation seems to violate the law (Miller 1956). In fact, that law would limit game complexities to 2 by 2 payoff matrices, if stress from mental overstrainment is to be ruled out.

It could be argued in favor of larger payoff matrices that a player, who applies a strategic heuristic, would not regard the whole payoff set. He preferably compares maximal two alternatives of one player's payoff set at a time. With a single cell opened at once, the player needs to store parallel 2 by 4 stimuli at most. On first sight, this seems to satisfy Miller's Law. De facto it would not: after comparing the payoffs of two cells, a piece of meta-information is generated (the orientation of the comparison's result). This piece would add to the number of stimuli and reach the capacity limit Miller proposed. The following comparison of the next two cells' payoff would already exceed the capacity limit.

However, arguing with production systems of heuristics (Section 4.2), Miller's law is not violated in case of larger payoff matrices. Heuristics only store a single bit of payoff information at once and compute it in the next step. This way they produce a piece of meta-information and actualize it in one of the next steps. Information, which is no longer in use, is eliminated from short-term memory. Repeating the lookup restores the information. Studies of information acquisition in decision-making tasks frequently report such behavior (among others Costa-Gomes et al. (2001, p. 1196, footnote 8)). Thus, it is possible to satisfy Miller's law, even in case of payoff matrices larger than 2 by 2 alternatives. Only such information units count that are visible at once or in focus at a time during the cognitive process.

Costa-Gomes et al. (2001) cover payoff information per se in the experimental design. The content of a single cell each can be visualized one at a time. Thus, one external stimulus is to be considered at a time. In that case, up to 6 ± 2-stimuli are left for internal representations of the task and computing a solution. Those stimuli are supposed to be meta-information in the sense that they contain intermediate steps of the solutions. They are derived from earlier calculations and stored in short-term memory for immediate or later use. Examples of meta-information are current maximum alternatives, current minimum payoffs, current sums of payoffs, and results of comparisons.

Following this argumentation, the total number of payoff information units in a game does not account for the total number of stimuli a subject faces at a time. Thus, Miller's 7 ± 2-rule can also be satisfied at game sizes larger 2 by 2, enabling investigations on more complex games.

However, a limitation in size seems practical for other reasons.[33] Task presentations, as well as having a potential deterring impact on decision makers, are reasonable limitations for the design of complexity (Ho and Weigelt 1996, p. 676). Both reasons can be evaluated and specified experimentally. Corresponding experiments in nonstrategic decision-making with an information presentation similar to a payoff matrix see an abrupt reduction of information acquisition at game sizes of six alternatives and larger (Payne 1976, p. 374). Five alternatives are thus a reasonable upper bound for the task complexity. Game sizes of 2 by 2 to 5 by 5 are regarded as a compromise between peoples' acceptance of complexity and this study's objectives. They are applied in this simulation.

The game size is used to construct normal-form game tasks of different complexity. Starting with the set of basic games introduced at the beginning of this section, all of size 2 by 2, the extension design partly follows a particular construction method. This method is applied to develop complex versions of all classic game types. Together with the types proposed by Costa-Gomes et al. (2001, p. 1203), they form a representative set. It provides a variety of parameters sufficient for this research (see Table 14). The method is developed in the spirit of Herbert Gintis' compilation of classic game-theoretic problems and his 'Col. Blotto'-scenarios[34] (Gintis 2009, p. 141). He extends the basic version of a game type with certain alternatives, without modifying the basic game itself, intending to demonstrate the variability of a game's solution (in game-theoretic terms). One part of his methodological scheme implicates adding neutral alternatives which do not alter the game-theoretic solution of the primary game. This procedure assures the existence of

[33] Note that these aspects principally do not affect simulation that could technically implement large numbers of alternatives per game and player. However, this data would lack experimentally derived equivalents to compare with and hence would be of limited practical use.
[34] The 'Col. Blotto'-game is a classic scenario in game theory. Two players independently distribute their limited forces over a specific number of battlefields. The higher number of forces dominates a field, and the higher number of dominating battlefields wins the war. It is a constant sum game. Its structure is also known as the 'Divide a Dollar' allocation class problem, originating from the French mathematician Emile Borel who proposed and solved it in 1921. For a more sophisticated consideration see for instance Roberson (2006).

dominated alternatives. A similar approach can be found at Brams (2000, p. 4). According to the Cognitive Hierarchy Model, proposed by Camerer et al. (2004, pp. 864 ff.), elimination of dominated alternatives is an indicator of the depth of reasoning.

Nonetheless, it also offers alternative solutions for other heuristics. This extends the solution space of the set of heuristics, which contributes to the differentiated evaluation of the heuristics and their performance. In the following, a method is presented to develop complex versions of the set of classic game types. The present study applies this method.

The development of complex versions need to consider two conditions to secure comparability of choice: the extended game shall possess the same key properties of the basic game regarding symmetry and constant sum. That means if the basic game is symmetric, the extended one is too. If the basic game is of the class 'constant sum', the extended one is too. The negations hold respectively.

Taking these conditions into account, the following design patterns are used to create complex versions of three to five alternatives per player from the basic game tasks:

1. Add a third alternative for each player that is a strictly Pareto efficient solution.[35]
2. Add a fourth alternative, which is (weakly) dominated by at least one of the other three or by a combination of them.
3. Add a fifth alternative that is exclusively dominated by the fourth.

However, a small number of game tasks developed do not fully meet these conditions because their payoff structures limit the application of the design patterns.

Applying those design patterns, a variety of games is generated that includes diverse types and sufficient reduction potentials (concerning dominated alternatives) for this study. Table 14 lists some key characteristics of the games regarding their game-theoretic solution, especially in the extensions.

[35] Pareto efficiency of an alternative a implies that no other alternative a' Pareto-dominates a. In other words, no alternative exists that would make one player better off without making the opponent player worse off concerning payoffs (utility). For more details compare for example Leyton-Brown and Shoham (2008, p. 10).

TABLE 14 – GAME TYPE CHARACTERISTIC: REDUCTION BY DOMINANCE

Size **Game**	**2×2**	**3×3**	**4×4**	**5×5**
Battle of Sexes	$(2\times2)^{p,m)[1]}$	$(2\times2)^{p)[1]}$	$(3\times3)^{p)[1]}$; $(2\times3)^{m)[2]}$	$(4\times4)^{p)[1]}$
Chicken game$^{*)}$	$(2\times2)^{p,m)[1]}$	$(3\times3)^{p,m)[1]}$	$(4\times4)^{p)[1]}$; $(4\times3),(3\times4)^{m)[1]}$	$(4\times4)^{p)[1]}$; $(4\times3),(3\times4)^{m)[2]}$
Costa 2A	$(1\times1)^{p)[1]}$			
Costa 2B	$(1\times1)^{p)[1]}$	------------------no extension------------------		
Costa 3A	$(1\times1)^{p)[1]}$			
Costa 3B	$(1\times1)^{p)[1]}$			
Hawk-Dove game$^{*)}$	$(2\times2)^{p,m)[1]}$	$(3\times3)^{p)[1]}$; $(1\times1)^{m)[2]}$	$(3\times3)^{p)[1]}$; $(1\times1)^{m)[3]}$	$(3\times3)^{p)[2]}$; $(1\times1)^{m)[4]}$
Prisoner's Dilemma$^{*)}$	$(1\times1)^{p)[1]}$	$(1\times1)^{p)[2]}$	$(4\times4)^{p,m)[1]}$;	$(5\times5)^{p,m)[1]}$;
Stackelberg's Leadership	$(2\times2)^{p,m)[1]}$	$(3\times3)^{p,m)[1]}$	$(4\times4)^{p)[1]}$; $(3\times3)^{m)[1]}$	$(4\times4)^{p)[1]}$; $(3\times3)^{m)[2]}$
Throwing Fingers	$(2\times2)^{p,m)[1]}$	$(3\times3)^{p,m)[1]}$	$(3\times3)^{p)[1]}$;	$(3\times4)^{p)[1]}$

Legend[36]

$(\boldsymbol{m} \times \boldsymbol{n})$ reduction to $\boldsymbol{m}$ row player and $\boldsymbol{n}$ column player alternatives;

Superscripted:

***)** game type is symmetric[37];

Elimination of dominated strategies by **p)** pure / by **m)** mixed dominating strategies;

Level-**[k]**-reasoning;

The resulting payoff matrices for the set of games are shown in Appendix B, having the basic game in the first two rows and columns (equivalent to a game complexity of 2 by 2) and the extensions in rows and columns three to five (equivalent to a game complexity of 3 by 3 to 5 by 5). A game of complexity 4 by 4 thus always contains the basic game, expanded by a third and a fourth alternative for each player, according to the patterns stated above. It should be noted that the original interpretation of a game type as a specific strategic conflict situation between two players may not apply to their extended versions.

[36] Example $(1\text{x}1)^{p)[1]}$: Game can be reduced to a single Nash-Equilibrium by pure dominating alternatives in a single step of reasoning (i.e. one round elimination of dominated alternatives).

[37] In symmetric games the role of the players is interchangeable without changing the payoff structure of the game.

5.3.3 Time Limitation

The time limitation is measured in number of EIPs, as proposed in Chapter 3. In the simulation, the range varies between 3 – 1400 EIPs, in steps of three EIPs. The lower bound is related to the absolute minimum number of EIPs for all heuristics in games of lowest complexity.[38] Apparently, zero EIPs is the trivial case with no decision at all, independent of complexity and game type. Regardless of the trivial case, the minimum is expected to occur in a *Random* choice. Here, a decision is available after one EIP, and the production system needs two EIPs to complete the task. This number is constant for all variations of task complexities and games. A random choice is per definition the replacement decision for all heuristics in case of effective time limitation. Thus, a limit of one respectively two EIPs leads to the same decision for all heuristics and three EIPs seem adequate as a lower bound.

The upper bound is derived from analyzing the maximum number of EIPs comparing the performance of all heuristics in all games of highest complexity (*D2* at 5 by 5 Stackelberg's Leadership – 913 EIPs). An increase of a single EIP does not affect the performance of the heuristics and is hence equivalent to the case of unlimited decision time.

Since a single EIP forms an atomic unit, limitation increases discretely. A step width of three EIPs proved to be sufficient to discriminate the heuristics' performance adequately and delivers a good time performance of the simulation at the same time.[39]

5.4 Random Variables in Simulation

Each heuristic has at least one of three random elements in its production system that affect performance. As mentioned earlier, under effective time limitation the heuristics provide a random decision between the (remaining) alternatives as ultimate action. This behavior is defined by the design of their production systems. It occurs until a single alternative is determined for choice and time limit is still effective. The point in time the (pre-)selection of an alternative occurs varies with the heuristics and the problem task. When the time limit is no longer effective, all heuristics, except for random, no longer rely on a random selection of an alternative.

[38] The relation between complexity, i.e., the number of alternatives per player, and expected number of EIPs is positive for any heuristic except *Random* (where it is constant).

[39] This step width, in turn, is a specific example of trading accuracy for speed in the proceeding.

The second random variable is the point within the payoff matrix at which the acquirement of information starts. Even though the heuristics' production systems scan the matrix systematically, the starting point is not determined. It could be any of the payoff matrixes' cells. All cells can be chosen as starting point with same probability. The value of this probability is thus depending on the task complexity.

Finally, the strategic heuristics, except for *L2*, exclusively contain a third random variable in their production systems' design: the starting focus, which determines whether the own payoff or the opponent's payoff is scanned first. The two focus options are chosen with same probability.

The random variables affect the heuristics' production systems quite differently and also change with time limitation. This fact influences the simulation setup regarding the number of runs per parameter constellation. An overview of the time limits for which the random variables need to be evaluated within the simulation range of [3 – 1400] EIPs is given in Table 15 for the set of heuristics.

TABLE 15 – EVALUATION OF RANDOM VARIABLES BY TIME LIMITS

Heuristic	**Starting focus**	**Starting point**	**Random choice**
Random	-	-	[3 – 1400] EIPs
Altruism	-	[5 – 1400] EIPs	[3 – 4] EIPs
Optimism	-	[4 – 1400] EIPs	[3] EIPs
Pessimism	-	[4 – 1400] EIPs	[3 – 4] EIPs
Naïve	-	[5 – 1400] EIPs	[3 – 4] EIPs
L2	-	[3 – 1400] EIPs	[3 – 1400] EIPs
D1	[3 – 1400] EIPs	[3 – 1400] EIPs	[3 – 1400] EIPs
D2	[3 – 1400] EIPs	[3 – 1400] EIPs	[3 – 1400] EIPs
Equilibrium	[3 – 1400] EIPs	[3 – 1400] EIPs	[3 – 1400] EIPs

The nonstrategic heuristics, with the exception of *Random*, are based on random decisions only for very small time limits, regardless of game type and complexity level. After the heuristics have determined a first payoff cell, the corresponding alternative is selected until a better one is identified. At this point, only the starting point for which an alternative is selected is important until the entire set of alternatives has been processed. For the strategic heuristics, the influence of the three random variables cannot be predicted as easily as for the nonstrategic ones. The starting focus is selected among the first actions of the production systems. The influence of the other two variables varies depending on the game type and complexity, until the production systems end their processes. Since the end point is difficult to calculate and thus not known *ex ante*, the influence of the random variables is evaluated for the entire simulation range of the time limit.

5.5 Simulation Procedure

The simulation approach is depicted in Figure 5. It comprises the parts 'simulation', 'calculation', and 'analysis'. The parts are processed one after the other within the approach (bold rounded rectangles in the middle of Figure 5).

The procedure for each part is depicted in more detail as a flow chart (large rounded rectangles above and below the main process). In the simulation part, every heuristic is applied to every decision task. The overlapping effects of the random variables on Goal 2 to Goal 4 of the heuristics' performance need to be taken into account. All three variables' randomness is based on discrete uniform distributions with respective numbers of possible alternatives. Their statistical inference makes an analytically computation of their effects difficult.

A stochastic simulation helps to evaluate the influence of those probabilistic elements numerically, reducing the complexity of the studied system (Domschke and Drexl 2005, p. 223; Robert and Casella 2010, p. 62). A parameter constellation is thus run for as many times as necessary to determine a convergence regarding the influence of the random variables.[40] For every parameter constellation the mean of the results (obtained from the simulation data of the final run when convergence is determined) are taken for analysis.

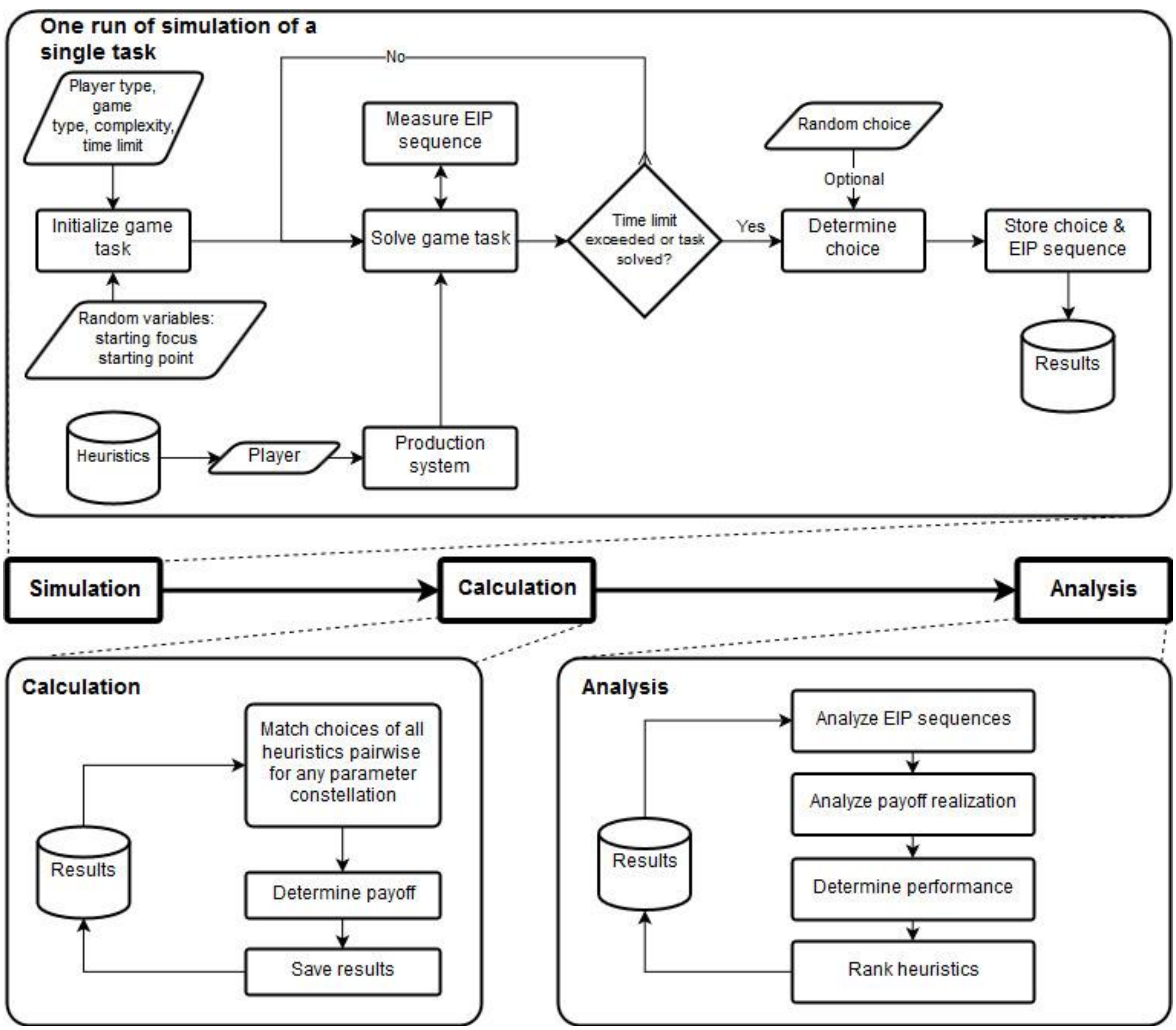
For a single run of one such decision task, the simulation procedure is depicted in Figure 5 (large rounded rectangles above the main process of the simulation approach). The heuristics' production systems represent the players' problem-solving processes. The games that serve as decision tasks in the simulation are implemented in their normal-form as matrices. Each game represents a single parameter constellation of player type, game type, complexity, and time limit. The parame-

[40] The performances for each goal and for every run of the simulation can be interpreted as results of an experiment. According to the Law of large numbers, the performances' means converge towards the true means with increasing numbers of runs. Cauchy's convergence test is used on the means per goal and per run. To determine convergence, two means of consecutive simulation runs $\bar{X}_m, \bar{X}_n$ need to satisfy the requirement $|\bar{X}_m - \bar{X}_n| < \varepsilon$, with $\varepsilon = 0.01$. The minimum number of runs is 7. This value ensures that for two randomly chosen alternatives (i.e. starting focus), each with a probability 0.5, the probability to choose the same alternative 7 times in a row ($= X$, $X \sim B\left(7; \frac{1}{2}; 7\right)$ is $(\frac{1}{2})^7 < 0.01$.

ters of the simulation are presented in Table 16. While task processing, the EIP sequence of the production system is recorded. After the time limit is reached or the heuristic production system completes the task, the EIP sequence and the resulting alternative choice are stored along with the game parameters.

In the calculation part, any heuristic's choice is pairwise matched with the choices of each of the other heuristics from the given set and for each game parameter combination. The payoff of the heuristics is determined for each game. In the next step, the values of a single heuristic are summarized as average over all games and saved. In the final part, the results are analyzed to determine successful patterns from the EIP sequence and the performance of the heuristics. From the performance results a ranking of the heuristics is created.

FIGURE 5 – SIMULATION PROCEDURE



In total, 16,776 games with unique parameters are studied.[41] Considering the application of the nine heuristics to these games, a data set resulting from 150,984 games is provided for analysis.

TABLE 16 – SIMULATION SETUP

Heuristics	**Game type**	**Player type**	**Task complexity**	**Time limit [step size]**
Random *Altruism* *Optimism* *Pessimism* *Naïve* *Equilibrium* *D1* *D2* *L2*	Hawk-Dove Chicken Prisoners' Dilemma Battle of Sexes Throwing Fingers Stackelberg's Leader-ship Costa 2A Costa 2B Costa 3A Costa 3B	row column	2 by 2 3 by 3 4 by 4 5 by 5	3 to 1400 [3]

5.6 Technical Implementation of Simulation

The technical implementation of the simulation requires a development and execution environment that provides specific properties regarding functionality, capability, and availability. Proper data handling, analysis functionalities, and presentational abilities are relevant here. Besides this, a professional introduction, as well as a large community with many support and tutorial sites, is undoubtedly of great help, too. Finally, handling knowledge, experiences, availability, and product price are essential arguments to consider. Several development environments meet those requirements.

In the end, *Matlab* is chosen, offering all of the opportunities above. The software is known to the author of this treatise from undergraduate studies and provided via an available campus account, evoking no additional costs. Based on matrix operations, it is optimized for multi-attributive pur-

[41] Half of the game types given in Table 16 have symmetric payoff matrices (Hawk-Dove, Chicken, Prisoner's Dilemma, Battle of Sexes, and Throwing Fingers). In this case, the player type hence need not to be considered. Those five game types are studied in four levels of complexities (equal to 20 constellations). Four game types are only played in a task complexity of 2 by 2 (Costa 2A, Costa 2B, Costa 3A, and Costa 3B). Since they are asymmetric, those four game types must consider both player types (equal to 8 constellations). Stackelberg's Leadership is the only game type that is asymmetric and studied in four levels of complexity (equal to 8 constellations). Those 20+8+8=36 constellations face certain time limits. With a step size of three EIPs, 466 time limits are realized. In total, 36*466=16,776 parameter constellations are studied.

poses, as found in the context of this research. These features make *Matlab* a powerful tool, already applied in many scientific projects and especially useful in the case of this study.

The simulation is executed in *Matlab* (in version R2014b) using the software-own syntax for programming all modules of the simulation and the analysis program. It is further used to generate diagrams and other figures to visualize results. The heuristics production systems are programmed as stand-alone modules that are encapsulated by a call procedure. That procedure traverses all parameter constellations. The results are stored in tabular form, which allows the various *Matlab* functions to calculate the realized payoffs and perform analyses. Both calculation and analysis are programmed as independent modules.

5.7 Expectations about Heuristics' Performance

In the present context, heuristics are understood as problem-solving strategies. They are interpreted as processes as proposed in the preparation time concept (Chapter 3). Following this information processing approach, a set of rules define the heuristics' way of proceeding containing structured operations to solve problem tasks. The application of the rules results in a sequence of operations characteristic of each heuristic. The sequence's length depends on the complexity of the given task and the time available. Thus, differences in performance are highly expectable within the given set of heuristics as well as with changing the games' task variables.

The production systems implicate that strategic heuristics apply substantially higher numbers of EIPs during their process compared to nonstrategic heuristics. Higher complexities of problem tasks promote this tendency due to the growing amount of information available to process. Thus, execution of the process of strategic heuristics generally takes more time than those of nonstrategic heuristics.

Under substantial, say effective time limitation, the process of strategic heuristics is interrupted, and the corresponding production systems define a random choice between remaining alternatives. If the set of alternatives could have been further reduced with more time available, the heuristics' choice is assumingly suboptimal compared to the case with no time limitation. It is thus generally expectable that the strategic heuristics' ability to generate a proper choice suffers from time limitation. Also, anticipating an equilibrium strategy – as far as it exists at all in a game – seems generally complicated under such circumstances.

In dominance-solvable games, players could figure out their best alternative by eliminating dominated alternatives[42] until only one for each player remains. Choosing a dominated alternative in such games causes inevitable losses in payoffs – regardless of the opponent's decision. Thus, especially in dominance-solvable games choice has a substantial impact on the resulting payoff. Since identifying dominated alternatives is a somewhat effortful procedure, payoff generation is also likely to suffer from time limitation. With the available processing time increasing, at least short before the time limit is no longer effective, the strategic heuristics should be generally better in payoff acquisition compared to nonstrategic heuristics, since they generate a strategic decision. However, as long as the procedure of eliminating dominated alternatives is running, a forced decision might be suboptimal. Of course, this issue depends on the game task and its concrete payoff structure.

It is expected that nonstrategic heuristics will make a quicker selection than strategic heuristics under comparable time pressure conditions as they have less information to process. However, the early termination of their production systems should also have an impact here. Also, nonstrategic heuristics cannot generally comprehensively evaluate the given information set. The recommended choice could thus be suboptimal in terms of payout generation compared to a possible equilibrium selection.

[42] A standard method in game theory to solve this issue is 'iterated elimination of dominated strategies'. For a formal description of this method see for example Leyton-Brown and Shoham (2008, p. 20), where also a definition of the term 'dominance' is provided.

6 Evaluation of the Simulation Results

This chapter presents in the first of two sections the data analysis approach for the simulation data. The approach follows the performance evaluation concept developed in Section 5.1. The following section presents the simulation results that take into account the heuristics' performance. Some of the results are considered in the following chapter, which presents successful EIP sequence patterns from the simulation.

6.1 Analysis of Simulation Data

The heuristics' performance is analyzed twofold regarding the proposed concept with its four goals: first, for each heuristic isolated, the fulfillment of each goal is examined by absolute values. Second, the heuristics are compared, resulting in a ranking for their performance.

The individual examination considers absolute values representing the arithmetic mean of the heuristic's performance overall game types for a particular constellation of time limitation, level of complexity and goal. The arithmetic mean depicts the dependent variable. Time limitation, level of complexity and goals depict the independent variables. For each parameter value of the two independent variables 'complexity' and 'goals', a single diagram is presented. In this, the arithmetic mean is only dependent on the discretely varying time limitation. A comparison of the resulting graphs helps to evaluate the influence of complexity and goals. This way, the complexity of data presentation is reduced, supporting the comprehension.

For interpreting data of Goal 3 (Generated payoff), it has proven helpful to conduct a linear regression. The graphs are supplemented by a first-degree-polynomial[43] with a form shown in Eq. (1), employing the method of Least Squares (abbr.: LS)[44]:

$$y = p_1 x + p_2 \tag{1}$$

with y as predictor data of Goal 3, $p_1, p_2 \in \mathbb{R}$ as coefficients of the linear function/fitting parameters and x as time limit data.

The LS follows the Gaussian principle of regression by minimizing the sum of squares of the residuals of observed data and predictor data (Eq. (2)).

[43] Although response data of the heuristics partly suggest higher, nonlinear polynomial-fitting, the linear version is chosen to show the underlying overall gradient as the tendency of the developing performance.
[44] The calculations are performed by deploying the Least-Squares-Fitting tool of *Matlab* which also provides introductions to Eqs. (1) to (4) in its documentation.

$$d = d(p_1, p_2) := \sum_{k=1}^{n} (y_k - p_1 x_k - p_2)^2 \tag{2}$$

with d : sum of squares of residuals, x_k , $y_k \in \mathbb{R}$, x_k : kth time limit, y_k : corresponding heuristic performance (regarding Goal 3) and n : number of response data points.

Following the general minimum condition, Eq. (2) is differentiated for each parameter p_1 and p_2 and set equal to zero. After solving the system of equations for each parameter, the estimator for the coefficients of the linear model, $\hat{p}_1$ and $\hat{p}_2$, can then be calculated by Eqs. (3) and (4):

$$\hat{p}_1 = \frac{n \sum_{k=1}^{n} x_k y_k - \sum_{k=1}^{n} x_k \sum_{k=1}^{n} y_k}{\sum_{k=1}^{n} x_k^2 - (\sum_{k=1}^{n} x_k)^2} \tag{3}$$

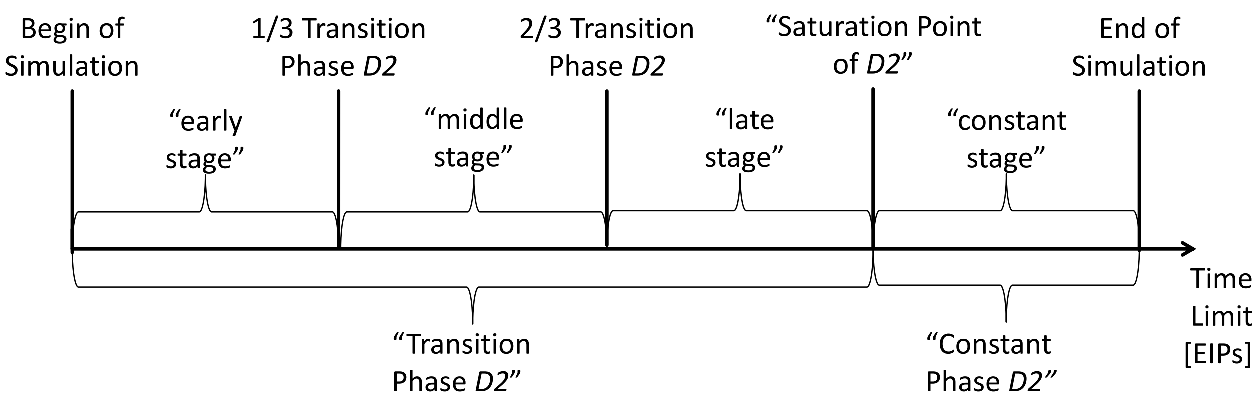
$$\hat{\mathrm{p}}_2 = \frac{1}{n}\left(\sum_{\mathrm{k}=1}^{\mathrm{n}} \mathrm{y_k} - \hat{\mathrm{p}}_1 \sum_{\mathrm{k}=1}^{\mathrm{n}} \mathrm{x_k}\right) \tag{4}$$

with $\hat{p}_2$ calculated by inserting the value of $\hat{p}_1$ of (3) in (4).

As mentioned above, the second interest is in comparing the heuristics' performances to each other. For the comparative analysis, relative numbers giving the rank of each heuristic per goal are studied. The ranks in Goal 3 (Generated payoff) follow equally weighted results from each heuristic playing only once against each other heuristic from the set. So, in the implicitly simulated population of players, each bound to one heuristic, the heuristics' share are equally distributed. That could be a stark simplification of reality to assume that players use one heuristic only. Despite that, other authors use a similar approach in their simulations. Johnson and Payne (1985, p. 400) for example relate the result of a heuristic's choice to a lower bound – the *Random* heuristic. All outcomes are interpreted as "relative improvement over random selection" (Johnson and Payne 1985, p. 397). They give no explicit assumptions concerning the distribution of heuristic proportions within their studied population. However, the simulation presented here is not aiming at reproducing a realistic population of player behavior. Furthermore, a lower bound is not regarded as necessary here, since all results are comparable.

The influence of time limits on the ranking is investigated within four ranges of particular interest. As will be emphasized in the following section, a 'transition' phase and a 'constant' phase define the behavior over time. The transition, in turn, is again categorized into three stages of equal length, namely the ‘early’, ‘middle’ and ‘late’ stage, with the late stage ending simultaneously with the transition phase, as illustrated in Figure 6. The minimum time limit represents the lower bound. The beginning of the saturation stage of the slowest heuristic, which has proven to be *D2*, defines the upper bound for all heuristics. The bounds frame the range of the transition phase for the set of heuristics. This segmentation of time limits is applied to all levels of complexity.

FIGURE 6 – PHASES AND STAGES OF HEURISTICS’ PERFORMANCE



The ranks of the heuristics regarding the three supplementing goals are presented in form of a two-dimensional figure. The first two goals are located on the Cartesian coordination system. The shades of gray illustrate the ranks regarding the third goal (see Figure 17 and following). Each of the axes represents one goal with the ranks as the dimension, starting with the highest rank numbered with '1'. This presentation form neglects a tremendous amount of information on the actual interaction of heuristics. However, this is traded in for clarity and comprehensibility. The so derived results can finally be interpreted.

As the examined time limitation range is divided equally into three parts and the goals are weighted equally, the ranks of the heuristics are determined by calculating the arithmetic mean of goal fulfillment for each of the time stages with Eq. (5):

$$\bar{x}_m = \frac{1}{n}\sum_{i=1}^{n} x_{m,i} \tag{5}$$

with $\bar{x}_m$: mean of goal m , n : number of measurements per stage and $x_{m,i}$: measured variable for goal m at time limit of i * [step width] * stage number.

In the next step, the mean values of all heuristics for each stage are ordered. The heuristic associated with the highest value of all means obtains rank number one and so on. In the case of equal values, the corresponding heuristics obtain the same rank. The subsequent heuristics then are ranked with the number increased by the number of equally ranked predecessors.[45]

6.2 Results from Simulation

The following figures presented here show the results of the heuristics regarding the three predefined goals over increasing levels of time limitation in the number of EIPs and for specific complexities.[46] The rankings are displayed in Figure 13 (p. 82) and following.

One similarity of the heuristics' performance curves throughout all goals and all complexities is their two-stage form. Whereas the first part shows a significant change of the goal's measurement value, the second shows static behavior. This behavior is expectable, when interpreting the beginning as a transition or adaptation phase, which later results in saturation or virtually infinite period of constancy. It seems quite evident that only in the first period the time limit is effective in the sense that it affects the goal achievements of the heuristics. The subsequent phase reflects gameplay without time limitation. The turning points, where one phase changes into the other, are characteristic for the heuristics and alter with complexity. That behavior is expectable.

6.2.1 Goal 2 – Reduce Number of Alternatives

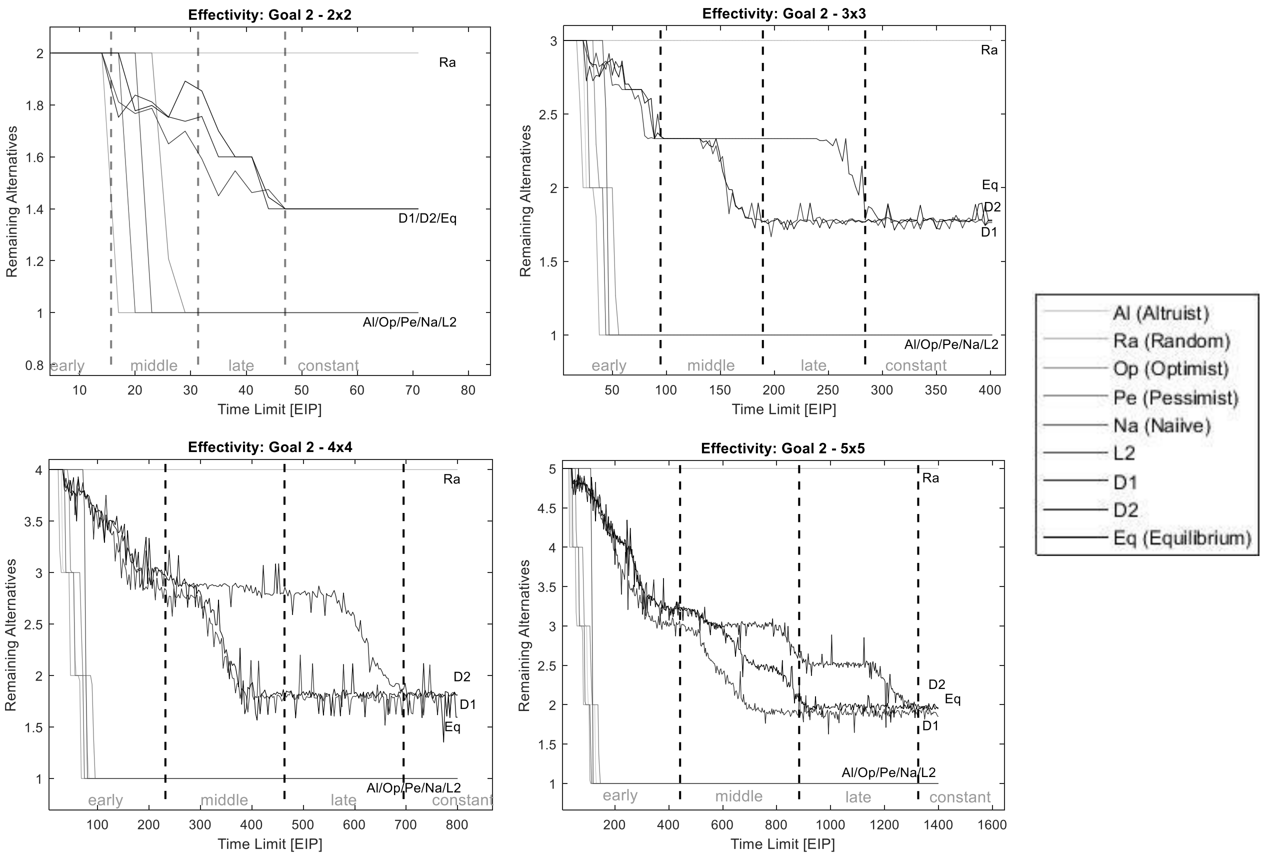
This goal reflects the effectiveness of removing irrelevant alternatives from consideration. Even though the underlying mechanisms for labeling an alternative as 'irrelevant' can be diverse, this term in its global meaning has the same implications throughout the various heuristics. The different mechanisms are hence the critical factor for the diversity of results in the set of heuristics.

A noticeable similarity between all heuristics – except *Random* – is the effective influence of time and complexity on the ability to reduce the set of alternatives. Figure 7 shows this relation for the set of heuristics, for different levels of complexity. Whereas time limitation reveals a negative influence, complexity shows a positive one. Apparently, the number of reduced alternatives

[45] Take two heuristics for example which share rank five because of their means. The subsequent heuristic would then earn rank (5+2 =) seven.
[46] It should be noted that the heuristics' performance is determined for discrete time limits, but is plotted as a line for easier visualization and hence for better comprehension.

heavily depends on the number of alternatives available at the beginning. This number increases with the level of complexity.

FIGURE 7 – GOAL 2 REDUCTION OF ALTERNATIVES



The constant value of *Random* is already mentioned in the context of developing the corresponding production system (Section 4.2) and now expectably confirmed by simulation. In the following, the emphasis is on a complexity level of 2 by 2 alternatives. Five heuristics pass through the adaptation process at time limits between 15 and 30 EIPs, reducing the set of alternatives to one (Altruist, Optimist, Pessimist, Naïve ,and L2; see Figure 7). With about 48 EIPs the heuristics *D1*, *D2* and *Equilibrium* require noticeably more EIPs to reduce the set of alternatives to a minimum. That minimum is well above the one of the other five heuristics. With increasing complexity, the differences in terms of effectivity between the two groups increase. Of the three slowly adopting heuristics, *D2* requires the most EIPs to minimize the set of alternatives. The five heuristics that adopt comparably fast are in the following referred to as 'early adopters'. These early adopters comprise all nonstrategic types of the heuristic set, complemented by the strategic heuristic *L2*. Complementary to the group of early adopters, the remaining strategic heuristics are referred to as 'late adopters', regarding their ability to reduce the set of alternatives. Among the group of early adopters, the heuristic *Optimist* shows the best performance throughout all levels of complexity in its ability to effectively (regarding quantity and velocity) reducing the set of alternatives in a game (also compare Figure 13).

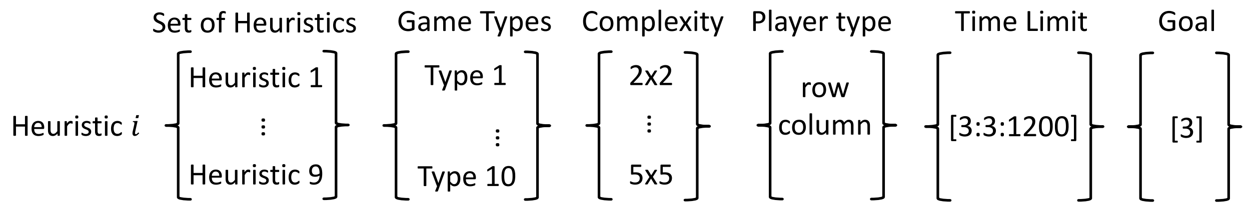
Another remarkable finding concerns the adaptation phase of the heuristics. Two different behaviors can be identified. A weakly monotonous decrease of the number of remaining alternatives characterizes the first one. The staircase structure hints to the discrete nature of the heuristics' adaption process. It is generally fast compared to the second examined behavior, which can be called oscillating. The oscillating behavior is altogether decreasing, yet not monotonously. Again, the observed behaviors match well the adopter classification. The early adopters show a steady and the late adopters show an oscillating performance behavior. The oscillation in turn is quite similar within the three strategic heuristics.

The oscillating performance graph of the late adopters, especially strong in the early stages of the games, is caused by the random variables that affect their production systems. In the early stage of a game, the production systems begin with randomly choosing the starting focus. It could either lie on the own payoff or the other player's payoff. This focus choice initializes the subsequent elimination process of dominated alternatives. Additionally, the starting point within the payoff matrix as well as choice between remaining alternatives is random. When the focus has changed once and one full round of eliminating dominated strategies is completed, the influence of the randomly chosen starting focus is leveled. The behavior of the performance graphs then converges towards the staircase form. However the influence of the other two random variables is still visible.

6.2.2 Goal 3 – Maximize Generated Payoff

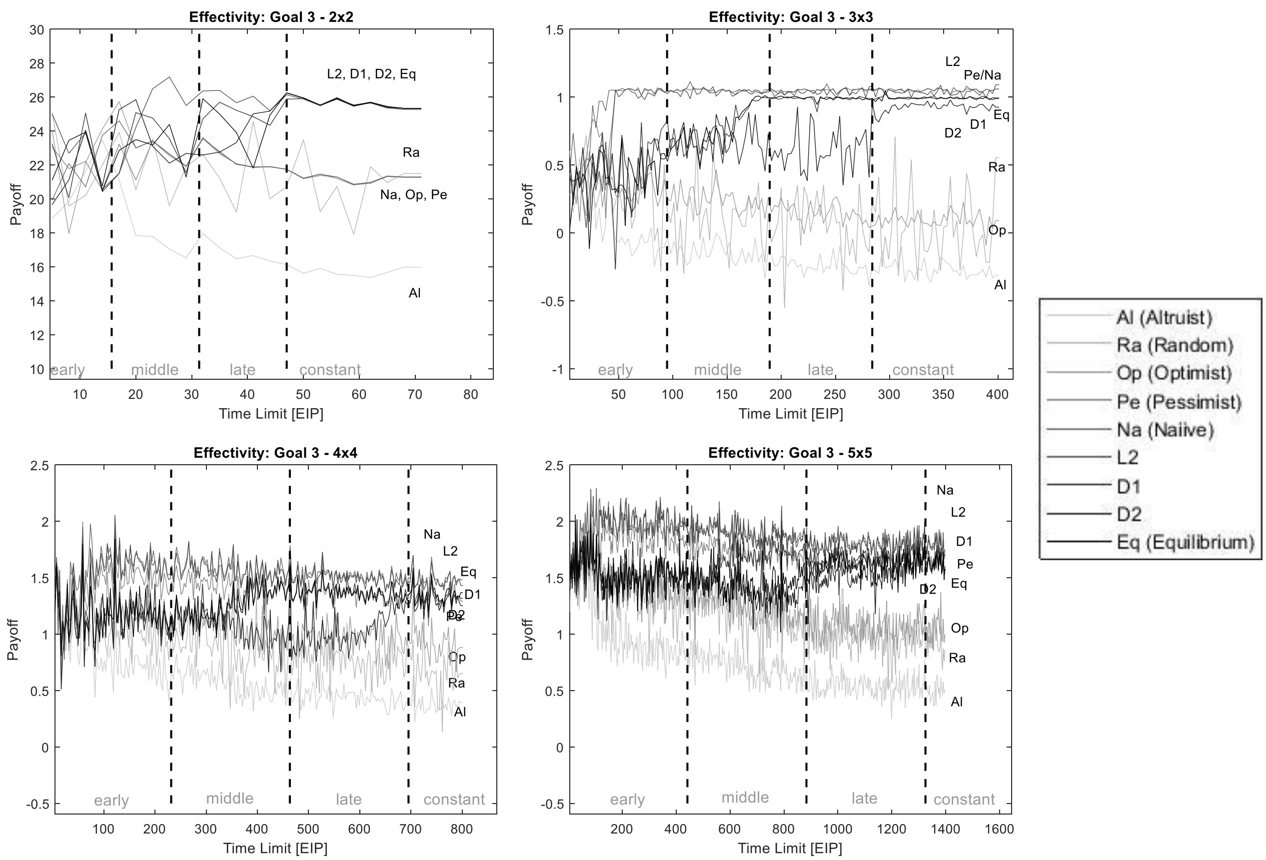
This goal represents the effectiveness of a heuristic concerning the resulting payoff it creates when playing against other heuristics. Therefore each heuristic of the given set is matched with any other, including itself, for any time limitation at any complexity and in any game type, as illustrated in the following scheme (Figure 8).

FIGURE 8 – GENERATING DATABASE FOR ESTIMATING PAYOFF



The performance of the heuristics against each other is very heterogenic. The shape of the performance curve is volatile, if not oscillating. For an overview see Figure 9.

FIGURE 9 – GOAL 3 GENERATED PAYOFF



Focussing on a complexity level of 2 by 2 first, one can roughly differentiate three kinds of paths in heuristic behavior. They differ in the level of payoff and begin right after a phase of substantial interference. The heuristic *Altruist* generates the smallest payoff (very light gray line). *Random*, *Pessimist*, *Optimist*, and *Naïve* generate medium payoffs. Finally, *L2*, *D1*, *D2*, and *Equilibrium* create the highest payoff values. Within each of the higher payoff levels, the results of the heuristics vary a lot, making it difficult to rank them by sight. At a complexity of 2 by 2, the payoffs range from about 15 to 28 units. For higher levels of complexity, this span is much smaller (about three units and less in total). However, the three paths can still be identified, yet less clearly. Also, the composition is slightly changing, and the phase of inference at the beginning lasts longer.

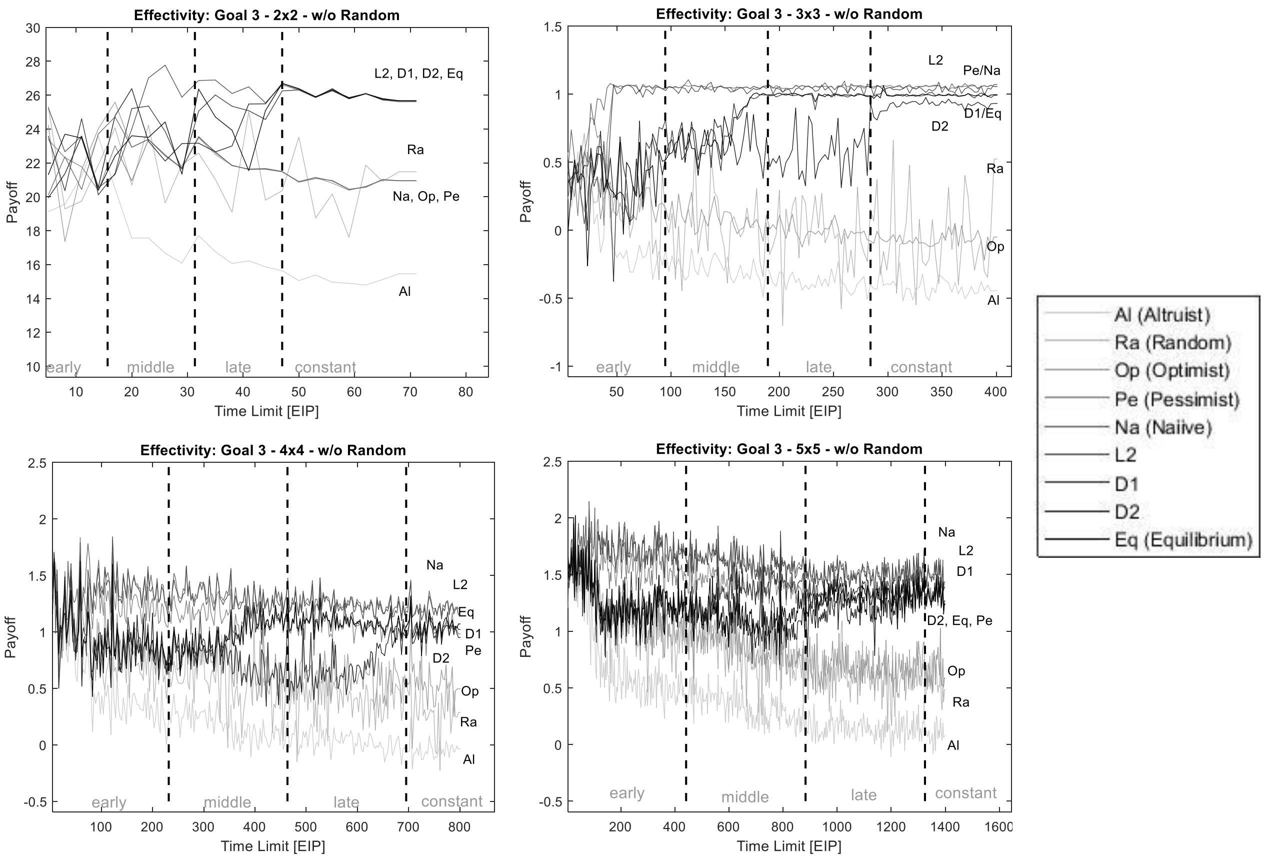
The influence of time limitation on payoff generation is expected to be noticeable. Their values, however, change at high frequency, most likely due to heuristics interactions. It is therefore difficult to recognize or evaluate the influence by analyzing the plots in this form. The shape of the plots differs greatly from the expected two-phase form. Neither the transition nor the constant phase can be distinguished as quickly as in case of Goal 2 or Goal 4. The time limit where both phases meet – henceforth referred to as saturation point – remains unclear.

The unsteadiness of the heuristics' performances, concerning growing time limitation is caused by the random variables that affect the production systems. They hence influence the heuristics' interactions: a random choice between remaining alternatives, as well as randomly chosen starting focus and starting points within the payoff matrix. The first variable is assumed to have the most significant impact on the oscillating shape with the *Random* heuristic being the main contributor regarding amplitude and frequency. Besides the *Random* heuristic, the strategic heuristics include this element of random choice in their procedures as ultimate action when time expires. For strategic heuristics this behavior is thus expected to be observed throughout the transition phase, but not in the constant phase. However, this impact is difficult to observe due to interferences with the impact of the *Random* heuristic. To better evaluate this assumption the influence of the heuristic *Random* is excluded from the performance calculations by discarding play against the *Random* heuristic (Figure 10).[47]

The curve form still reveals an observable oscillation. Here, the other two reasons additionally take effect. While in the transition phase the oscillation results from the interference of both reasons, in the constant phase the slightly moderate oscillation (regarding amplitude and compared to the interference effect of both reasons) is exclusively based on the also randomly chosen starting focus and starting point of observation.[48] In analogy to the case of *Random* choice, all heuristics are affected again since all heuristics play against each other. The amplitude of the oscillation is in this context determined by the interference with the three strategic heuristics.

[47] This figure still contains the results for the *Random* heuristic as own graph.
[48]. Section 5.4 introduces the terms 'starting focus' and 'starting point'.

FIGURE 10 – GOAL 3 GENERATED PAYOFF WITHOUT *RANDOM*



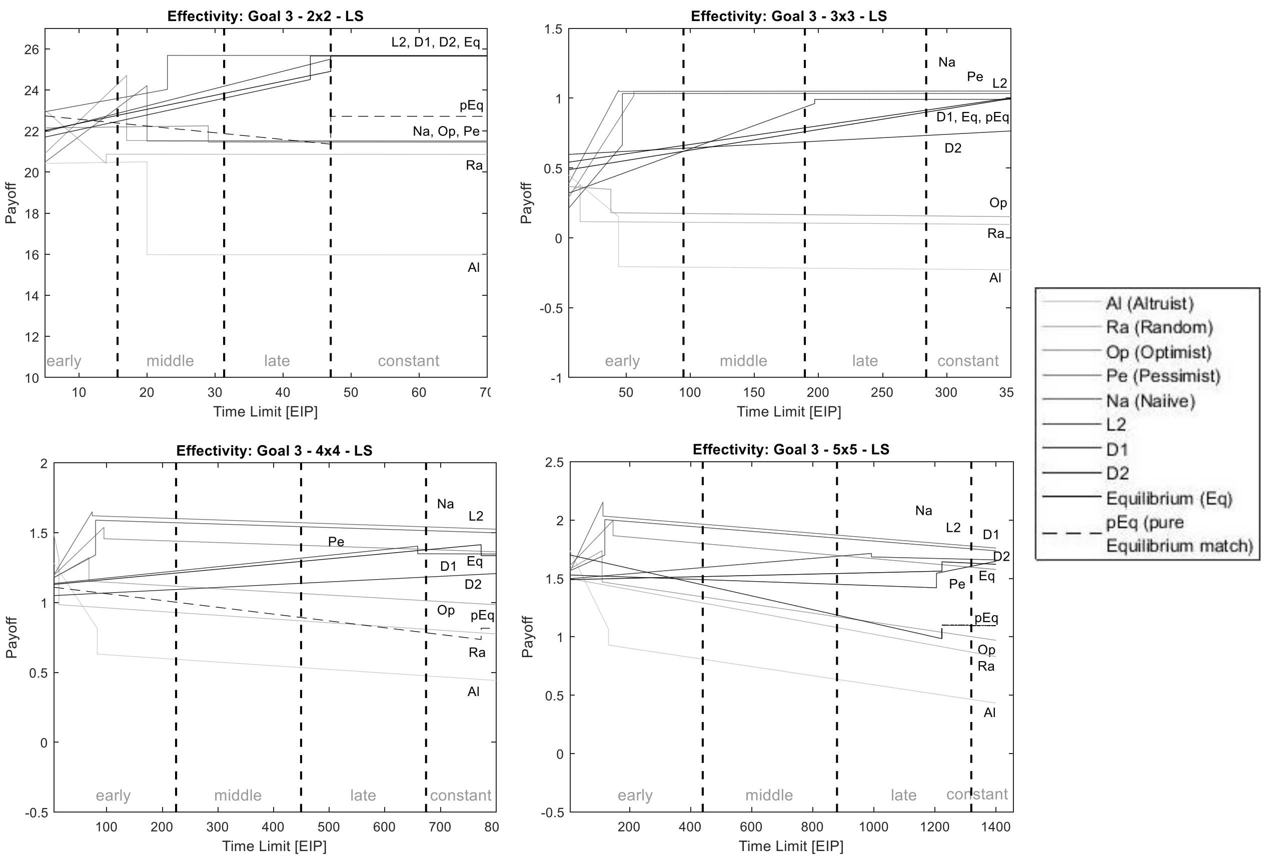
The graphs depicted in Figure 10 are still confusing due to interfering oscillation. Furthermore, it can hardly be determined whether the heuristics' graphs decrease or increase with growing time limitation, making it difficult to describe the performance. Hence, the single heuristics' curves are transformed by the method of LS and fitted to a linear model with Eq. (1).

In analogy to the examination of Goal 2 and Goal 4 it proved useful to divide each curve in its transition phase and its constant phase and transform both phases individually. Assuming that the saturation point is identical to the one in case of Goal 2 and 4, it is further applied to determine the dataset that belongs to each of the two phases.

The combined results are depicted in Figure 11. Both phases are of a linear character and hence are markedly distinguishable throughout all heuristics and for all levels of complexity. Since they share the saturation point, the two phases connect through a vertical line. As a benchmark, the payoff curve of the *Equilibrium* heuristic exclusively playing against itself is added to the set of performance curves (black dashed line).

The bipartite shape is still noticeable, revealing the compelling influence of time limitation. The outcome of the transition phase results from the ongoing reduction of the set of alternatives and the random choice at the moment the time expires. Depending on the combination of several game task parameters (see Table 16), especially the degree of complexity, this has different implications for the generated outcome.

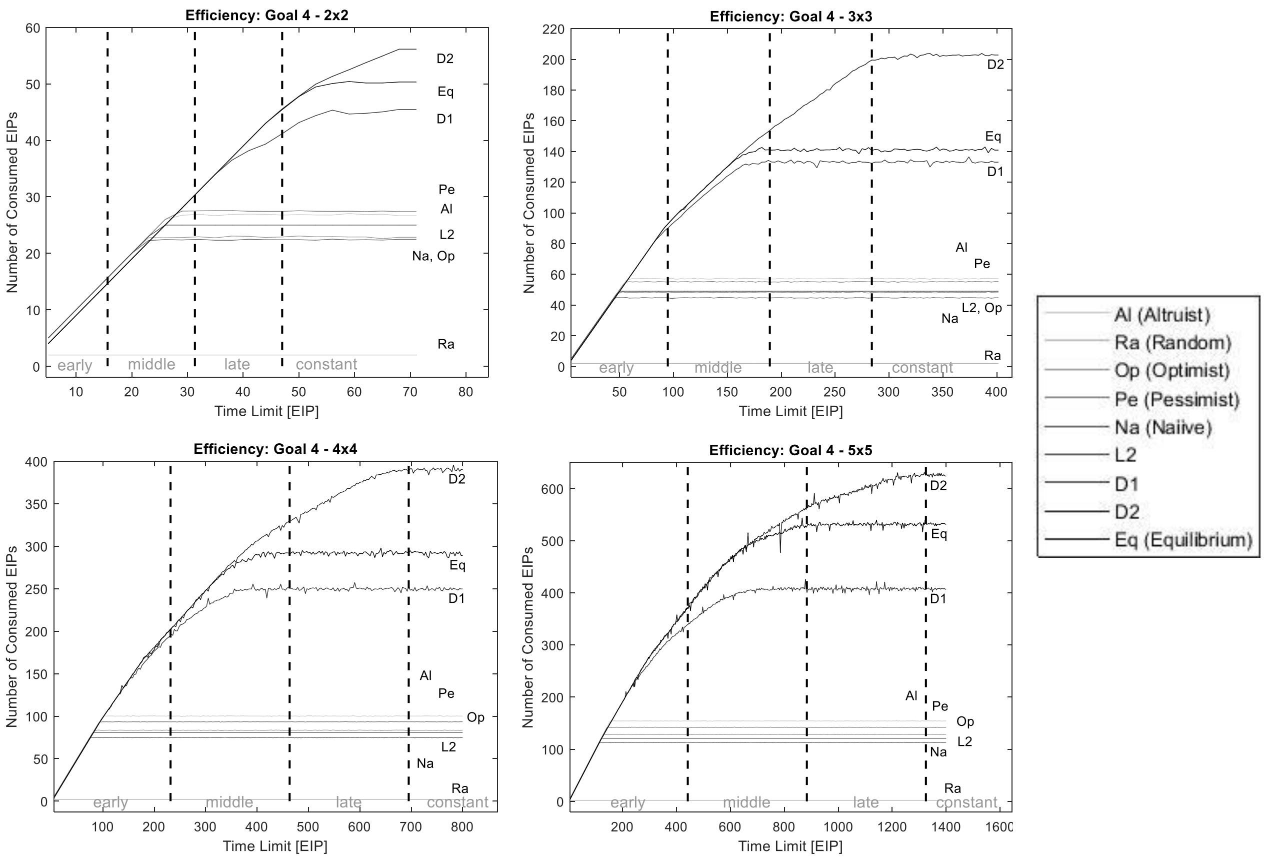
FIGURE 11 – GOAL 3 GENERATED PAYOFF "LEAST SQUARES" - PROJECTION



6.2.3 Goal 4 – Minimize Number of Applied EIPs

This goal reflects the efficiency, with which a decision is generated: the fewer EIPs are applied during a heuristic's procedure, the more efficient it is. The two-phase shape of the heuristic's performance, transition and saturation, is also noticeable here (Figure 12). Solely *Random* shows a static behavior again – this time on a shallow level, due to minimal use of EIPs.

FIGURE 12 – GOAL 4 APPLIED EIPS



The graphs of the nonstrategic heuristics along with *L2* grow linearly at the beginning until they reach their saturation points. At these points, the slope is changing abruptly to zero. An obvious explanation is offered by the heuristics' production systems, which apply a certain amount of EIPs for a certain complexity and game type. While the number of available EIPs (symbolizing an effective time limit) is smaller than this amount, all available EIPs are used to proceed, with the last one always reserved for the respective last action[49] of a heuristic. A number of EIPs equal to or larger than the number of EIPs needed to end the procedure regularly (i.e., no random choice) symbolizes an ineffective time limitation. The heuristics' procedures apply EIPs until its completion, but no more afterward.

The linearity of the slope shall be discussed in more detail now. If the nonstrategic heuristics' behavior is sensitive to the game types examined here, one would expect a heuristic's procedure to apply variable amounts of EIPs. The slope of the curve would hence change over time since it

[49] In case of non-strategic heuristics, together with *L2*, the last action is choosing the currently best alternative with the last EIP available. The other strategic heuristics randomly choose one alternative among the remaining in that case.

is estimated as a mean for the heuristic's performance throughout all game types at a certain complexity. A change of slope cannot be detected. This leads to a remarkable conclusion: the nonstrategic heuristics including *L2* are not sensitive to game types in terms of EIP requirements.[50] The relation to the games' complexity is evident, suggesting quadratic relationship between the number of EIPs and the number of alternatives, given a fixed time limit. That relation results from the number of elements in a payoff matrix which increases quadratically with growing level of complexity. The number of elements directly affects the heuristics' information gathering. The linear form within the transition phase points to a perfect correlation. It underlines the influence of time limitation on Goal 4.

The other three strategic heuristics show this initial behavior, too. However, the transition slope is not constant, but changes at least once (and becoming smaller) before it reaches the saturation phase. In opposition to nonstrategic heuristics, the various game types induce such behavior. The strategic heuristics need different amounts of EIPs for their procedures. This is because of the game types' different structures. Every time a heuristic fully processes at least one game type, its performance graph changes to a smaller slope. Such behavior is perceived until the heuristics complete all game types, and the graph ultimately changes into the constant phase. This result now implies two conclusions: first, the set of game types is heterogeneous enough to cause different amounts of EIPs per strategic heuristic. Second, the procedures of the strategic heuristics are sensitive to game types. Again, the linear slope of the performance curves in the transition phase leads to the conclusion that time limitation and fulfillment in Goal 4 is (almost) perfectly correlated. Drops in the slope are caused by the payoff structures of the various game types.

The graphs of the nonstrategic heuristics stay close together throughout all levels of complexity compared to the block of strategic heuristics. Nonetheless, the distance between the heuristic with the lowest number of applied EIPs and the one with the highest when reached the saturation phase is nearly doubling with increasing level of complexity. Herewith, the degree of inference successively decreases. The heuristics further change ranks in between, with *Altruist* and *Optimist* losing and *L2* winning ranks. In case of the strategic heuristics, the development is entirely different. Whereas the ranks within the three heuristics stay unaltered, the distance between high-

[50] This result is limited to the set of game types examined here. Nonetheless, the set contains a broad variety of types, as can be seen in Table 14. Hence, the generality of this result can be expected.

est and lowest number of applied EIPs is tripling in the first step, less than doubling in the second and less than 1.5 times in the third.

At Goal 4 one can once again remark the oscillating forms of the strategic heuristics *D1*, *D2* and *Equilibrium.* It is clearly visible only for higher complexities (starting with 3 by 3). However, for a complexity of 2 by 2 the effect occurs on a smaller scale. The oscillation starts at the end of the early period of the transition phase and shows a more or less gradual transition to the constant phase. The underlying slope is slowly approximating zero. Once having reached this phase, the number of EIPs is oscillating within a small range. Again the reason for this behavior can be found by taking the game types in combination with the random variables into account. For non-symmetric game types, a dominated alternative sometimes exists for just one player's set of pay-off. Thus, the starting focus and the starting point sometimes determine how fast the procedure terminates and hence how many EIPs are used in total. These random choices can be discovered in the oscillating behavior of the heuristic's performance graph.

The production system of *D2* includes two rounds of checking for iterated dominance by definition – regardless of results from round one (which for example has shown that there is no dominated strategy in the game). The maximum level of reasoning of the games examined in this study is two, with the elimination procedure solely concentrating on pure strategies. Hence, applying *D2*, similar to *Equilibrium*, always means eliminating the maximum number of alternatives by employing a game-theoretic solution concept. However, the production system of *D2* is, unlike the *Equilibrium* heuristic, inflexible in its termination criteria. The constantly larger amount of applied EIPs during the saturation phase compared to *D1* and *Equilibrium* illustrates this predetermination.

6.2.4 Effectiveness and Efficiency in the Heuristics' Performance

The detailed analysis of the heuristics' performances for each goal is now summarized regarding effectiveness and efficiency as proposed in Section 5.1. The earlier introduced classification of heuristics seems helpful again in the following consideration. Goals 1 to 3 contribute to the effectiveness aspect of the performance, whereas Goal 4 symbolizes the efficiency as argued earlier. In the following, the sensitivity of the heuristic's performances towards a reduction of time limits is examined. The argumentation for increasing the time limit is analogous.

Goal 1 (Propose a decision) is fulfilled by all heuristics by definition already for very small time limits which in the following are characterized as reasonable. With this, a decreasing time limit has no effect on the heuristics' effectiveness regarding Goal 1. It is thus characterized as nondiscriminatory among the given set of heuristics and not further considered when comparing the heuristics' performances.

Regarding Goal 2 (Reduce the set of alternatives) all heuristics show dependency towards time limitation. With decreasing time limits (the time that can be used to execute the production system of a heuristic), the fulfillment of Goal 2 decreases, too. The set of heuristics can be grouped according to their performance at Goal 2. Nonstrategic heuristics, together with *L2*, reduce the set of alternatives faster than the remaining strategic heuristics. This reduction speed makes them more effective under decreasing time limits. However, a growing complexity generally seems to increase this dependency.

Goal 3 (Maximize certain payoff) shows intense oscillation among all heuristics in their transition phases. It is thus hard to determine how a heuristic's payoff generation is developing over changing time limitation. As suggested earlier, conducting a linear regression on the data helps to ascertain the impact's tendency. The slope of the resulting first-degree polynomial expresses the tendency. A regression is conducted for both phases and for every heuristic individually.

In the transition phase the performance graphs of the heuristics show dissimilarities regarding its sensitivity to a decreasing time limit. Most of the heuristics face a declining payoff output under such conditions. Only the heuristics *Altruist* and *Random* improve their output with sinking time limit over all levels of complexity. *Optimist* (3 by 3) and *Pessimist* (2 by 2) improve their payoffs for one level of complexity. However, the payoff generated in the early stages of a game strongly depends on the randomness of choice. Increasing the outcome with lesser strict time limits reflects the heuristics' ability to reduce the influence of randomness on their choices. This is a property that is more pronounced in strategic heuristics.

All in all, the heuristics' ability to adapt to increasing time limitation is individually different and dependent on task variables as well as random variables. This heterogenic behavior disables a general statement on the influence of time limitation on effectiveness regarding Goal 3. Also, with this examination of the individual heuristics and their adaptability, nothing is said about the heuristics' performance when compared to each other. This estimate is discussed in Subsections 6.2.7 to 6.2.10, where the heuristics are ranked, enabling reasonable recommendations.

Goal 4 (Minimize Effort) is directly related to the number of EIPs the proceeding of a heuristic applies. Since the number of EIPs measure the time limitation in this simulation, the consumption and hence the effort directly links to the time limit. As presented earlier, all heuristics show a direct relation to the time limit during their transition phase. It is thus not surprising that a reduction of the time limit affects a reduction of the applied EIPs. All heuristics become more efficient when the available time shortens.[51] Nonetheless, the two groups of heuristics mentioned above show significant differences. The nonstrategic heuristics, including *L2* reach their constant phase comparatively early and the remaining strategic heuristics late. The latter are thus at least equal or more sensitive to a reduction of time limitation as will be shown now in more detail.

Assume that a time limit is given which applies only to the group of late adopter heuristics.[52] Assume also that an extent of reduction is given that leaves a time limit equal to or greater than zero[53] after reduction. If the extent of the reduction is too small to reach the heuristics' transition phase, compression of the time limit will not affect the early adopter heuristics. Alternatively, if the reduction is substantial enough, an effect can be observed. Either way, it has in every case an effect on the remaining late adopter heuristics. This inequality can be generalized for larger time limits (reaching the constant phase of the late adopters) without further assumptions.

This consideration determined the qualitative influence of time limitation on every heuristic's effectiveness and efficiency. Table 17 provides a summary of nonstrategic and strategic heuristics. The results for Goals 1 to 3 could be combined with various combinations of weights to derive an overall result for effectiveness. However, there is no known argumentation in the literature that favors one goal over another. Hence an equal weighting of goals is suggested.

[51] Recall that a time limit can only cause an effect on the performance of a heuristic when the heuristic is still in its transition phase. Increasing or decreasing the time limit in the constant phase does not affect the performance.

[52] Recall that the group of late adopters is formed by *D1*, *D2*, and *Equilibrium*, whereas all other heuristics form the group of early adopters, following the notion introduced in Subsection 6.2.1.

[53] A time limitation that is smaller than zero EIPs / zero seconds is not defined.

TABLE 17 – INFLUENCE OF DECREASING TIME LIMIT AND INCREASING COMPLEXITY ON EFFECTIVENESS AND EFFICIENCY OF HEURISTICS

Criteria	**Effectiveness**								**Efficiency**
Goal	Ø G1-G3	G1	G2	G3					G4
Complexity	all	all	all	Ø all	2×2	3×3	4×4	5×5	all
Heuristics									
Ø non-strat. heuristics	−/0	0	−	+/−	+	+	−	−	+
Altruist	0	0	−	+	+	+	+	+	+
Random	0	0	−	+	+	+	+	+	+
Optimist	−	0	−	−	−	+	−	−	+
Pessimist	−	0	−	−	+	−	−	−	+
Naïve	−	0	−	−	−	−	−	−	+
Ø strat. heuristics	−	0	− −	−	−	−	−	−	+
L2	−	0	−	−	−	−	−	−	+
D1	−	0	− −	−	−	−	−	−	+ +
D2	−	0	− −	−	−	−	−	−	+ +
Equilibrium	−	0	− −	−	−	−	−	−	+ +

Summarizing Table 17 leads to general findings of the influence of time limitation and complexity on heuristics' performances: whereas effectiveness tends to be negatively impacted, the opposite is true for efficiency for all heuristics. Looking at each heuristic individually, some perform more robust under decreasing time limits (*Altruist* and *Random*). The effectiveness of those heuristics shows an invariant sensitivity (mean over Goals 1 – 3). No clear tendency for one of the two groups (nonstrategic and strategic) is observed either.

6.2.5 Comparison of the Heuristics' Performances

Based on the heuristics' results regarding their performances a ranking is determined for Goal 2 to Goal 4 and for each time limit.[54] Appendix C provides one-to-one comparisons of the heuristics and their performance for all goals. As mentioned in Section 5.2, the observation of the time limitation's influence concentrates on the four stages 'early', 'middle', 'late' and 'constant'. The heuristics' rank developments over time limitation are projected onto these four categories. This

[54] Recall that Goal 1 (Recommend a Choice) is not discriminatory.

provides a perspective that gives a good overview of the heuristics' achievements during the changing time limit. Figure 17 and following depict the results.

The diagrams display the ranks of Goal 2 (Reduce alternatives) on the horizontal axis and of Goal 3 (Generate payoff) on the vertical axis. Ranks of Goal 4 (Applied EIPs) are represented by shades of gray, associated with the underlying color map. All diagrams additionally contain two orthogonal lines parallel to each axis, intersecting in point $(x, y) = (5, 5)$. Those lines represent the median ranks of Goal 2 and 3. The median rank of Goal 4 is gray, with decreasing rank numbers (which are similar to better ranks) towards light gray and increasing towards dark gray. The median lines form four quadrants, helping to visualize the interpretation of the heuristics' performance. Heuristics in the lower left quadrant are performing better than the median for both goals 2 and 3, whereas those in the upper right quadrant perform worse than the median. The other quadrants' interpretation is straightforward. It should be noted here that a heuristic's ranking trajectory must not be viewed isolated. All heuristics influence the ranking with their developments over time. Hence, it is just a relational shift within the set of heuristics.

6.2.6 General Consideration

At the very beginning of the early stage, the heuristics perform quite similar throughout all goals. The graphs of the heuristics overlap much here, and the rankings reflect more or less marginal variances concerning goal fulfillment, with Goal 4 (Figure 12) showing the least differences and Goal 2 (Figure 7) and Goal 3 (Figure 11) following. The nonstrategic heuristics then often reach their saturation phase around halftime of the early phase of (the benchmark) *D2*, a little later for a complexity level of 2 by 2 and a little earlier for a complexity level of 5 by 5. In the middle stage of the time performance, the production systems of the nonstrategic heuristics are already finished. A change of ranks in the late and constant phase is thus exclusively resulting from the strategic heuristics' ongoing consolidation process.

This consolidation process reveals significant time differences throughout the group of strategic heuristics. Whereas *D1* and *Equilibrium* conclude their procedures in the outgoing middle and beginning late stage, *D2* ends by definition (of the stages) on the turning point from the outgoing late stage to the beginning constant phase and thus by far lasts longest of the set of heuristics.

Regarding time duration of procedures, the already introduced grouping of the set of heuristics in early adopters and late adopters is valid here. However, since fully processing *D2* requires markedly more time than for *D1* and *Equilibrium*, the group of late adopters needs to be split into two (group 1: *D1 + Equilibrium*; group 2: *D2*) and discussed separately. The groups maintain their structure throughout all levels of complexity. Of course, within each group differences exist on a smaller scale and those differences vary with growing complexity, too. However, the time differences between the groups are markedly higher compared to those differences observed within each group: with increasing level of complexity, the early adopters end its procedures about three times (at 2 by 2) up to six times (at 5 by 5) earlier than *D2*.

The time span between the two groups finishing their procedures rises with the level of complexity. The early adopters have significant advantages in the time needed to finish the procedures, even though the impact on effectiveness and efficiency is diverse. *D1 + Equilibrium* barely need half the time of *D2* for low complexities. In opposition to the early adopters, the relation then increases from about ½ (at 2 by 2, 3 by 3) up to about ⅔ (at 5 by 5), meaning that the time the procedures of the strategic heuristics need converges with growing complexity.

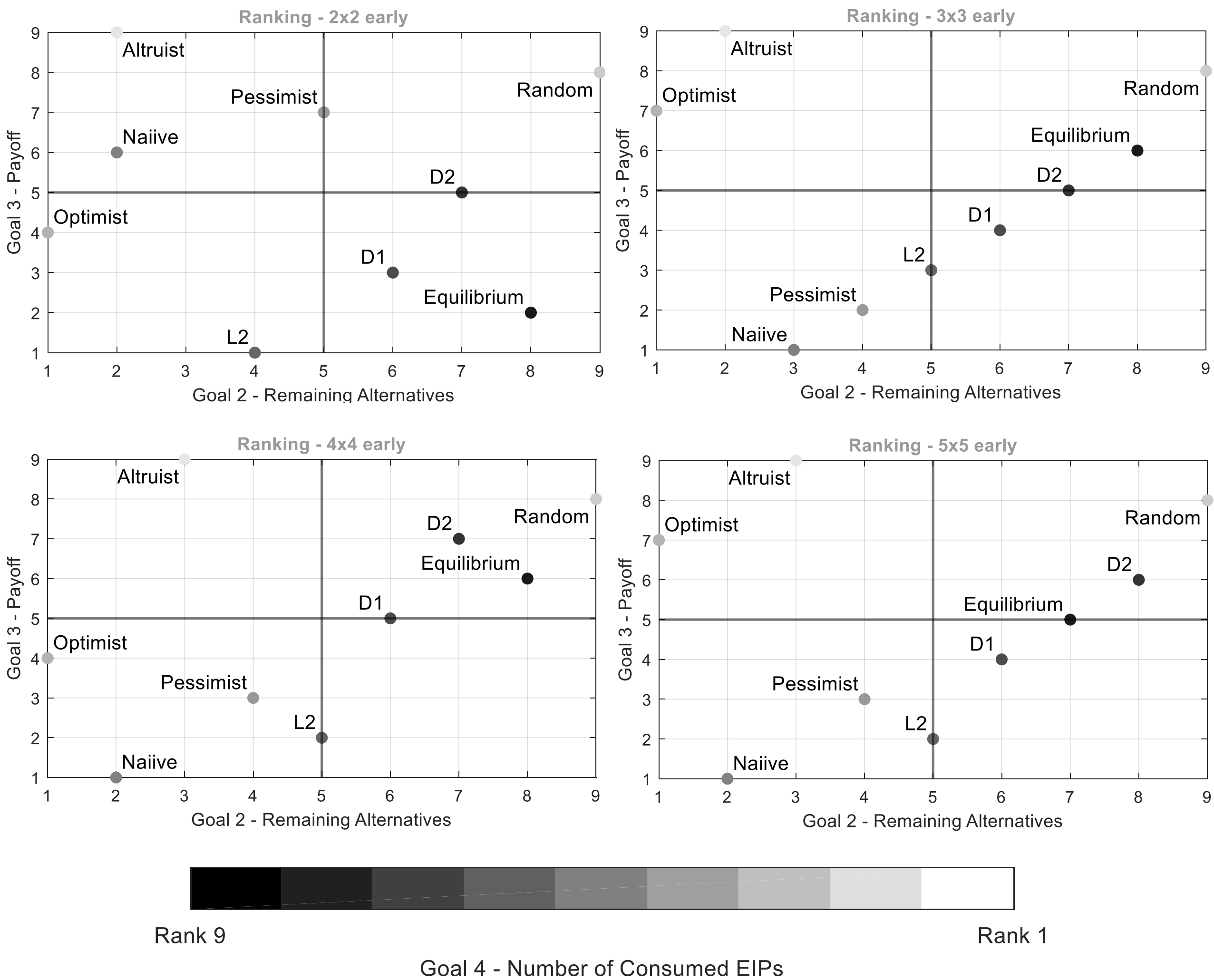
It is worth noticing here that for the interpretation of the heuristics' overall performance at each time stage the results in Goal 4 (Applied EIPs) play a minor role. The evaluation of one stage is based on the assumption that the effective time limit lies in this stage, too, with all heuristics evolving until this very point in time. If some heuristics comparably score at Goal 2 and 3, Goal 4 serves as a means to evaluate the heuristics efficiency and hence as a reasonable critical factor.

6.2.7 Ranking in the Early Stage

The rankings of the early stage are presented for all levels of complexity in Figure 13, beginning with 2 by 2 in the upper left corner and ending with 5 by 5 in the lower right. This order is retained for the forthcoming illustrations of all other stages.

Looking at the performance of the heuristics in absolute terms, this stage is by far the most dynamic, as one can observe in Figure 13. As mentioned earlier, it thus seems hardly adequate to transform the heuristics' dynamic performance under varying time limitation into a single ranking. On the contrary, at the very beginning of the heuristics' production systems all decisions predominantly rely on chance: *Random* and all strategic heuristics select a random decision as the initial choice. The other nonstrategic heuristics usually chose the start alternative of their procedures. Heuristics' procedures determine the starting point of observation (within the payoff matrix) by chance and without any pre-knowledge about the payoff matrix. Thus, an interpretation of the rankings is of limited use.

FIGURE 13 – RANKING IN THE EARLY STAGE



At about half of the time of the early stage, the first reasonable results of the heuristics' procedures can be obtained for interpretation. With increasing time limits in the second half of the early stage, the heuristics' performance results change successively, too.[55] However, it seems to be acceptable here to concentrate on the mean performance of the heuristics within the early stage. This concentration keeps the evaluation manageable concerning size and accuracy.

The *Optimist* heuristic, an early adopter, starts strongly with rank 1 in reducing alternatives, rank 4 in generating payoff and rank 3 in applied EIPs for a complexity level of 2 by 2. However, with

[55] Indeed, when concentrating on one particular time limit or span within the second half of the early stage, the results differ slightly from the computed mean. In that case, an own analysis for this very time limit is recommended.

growing level of complexity, it subsequently loses ranks in generating payoffs (rank 2 at 2 by 2 to rank 7 at 5 by 5).

Other than *Optimist*, the heuristic *L2* remains located in the lower left quadrant for all levels of complexity. The heuristic shows very good payoff generation (rank 2 to rank 3) and good reduction of alternatives as well as moderate EIP consumption (ranks 4 and 5). At complexity levels higher 2 by 2, *Pessimist* and *Naïve* have very good ranks for all goals, making them a good choice in such conditions.

Altruist shows a very good performance in reducing alternatives and EIP consumption (ranks 2 and 3) at all levels of complexity. However, the payoff generation is poor (rank 9). This performance lets it reaching the upper left quadrant.

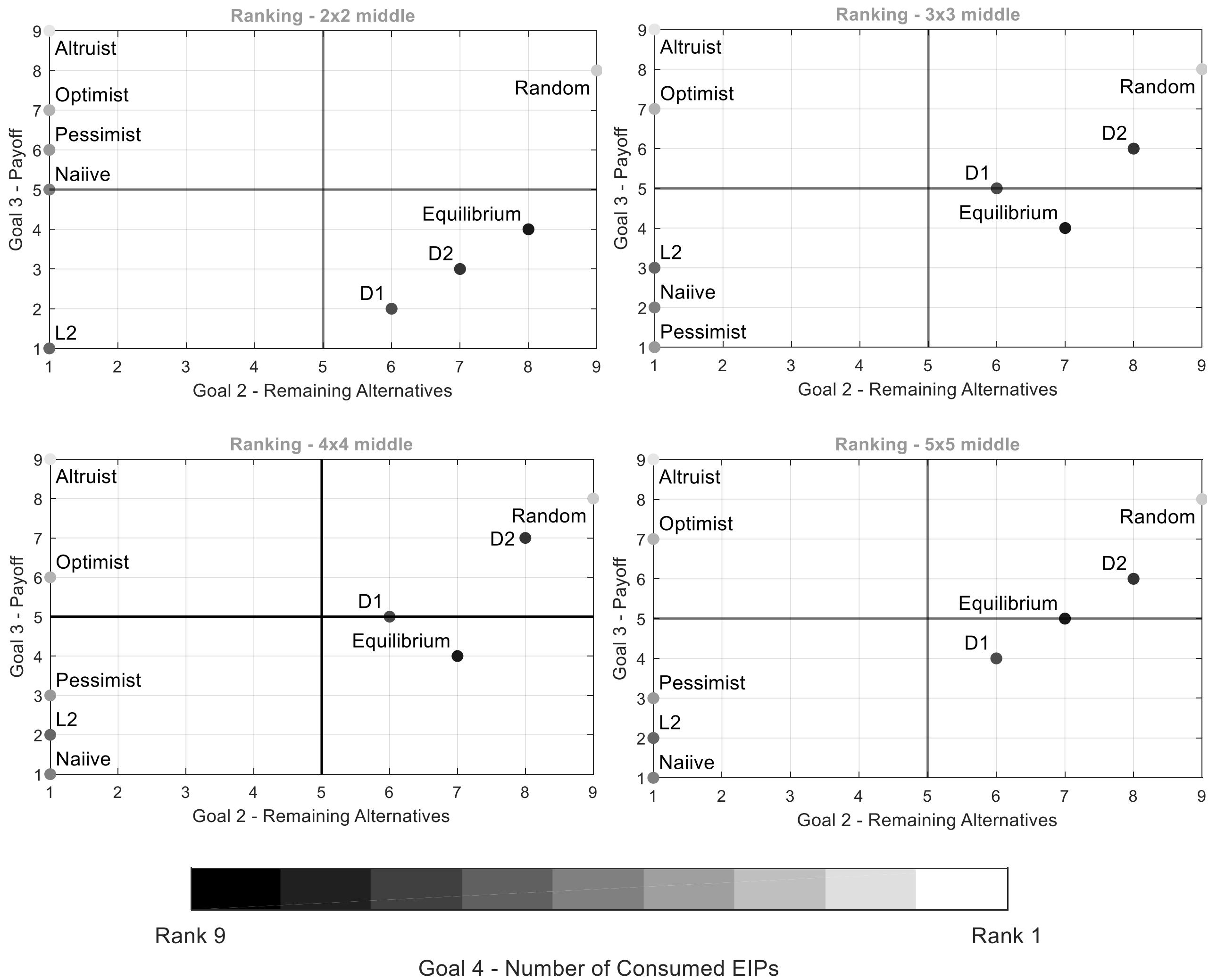
The strategic heuristics are frequently located around the payoff median in both of the right-sided quadrants with just marginal differences in ranks between them. In the fulfillment of the other two goals, the strategic heuristics show an even weaker performance occupying ranks well above the median (ranks 6 to 8 for Goal 2 and ranks 7 to 9 for Goal 4).

Summarizing the relative performances, the *Optimist* heuristic seems recommendable for the smallest level of complexity at an early time stage. For complexities of 3 by 3 and higher, *Naïve* and *Pessimist* top the ranking.

6.2.8 Ranking in the Middle Stage

Figure 14 depicts the rankings of the heuristics observed at this stage. In the middle stage, the results are quite similar to the early stage. Concerning Goal 4, no differences occur. *Random* performs best followed by the other nonstrategic heuristics and, with some distance, by the strategic heuristics. However, *L2* markedly improve its rankings regarding Goal 2 and Goal 3. The nonstrategic heuristics approached their maximum alternative reduction (with one alternative remaining), letting them share rank 1 together with *L2*. At complexity levels higher 2 by 2 *Naïve* and *Pessimist* improve their payoff generation, displacing the strategic heuristics from their ranks. They are thus to favor under such conditions, even though *L2* is close to their performance and at an advantage for smallest complexities.

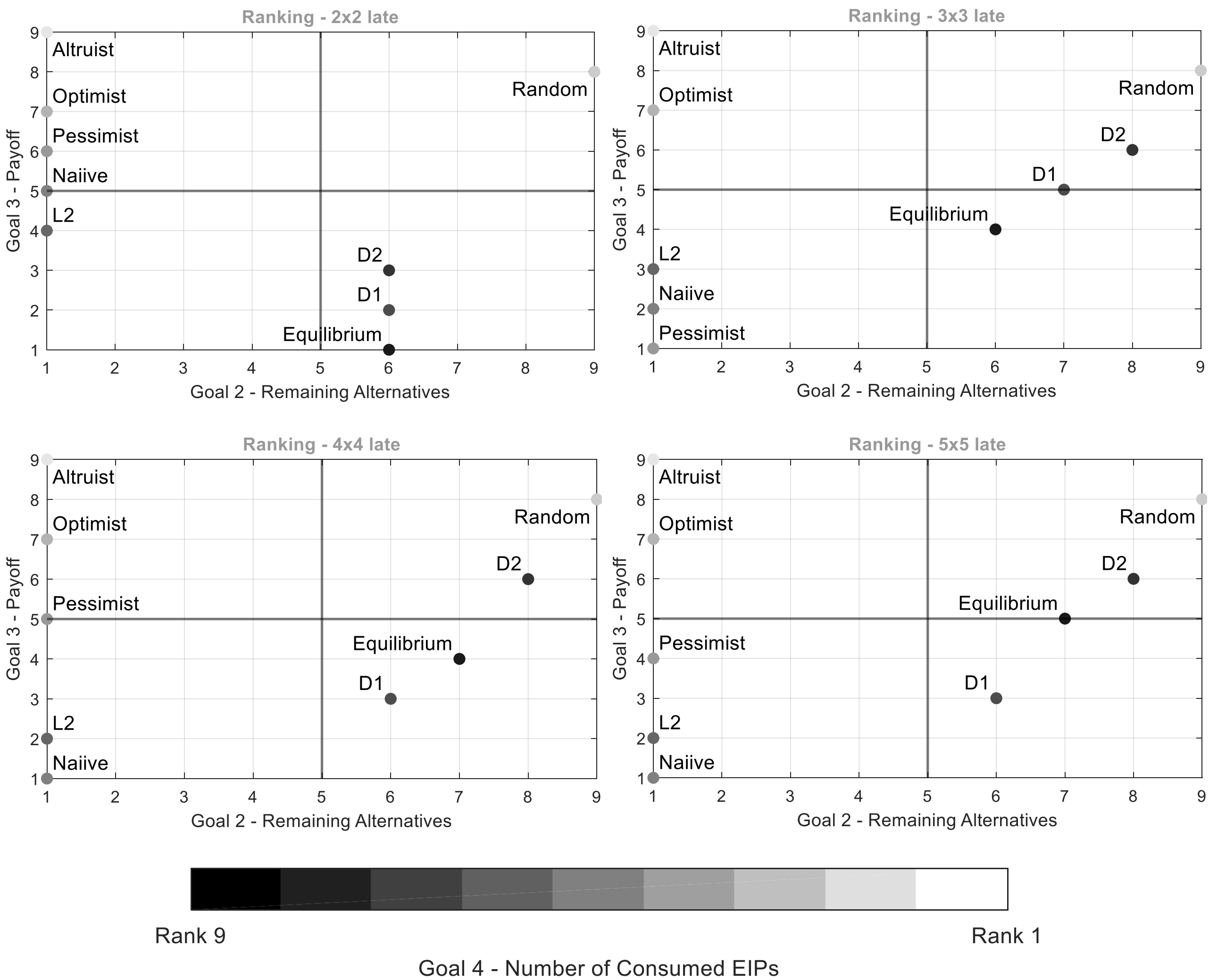
FIGURE 14 – RANKING IN THE MIDDLE STAGE



6.2.9 Ranking in the Late Stage

Figure 15 illustrates the rankings corresponding to the late stage. In this stage the rankings are again quite similar to the ones of the two stages analyzed before. Concerning Goal 4, no differences occur. As expected, the strategic heuristics still require the most EIPs for processing their production systems, resulting in ranks 7 to 9. For the other two goals *L2* throughout occupies best rankings in the lower left quadrant. The heuristic is accompanied by *Naïve* at complexity levels higher than 2 by 2 and by *Pessimist* at levels 3 by 3 and 5 by 5. In terms of payoff generation, the other strategic heuristics occupy at least one of the first four ranks throughout all levels of complexity. If the focus of optimization is on payoff generation only, *Equilibrium*, *Naïve* and *Pessimist* shall be selected, depending on the complexity level. *Naïve* is to favor for complexity levels of 4 by 4 and higher. However, of the strategic heuristics *L2* is recommended for all levels of complexity regarding alternative reduction and payoff generation.

FIGURE 15 – RANKING IN THE LATE STAGE

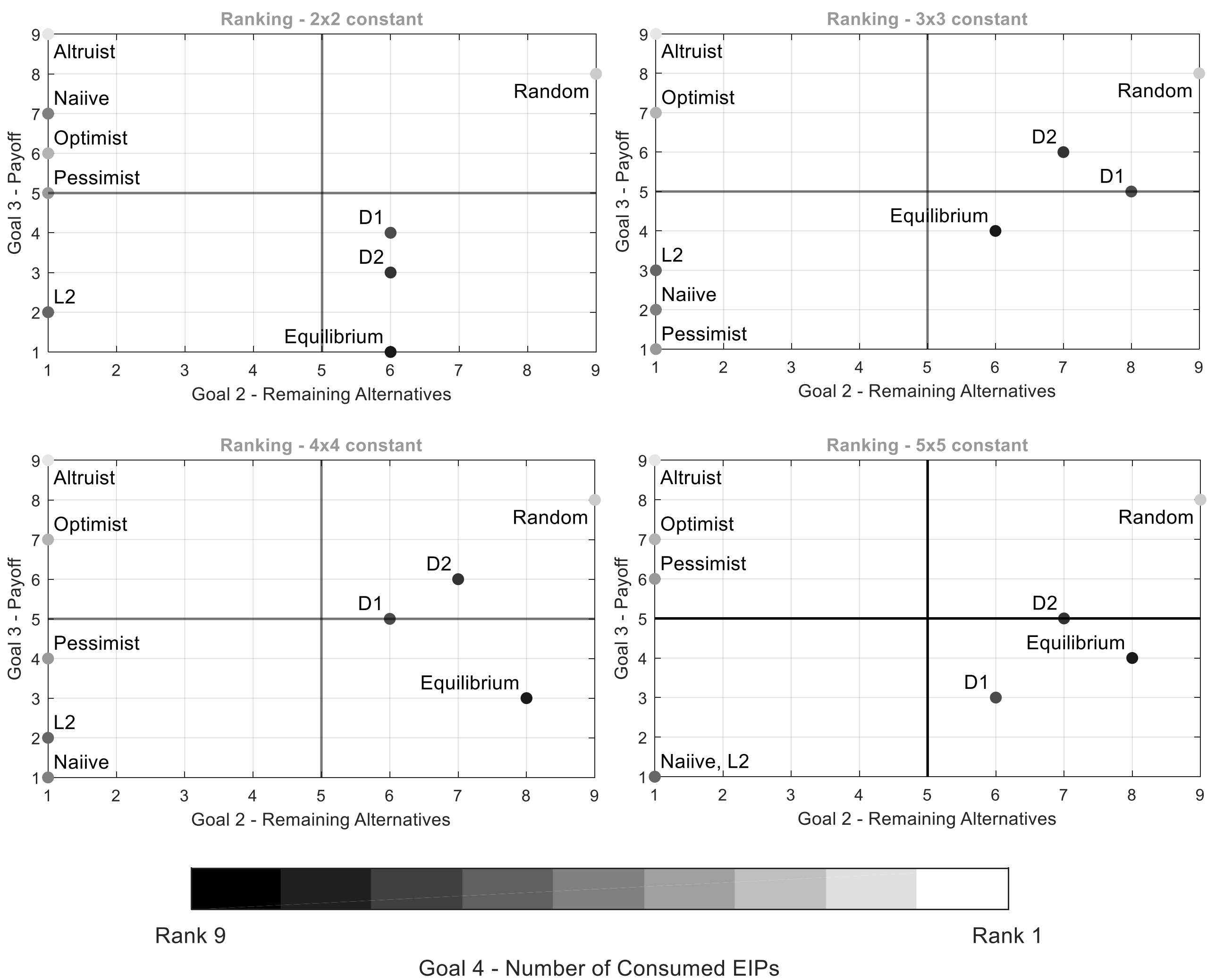


6.2.10 Ranking in the Constant Stage

At this stage, time limitation is no longer effective. The rankings of this stage thus illustrate the case of decision-making under no time limitation. Even though the level of complexity is varying, without any effective time limit, no time pressure exists. Thus, the heuristics complete their proceedings. The performance is hence expected to be similar to the one described in the literature (see Costa-Gomes et al. (2001, p. 1194)). Figure 16 illustrates the corresponding rankings.

After finishing task processing with their production systems, the nonstrategic heuristics (except Random) still remain ranked 1 in terms of 'alternative reduction'. *Naïve* once again shows a comparably strong payoff generation for complexity levels 3 by 3 and higher. The strategic heuristics do quite well on this goal, too. At least one of the three strategic heuristics *D1*, *D2* and *Equilibrium* can be found among the top four ranked heuristics. *L2* is throughout ranked among the top three. In combination with a shared first rank in 'alternative reduction', it is located in the lower left quadrant for all levels of complexity. However, the strategic heuristics traditionally share low ranks at Goal 4, where *Random* achieves without any competition rank 1. Overall, L2 and Naïve (at least for higher levels of complexity) are recommended to use for task processing within this stage of time limitation.

FIGURE 16 – RANKING IN THE CONSTANT STAGE



Interestingly, not the group of late adopters, but *L2*, *Naïve*, and *Pessimist* dominate the ranking concerning payoff generation for higher levels of complexity. This could not be expected from a normative point of view, where an equilibrium choice should optimize ones payoff. However, the different game types in combination with equally weighting the influence of each heuristic on the calculated payoff mean seems to be a disadvantage for the strategic heuristics. Additionally, when looking at the generated payoffs, the differences between values of top ranked heuristics and the ones of the strategic heuristics are comparably small. Considering these points, the results presented here are remarkable, but not contradictory to the equilibrium model of game theory and in good compliance to findings of Costa-Gomes et al. (2001, p. 1209).

6.2.11 Selected Trajectories

To better illustrate the evolution of the ranks, the charts of Figure 13 to Figure 16 are sorted in a new order. They form groups of same levels of complexities (Figure 17 to Figure 20). Each of the figures contains the ranking for one complexity level of all four stages, beginning with the early stage in the upper left corner and ending with the constant phase in the lower right. It follows a discussion of selected ranking trajectories. Appendix C provides a comprehensive overview of all trajectories in tabular form.

Nonstrategic heuristics are generally fast to identify an alternative to choose which is similar to effectively reducing the set of remaining alternatives to an absolute minimum ($= 1$ alternative). They thus occupy the first ranks right from the beginning with the heuristic *Optimist* dominating the early time stage. Due to their production systems, which require only a small share of EIPs of what the strategic heuristics (except *L2*) need, the nonstrategic heuristics also dominate the ranks of Goal 4. In terms of generating payoff they in most cases perform worse than the strategic heuristics. Frequent exceptions are *Pessimist* and *Naïve* which occupy one of the top five ranks.

Among the set of heuristics, *L2* occupies a special position. Regardless of time limitation and level of complexity this heuristic always performs very well in terms of Goal 2 and Goal3. It reduces the set of alternatives as fast as most of nonstrategic heuristics and its payoff generation is frequently among the highest values. This heuristic is therefore recommended for all time pressure conditions.

The development of ranks for the other three strategic heuristics is different from the one of *L2*. For Goal 2 and Goal 4, no changes in the ranks occur. In Figure 17 to Figure 20 the strategic heuristics are thus consequently located in the two right sided quadrants with darker shades of gray. The three strategic heuristics are incapable of completely catching up on performance versus nonstrategic heuristics. That is also partly true for Goal 3. However, as available time increases, they can increase their payoff generations to improve their ranks by at least one position. This behavior illustrates the strategic heuristics' characteristic as late adopters.

Despite such differences between the heuristics, two characteristics have been identified which all of them share. First, the heuristics' respective rank evolution shows a discrepancy with respect to time between Goal 2 and Goal 3. The ranks of Goal 3 change before the ones of Goal 2. This outcome surprises since in case of the strategic heuristics the performance for Goal 3 is expected to be directly linked to the performance for Goal 2. Thus, this point is to be discussed in more detail in the following by examining whether this phenomenon coincides with absolute numbers. Revisiting Figure 7 reveals a declining number of alternatives for the strategic heuristics in the middle stage. Here, the nonstrategic heuristics have reached their constant phase already. The point of intersection, if it exists at all, between the performance graphs of strategic and nonstrategic heuristics is located in the later phase of the middle stage. Beginning with the moment of intersection, the strategic heuristics perform better than the nonstrategics. This point in time is decisive for the rankings since the mean performance of a particular stage decides on the rankings at this stage. The later this moment occurs at this stage, the smaller the time span where strategic heuristics perform better than the nonstrategics. Thus, the advantageous performance of the strategic heuristics does not become effective before the late stage concerning improved ranks.

On the other hand, when examining the middle stage of Goal 3 in Figure 9 one observes a tendentially increasing payoff for the strategic heuristics. Nonstrategic heuristics do not behave homogenously here. With *Random*, *Altruist*, *Optimist*, and *Pessimist*, four out of five heuristics already have lower ranks than the strategic heuristics which later even continue to decline. *Naïve* is the prominent exception which shares the top rank with *L2* from the early to the constant stage for all but one level of complexity (i.e., 2 by 2). So, the strategic heuristics behave as initially expected with a drop in remaining alternatives and a simultaneous growth in the payoff. The influence of the other heuristics' performance is still notable. With the reduction of dominated alternatives in the middle stage, the strategic heuristics' performance is better than most of the nonstrategic heuristics (except *Naïve*). Figure 10 clarifies this relation for a complexity level of 2 by 2. The influence of *Random* on the payoff of the set of heuristics is neglected in this picture. The nonstrategic heuristics' performance for Goal 3 deteriorates in the middle stage successively with the strategic heuristics reducing their set of alternatives (and thus increasing their payoff).

And second, the ranks do not change much with the time stages. The respective ranks the heuristics occupy at the early stages are frequently similar to the ones they occupy at the other time stages. Some changes occur within the group of strategic heuristics regarding Goal 2 and Goal 3. Some changes occur within the first five ranks regarding Goal 3 over time between *Naïve*, *Pessimist* and the strategic heuristics. Despite that the ranks are stable. Again, the diverse time sensitivity of the heuristics' production systems can be accounted for this. The incapability of the late adopters to frequently produce performance values at Goal 2 and Goal 3 better than the ones of the others is a second cause. A different compilation of heuristics' weight within the calculation of the payoff generation might have changed the results. However, this interrelation is not further studied here.

FIGURE 17 – RANKING AT A COMPLEXITY LEVEL OF 2 BY 2

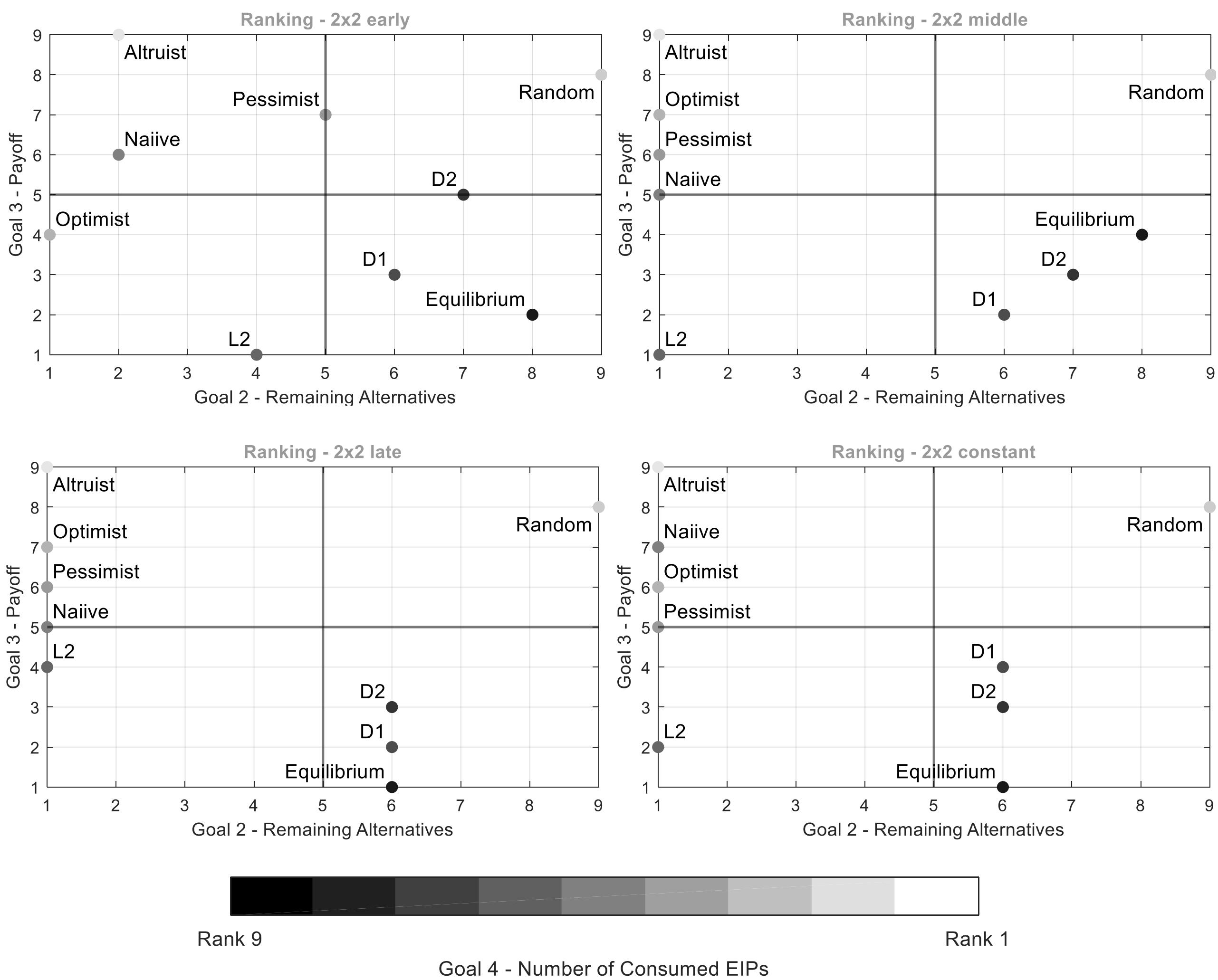


FIGURE 18 – RANKING AT A COMPLEXITY LEVEL OF 3 BY 3

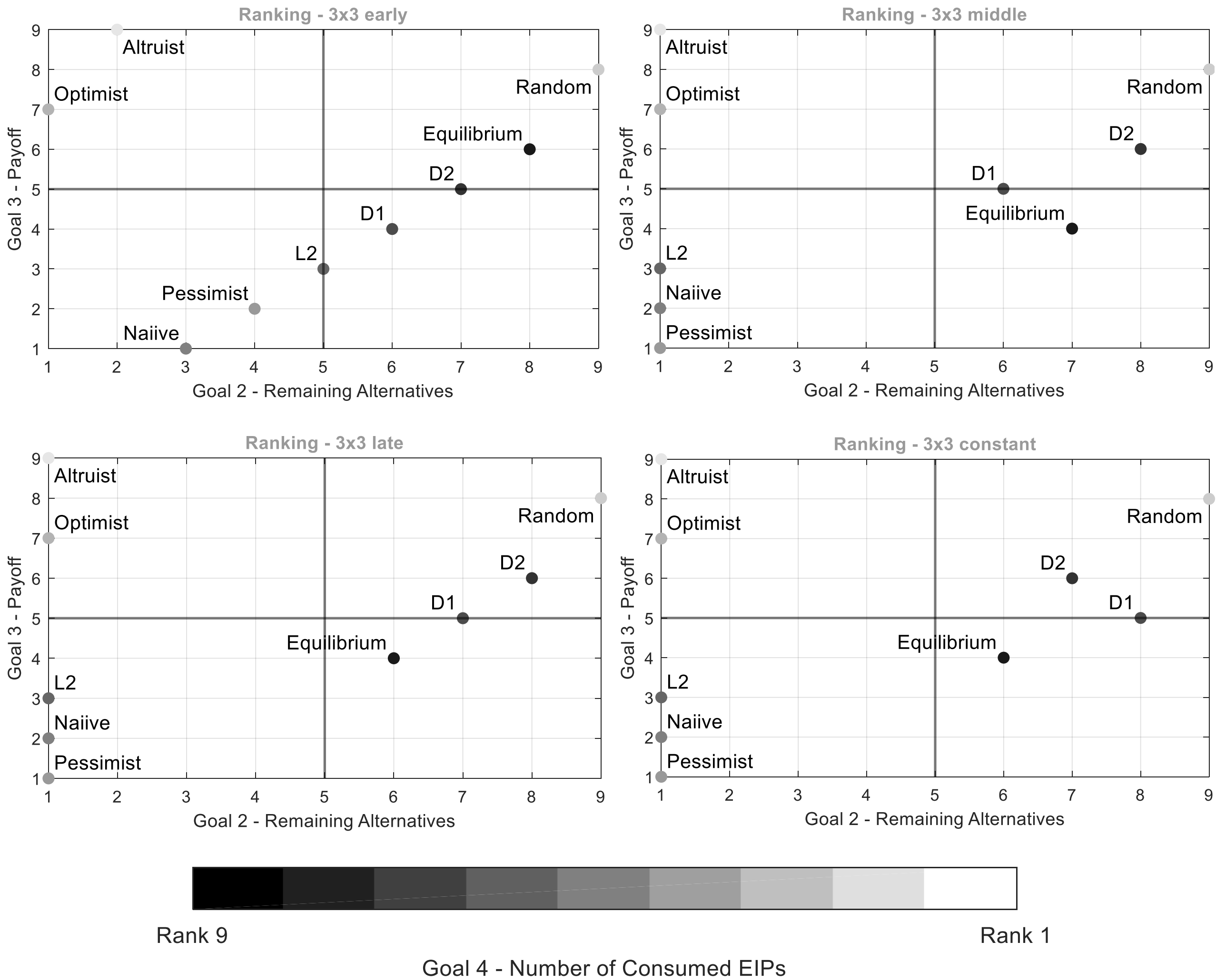
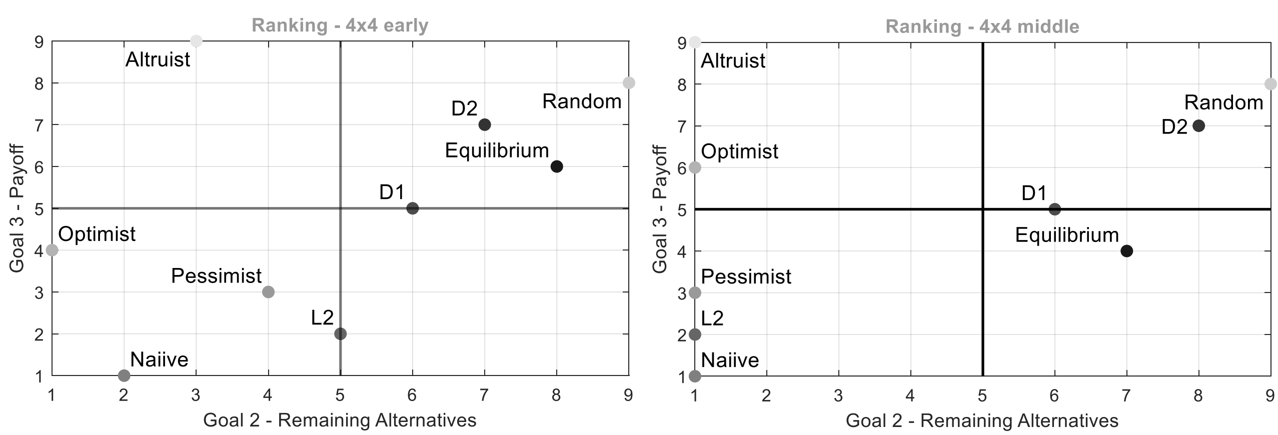


FIGURE 19 – RANKING AT A COMPLEXITY LEVEL OF 4 BY 4



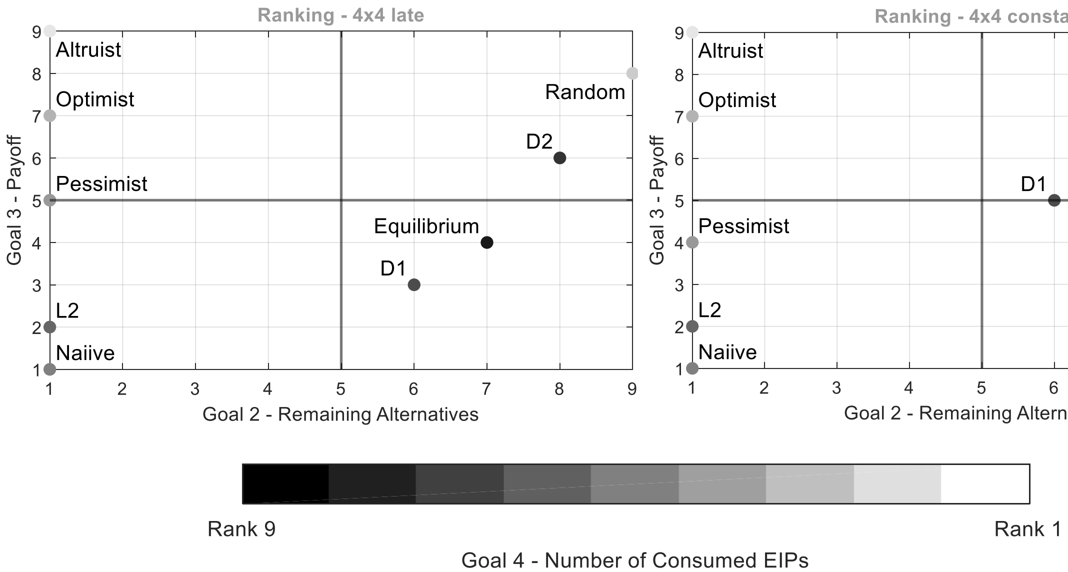
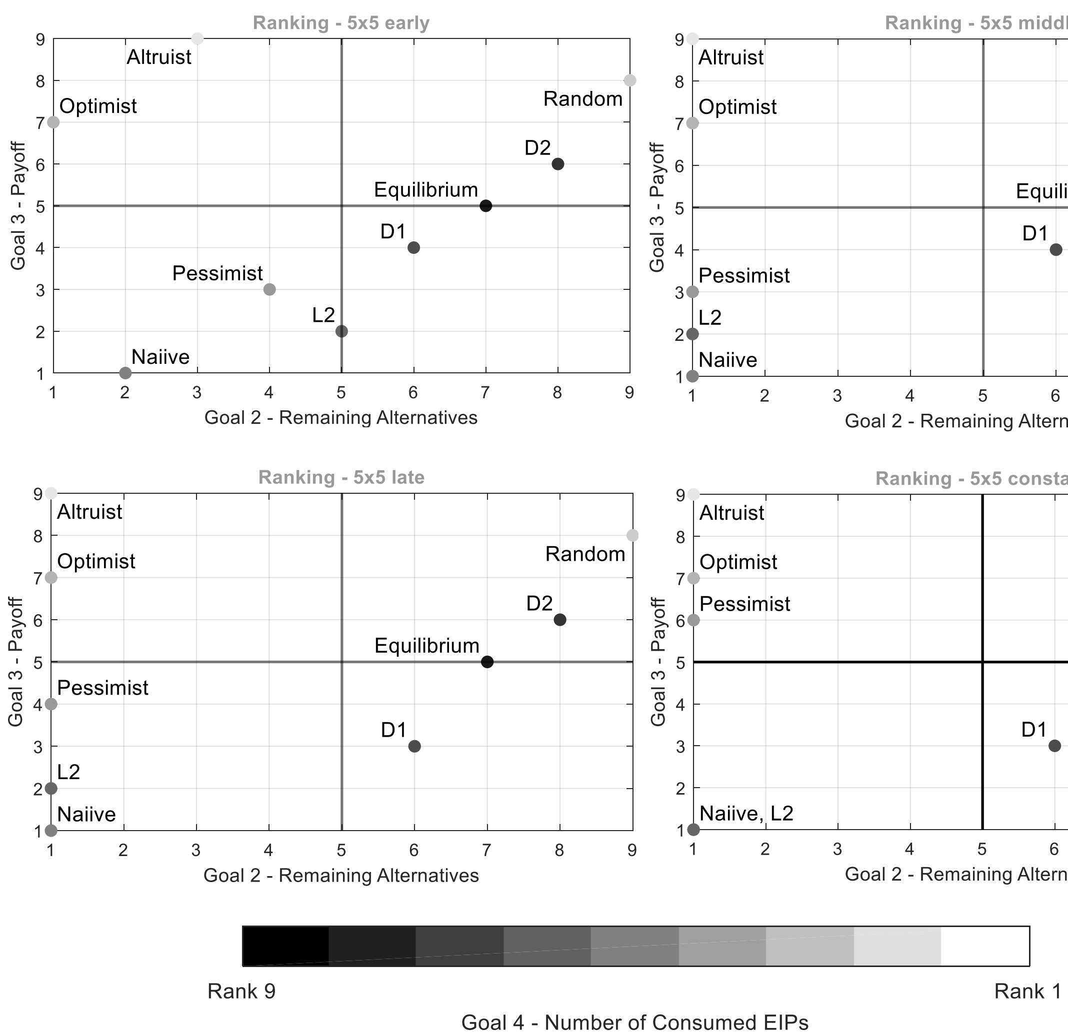


FIGURE 20 – RANKING AT A COMPLEXITY LEVEL OF 5 BY 5



In this chapter the simulation data has been analyzed regarding the heuristics' performances under varying time pressure conditions. Three goals reflect effectivity as well as efficiency of the choice and the process of choice generation. The heuristics' performances were studied in absolute terms and compared to each other in a ranking. The heuristics show remarkable differences in their ability to reduce the set of alternatives, to generate payoff and to process the decision task with as few EIPs as possible. While the nonstrategic heuristics reduce the set of alternatives the fastest, the strategic heuristics achieve good payoff values. A special position occupies the heuristic *L2* which combines the advantages of both groups. Its features, as well as the ones of other behavioral patterns that proved successful under time pressure conditions, are the subject of the following chapter.

7 Successful Heuristic Patterns

This chapter deals with heuristic patterns that are regarded as successful under varying time pressure conditions. The discussion aims at describing the patterns' characteristics and the conditions under which they are successful in the light of the preparation time model and its predictions (c.f. Chapter 3). The discussion is based on the results from Chapter 6, where the performance of selected heuristics over varying time limitation and complexity was determined. The set of heuristics is regarded as commonly used in the human decision-making process. The heuristics have shown a significant sensitivity to both complexity and time limitation in strategic normal-form games. From an application perspective, it therefore seems appropriate and reasonable to design the time pressure based on these two factors.

An evaluation framework to measure the heuristics' performance in the decision process was set up. This concept offers the possibility of comparing the performance of the heuristics regarding their effectiveness and efficiency. Thus, one of this treatise's goals could be met: identifying successful behavioral patterns under time pressure conditions as benchmarks to compare with other heuristic' performances and human decision-making.[56]

Some simplifications define the following overall consideration. The heuristics, as well as the games, represent a small selection compared to the richness of variations present in literature. All findings of this research and their validation are hence limited to this set of games and heuristics. Those cannot claim universality. However, the selection of games and heuristics is based on their frequentness and generality. Considering these limitations, the applicability of the concept of preparation time in game theory can be assumed.

Another simplification deals with the evaluation of the heuristics' performance. As can be seen in Table 14, the set of games applied in this study shows a variety of form and level of reasoning. Still, for presentational reasons the performance is calculated as the mean over all games per level of complexity. That way a manageable overview is generated out of the rich dataset from the simulation.

For the same reasons, the examination of the influence of time limit is reduced to four stages, with the performance means of each stage determining the ranks. With 466 time-limit measurements many more combinations exist that await a more detailed study. For example, the majority

[56] For results see Appendix C. A digital version of the comprehensive simulation data set can be obtained on request.

of heuristics already finishes procedures and reaches the eventless constant phase in the early stage. Examining time limits of this stage is thus likely to reveal more interesting results and relations. The performance of the heuristics for one particular game type or level of reasoning and for a particular time limit thus might be relevant in further studies.[57]

As already mentioned in Section 6.1, the heuristics' performances for Goal 3 result from playing against each other. Here, a heuristic is matched with each heuristic from the set, including itself. The sum then is uniformly weighted by the number of heuristics. In doing so, one implicitly assumes that the set of heuristics is evenly represented in the simulated population of decision-makers and thus in the performance evaluation. This assumption might be a limitation of reality, strongly influencing the heuristics' outcome for Goal 3.

On the other hand, the experimentally derived probability distribution of heuristics as a solution concept for decision-making is still a controversial and lively part of scientific discussions.[58] The assumption made here is thus to be seen as an adequate first assessment, with too few solid arguments presented in literature standing against it. The need for further experimental studies, applying the latest technical developments in process tracing, combined with an adequate data analysis concept is quite apparent here.

Considering those simplifications as reasonable and adequate, the results from the simulation can now be discussed. The focus lies on the heuristics' performance under time pressure and its relation to the implications of the preparation time concept, introduced in Chapter 3. For heuristic-based decision-making under time pressure, this concept predicts that choices deviate from those made without time pressure. This effect can be related to the heuristics' ability to minimize the set of alternatives and to generate payoff. It thus can be directly linked to the second and the third performance goal. The preparation time concept further implies that the number of EIPs required for executing a heuristic's production system decreases under time pressure conditions. This reduction affects the heuristics' efficiency and thus links to the fourth performance goal. Both implications of the preparation time concept are addressed alongside the discussion of the influence

[57] Performance data will be provided by the author on request, enabling specific data aggregation by choice for further research projects.

[58] The studies of Camerer et al. (2004) and Costa-Gomes and Crawford (2006), for example, offer two fundamentally different approaches to examine decision-making by heuristics in a strategic context. Critical studies concerning this topic are among others Bettman (1979) and Hahn et al. (2010).

of time pressure on the heuristics' performances. The impact of time pressure is at this point discussed by evaluating the impact of its components time limitation and complexity.

At first, concentrating on the time sensitivity of the heuristics performance, they show a similar adaptation behavior with slight variations, supporting the predictions from the preparation time concept: in the heuristics' performance graphs a transition phase and a constant phase can be determined. In the transition phase, the time limit is effective in the way that the heuristics are not able to finish their procedures other than with a random decision. At the point of saturation the transition phase changes over to the constant phase. In the constant phase, time pressure is no longer effective – the number of applied EIPs, as well as the number of remaining alternatives, is not changing anymore, even though the time limit increases. However, because of the interactivity, the payoff for each heuristic differs with time until the last heuristic reaches its constant phase. As expected from the preparation time model, the concrete point of saturation is a characteristic of the heuristics. It directly depends on the level of complexity.

Despite similarities in the shape of the adaptation process, the heuristics can be categorized from a performance-based point of view concerning different aspects. The basic categorization of the set follows Costa-Gomes et al. (2001), who defined *Random*, *Altruist*, *Optimist*, *Pessimist*, and *Naïve* as nonstrategic heuristics and *L2*, *D1*, *D2* and *Equilibrium* as strategic heuristics. All discrepancies between the heuristics' performances are based on differences in their production systems and the systems' complexity. The differences in the production systems' complexity are clarified and operationalized by simulation. The nonstrategic heuristics apply simple information search and gathering steps, not focussing on the opponents' set of choices. Compared to the strategic heuristics this simplicity pays off with respect to the amount of EIPs necessary to finish the procedure and a faster reduction of the number of alternatives.

The nonstrategic heuristics *Optimist*, *Pessimist* and *Naïve* even tend to be almost invariant to time limitation regarding their effectiveness. They show a distinct ability to develop a reliable choice in a short time. This decision speed leads to a comparatively good performance in generating payoff under comparatively short time limits. Although *Random* is even faster, its choice is not based on any information. The use of even little information seems advantageous here, regarding payoff generation and thus effectiveness. *Altruist* is comparably fast, but the chosen alternatives lead to poor payoffs.

The complexity of the heuristics' production systems is, in turn, a strong argument to explain why the strategic heuristics need so much more EIPs. As indicated on several other occasions, *L2* again is in an intermediate position: opponent's payoff information is used to form a decision. *L2* thus can be defined as strategic (similar to Costa-Gomes et al. (2001, p. 1195)). However, it behaves nonstrategic in its demand for EIPs, making it more efficient than the other strategic heuristics.

Goal 4 (Applied EIPs) illustrates very well the time aspect of the heuristics' procedures (see Figure 12). Here, one can identify two groups with different time requirements. The nonstrategic heuristics, together with *L2*, can be labeled as 'early adopters'. The strategic heuristics *D1*, *D2*, and *Equilibrium* follow with a marked time difference. They are hence labeled as 'late adopters'. Within this group, *D2* again needs significantly more time – regardless of the game task's complexity. The game tasks applied in this paper require a level of reasoning of two at the maximum[59], meaning that the normally optimal strategy might require two rounds of iterated eliminating of dominated alternatives[60]. The mean level of reasoning of all game tasks is lower than two. However, the solution concept of *D2* is rigid in the sense that regardless of the particular game type always two rounds of iterated elimination of dominated alternatives are processed. It is thus performing worse than other heuristics concerning effectiveness and efficiency.[61]

The complexity of game tasks has a substantial impact here, with the early adopters increasing their time demand almost quadratic with the level and linear with the growing number of elements in a payoff matrix. The late adopters show a noticeably bigger factor at this point. This circumstance hints to the stronger dependency on the context variables 'game type' and 'payoff information' compared to the early adopter heuristics. Whereas the early adopters generally seem to be game type invariant, the late adopters show specific sensitivity. The reasons for this can be found again in the properties of the production systems: nonstrategic heuristics check each cell of the payoff matrix with the same frequency (usually once). *L2* again is behaving as a nonstrategic heuristic here, even though it is checking a certain amount of cells more than once. The late adopter heuristics do not show that congruence, with the frequency principally depending on the concrete elimination process. It is not surprising that the context variable 'payoff information'

[59] A maximum level of reasoning of two for game types is sufficient for this study. Section 5.2 provided an overview of applied selection criteria for game types.
[60] Section 5.2 provided a detailed discussion of this game-theoretical solution concept.
[61] Because of its processing time, *D2* is selected as reference heuristic in the analysis section of this study, framing the scene with determining the examined time stages.

that links to the game type also has a remarkable influence on the number of EIPs heuristics require during their procedure. Except for *Random*, the complete set of heuristics shows sensitivity here. The number of remaining alternatives in the constant phase is a good indicator for the influence. This number is changing with the degree of complexity, even though it should be the same or at least directly related to this level. Time dependence is noticeable yet until the constant phase is reached.

Concentrating on Goal 3 (Generated payoff), the influence of task and context variables is also clearly observable. Since the payoff is determined for all heuristics by playing against each other, especially at early stages where several heuristics rely on random decision-making, the effective payoff is influenced by chance. The impact of the element of chance decreases with increasing processing time of the heuristics – primarily when the late adopters can develop decisions based on higher levels of reasoning. This behavior is in reasonable compliance with the implications of the preparation time concept.

Surprisingly, the decisions proposed by *Naïve* and *L2* are even more effective than the ones of the late adopters, considering the mean payoff acquired from playing all game types. Those two heuristics perform very well for all goals, with *L2* generally being advantageous. Since *Naïve* identifies the alternative optimally responding [62] to a random choice and *L2* identifies the best alternative to respond on a *Naïve* choice, *L2* reaches the second level of reasoning according to the CH-model from Camerer et al. (2004).[63] Effectively, *L2* does not employ the normative game theoretic solution concept of eliminating dominated alternatives that is based on the assumption of rational opponents. Using *L2*, one simply assumes that the opponent is rationally bounded[64], at least nonstrategic.

With the behavioral pattern described above, *L2* is a remarkable exception from the group of strategic heuristics. It does not rely on the comparatively EIP-intensive search for dominated strategies. That significantly reduces its EIP demand towards scales that are comparable to the ones of nonstrategic heuristics. Since *L2* has a strategic character, it regularly performs better than the nonstrategic heuristics regarding payoff generating. Its ability to reduce alternatives in compara-

---

[62], *Naïve* is thus of a first-level-of-reasoning type, following the classification system of Camerer (2003).
[63] Compare the explanations in footnote 6, p. 13.
[64] The game theoretic term ‘bounded rationality’ refers to a player's limited ability to generate a strategic decision. Also, see the description in the glossary (0).

tively less time is in turn similar to the nonstrategic and thus definitely advantageous to the strategic heuristics.

Whether the success of *L2* within this set of heuristics is heavily influenced by the played games, by the structure of the set of heuristic or other variables is difficult to determine and necessary to be investigated in further studies, explicitly not limited to simulation. Costa-Gomes et al. (2001, p. 1219, footnote 58) report similar tendencies for this heuristic, although the strategic heuristics in their study perform better compared to the results in the constant phase here. From the current point of view, *L2* is recommendable for almost the whole parameter set studied here, even though some uncertainty remains.

With this, one can finally conclude the time pressure sensitivity of heuristics' performances regarding effectiveness and efficiency. Table 17 provides an overview of the influence of time limitation and complexity. Some heuristics naturally perform more robust under a decreasing time limit than others with respect to effectiveness. This fact does not necessarily depend on the adopter character of the heuristic. However, this characteristic can be a supportive argument. Especially the group of late adopters faces decreasing effectiveness under increasing time limitation. A growing complexity even scales this effect. The reason probably can be found in the complexity of their production systems and thus in their comparatively intense need for EIPs. Another strategic heuristic, *L2,* performs surprisingly well under time pressure and is thus recommendable from the given set of heuristics.

Concluding the second part of this treatise, the findings can now be summarized. Decision-makers' behavior in tasks under varying time pressure conditions was simulated through a set of common heuristics. Therefore, the heuristics' production systems, developed in Part I, were applied to the strategic decision-making tasks. Time pressure was modeled as a function of time limitation and task complexity. The performance of the heuristics, measured in terms of effectiveness and efficiency, was determined.

It was shown that the approach produces results in line with Costa-Gomes et al. (2001) and others for the extreme case of no time limitation. The simulation generated reasonable outcomes for the heuristics' performances, strongly depending on the production system's progression under time limitation. The heuristics presented significant differences in their general performance. Surprisingly, none of the established sophisticated heuristics tops the overall ranks, but one which combines non-sophistic and sophistic elements in its procedure. Even in task environments of high complexity, where sophisticated heuristics are assumed to surpass the nonstrategic ones markedly, some others performed comparatively strong, too. Hereby, especially under time pressure, the sophisticated heuristics are not unconditionally recommended from a normative point of view. Altogether, the results back up the presumptions of the preparation time concept and further support the applicability of both model and method. However, the results await their experimental validation. They thus find consideration in the hypotheses formulation in the following of part this treatise. In Part III, a decision-making experiment is conducted to study the predictions of the preparation time model under more realistic conditions.

Part III: Time Pressure in Experimental Normal-Form Games

The third part of this treatise provides the description of an online decision-making experiment. Main objective is to evaluate the predictions from the preparation time model, elaborated in the two preceding parts of this treatise. The predictions are formulated as hypotheses which are tested in the experiment. In this context, behavioral patterns in decision-making processes when executing normal-form games under various time pressure conditions are identified. It is further evaluated on how time pressure, as a function of time limitation and task complexity, influences effectiveness and efficiency of cognitive processes.

Mouse tracking is used to obtain process data. For this purpose, a detailed discussion of mouse tracking along with other relevant process tracing techniques (PTMs) is given in Chapter 8. Improvements in mouse tracking are described that result from comparison of relevant PTMs. In the same chapter a metric is developed that allows for interpreting behavioral data, observed by mouse tracking, as cognitive operations (EIPs).

The following three chapters are dedicated to the online experiment. Method and design are explained in Chapter 9. This chapter also provides some general issues regarding experiment and data analysis. Findings are presented in the following Chapter 10 along with results of hypothesis testing and classifying behavior.

In the subsequent Chapter 11 results are discussed, concentrating on the experimental objectives, hypotheses, and classifying behavior. It also provides a reflection on identified cognition types in the light of the earlier introduced model of preparation time.

8 Tracing the Cognitive Process

In this chapter, techniques are introduced in brief that allow inferring on the cognitive process from observed behavior and discussing mouse tracking in more detail. Mouse tracking is compared to other relevant techniques in substantial categories. The comparison shall reveal potential improvements for mouse tracking regarding implementable functionalities. Examined categories refer to the purpose of this experiment. On this technological base, a metric is developed that enables interpreting the observed behavior as a cognitive process, applying the concept of EIPs. This metric is further used for analyzing the experimentally derived behavior data.

8.1 Purpose of Process Tracing Methods

The Process Tracing Methods (PTMs) comprise numerous techniques designed and developed to gather and provide appropriate data to examine cognitive processes which trigger human behavior. Following early attempts of Ford et al. (1989) and Woods (1993) to structure this comparatively young field of research, Kühberger et al. (2011, p. 3) recently offered a comprehensive introduction to the field of PTMs. The authors include a compilation of numerous studies, presenting experiences and implementations of several techniques and give a convincing categorization of the methods discussed. In this paper, their categorization is borrowed. The methods are classified regarding their intention of helping to examine:

- information gathering and acquisition,
- information evaluation and integration, and
- physiological and neurological aspects

of a cognitive process.

PTMs partly fulfill more than one purpose, making their classification less strict. That in turn principally broadens the range of applications for a single method. On the other hand, the number of potential techniques for a particular application increases with this, too. Choosing one out of the pool of potentials is thus inevitable, as long as the research concept does not focus on a specific method from the beginning.[65]

In general, PTM-related literature frequently addresses two significant points of concerns. Consequently, one needs to consider these potential flaws prior to selecting a method in the context of a research project:

1. One major issue is to assess the potential influence of the technique on the cognitive process. That implies both, the accuracy of data regarding its ability to reflect on the underlying cognitive

[65] Even in that case, comparing relevant techniques about performance and fitting could generate stimuli for the selected one and its experimental application. That in turn potentially leads to a further development of the method itself. The development has often been influenced by technical evolution, generating new instruments which are frequently adopted by several process-tracing methods. That triggered similar changes in design and development of the techniques. The availability of personal computers in combination with growing data storages is a prominent example. These technologies strongly influence experimental designs with regard to gathering, storing and analyzing process data. They offer new opportunities in data richness and precision.

process as well as minimizing distortions as reported in Russo et al. (1989). The results of this reflections especially influence the experimental design, which needs to consider possibilities to identify quality and quantity of such "sources of invalidity" (Russo et al. 1989, p. 760). This critical point is part of the related discussion of relevant process tracing methods within this study.

2. Kühberger et al. (2011) mention the general flaw in the methods' working mechanisms which equally limits their applicability: PTMs are bound to information given within an examination. They cannot supervise brain-internal data such as knowledge and experience. Its use and influence largely remain unclear (pp. 12 f.).[66] Kaplan et al. (1993, p. 265) dedicate an entire study to this issue in the context of judging people from specific presented characteristics under time pressure. They found evidence for the existence of brain-internal representations of information.

The challenge to determine and evaluate the brain-internal data and its influence on decision-making in a particular task generally exists in every human decision-making study. A promising way is to formulate decision situations knowledge-neutral, such that they most likely do not belong to the (many) experience horizons of people.[67] Thus, the absence of that kind of experience could be primarily assumed. Also, the ability to deal with that kind of problems by methodology would remain and would not be affected. In case of the current research, this approach is an option, as will be shown later when decision tasks are defined. The uncontrollable internal data are hence of minor importance in the following selection process.

Aspects said to be relevant when selecting a technique can be formulated, considering these points of criticism. These aspects build a framework that is employed in a second step to evaluate possible methods from a comparative point of view.

The goal of such a selection is to choose a method that serves the research objectives well for:

1. generating proper data to examine relevant characteristics of the underlying model[68],
2. offering a sufficient cost-benefit perspective in an application,
3. being accepted as a method in scientific application context or at least offering strong arguments for its acceptance, and

[66] It needs to be added here, that especially neurologically-based process tracing methods mainly evolve towards visualizing and analyzing brain activity. This method can be used to determine whether an examined activity is connected to an external or an internal stimulus.
[67] Consider that the neutral formulation might not be possible in every case, depending on the research concept.
[68] This point also comprises proper means to minimize distortions of the PTM on the behavior according to the first point of criticism mentioned above. It is further necessary to note that no technique is ready to be applied from scratch. In every case, researchers need to further adjust it regarding their particular approach.

4. providing sufficiently documented application experiences and reference experiments in the literature that prove an adequate level of elaborateness and enables comparison of results.

Especially the second aspect is a decisive factor when research is bound to a given pool of (monetary and technical) resources. This scheme of aspects forms the basis of the following evaluation and comparison of relevant methods.

8.2 Discussion of Relevant PTM

This evaluation shall identify potential improvements for mouse tracking from other PTM. The author focuses on adequate functionalities he can implement in mouse tracking. Mouse tracking is compared with three popular techniques that nominally have the potential to substitute it for the given experimental purpose: eye tracking, verbal protocols and functional Magnetic Resonance Imaging (fMRI).[69] Since researchers partly develop methods for different application areas, their performances are assumed to differ, given the particular intention of the present research. A description of the techniques identified regarding general functionality, working mechanism as well as traditional fields of application is subject of the following four subsections. The fifth presents the results of the comparison.

8.2.1 Information Board and Mouse Tracking

Purpose

The process tracing method ‘Information Board’ belongs to the class of information gathering and processing tools that trace subjects’ behavior during the act of decision-making. Payne initially presents it in an information search context. He also coins the technical term of this method (Payne 1976, p. 370).

Functionality and Points of criticism

In his work, Payne describes a method to examine subjects' information gathering procedures in a predefined experimental information setup. This setup contains a board on which information is

[69] The PTMs are selected because they apply to normal-form games and are capable of tracing relevant information about the decision-making process. Moreover, they are used in relevant studies, showing their potential in decision-making research. Other PTMs, studied in a preliminary consideration to this treatise, missed at least one of the two criteria and are hence omitted. A good overview and thorough introduction to the latest process tracing methods in use, as well as advantages and limitations, offers the compilation of Schulte-Mecklenbeck et al. (2011b).

presented, hidden in different envelopes. A subject then is asked to fulfill specific tasks, based on the processing of the information stated in those envelopes. The information gathering in this context is restricted to a single envelope that is allowed to be opened and watched at a time. This way Payne recorded time and order the information is monitored and draw conclusions about the underlying information processing and thus on the cognitive process itself (Payne 1976, pp. 370 ff.). He later improved his approach with a digital version he named Mouselab. Based on the same procedure, the software solution employed a graphical user interface (GUI) corresponding with a PC as information board and a mouse as input device. Mouse movement proves to be an accurate mean for measuring information processing, comparable to the functionality of the eye tracker technique (Willemsen and Johnson 2011, p. 38). With some modification this software is made internet-ready, firming under the current title *MouselabWeb*. Other software solutions, principally based on the same idea, are *MouseTracker* (Freeman and Ambady 2010) and with the generally broader goals to support behavioral experiment *z-Tree* (Fischbacher 2007).

This technique principally faces two points of criticism as noted in Kühberger et al. (2011, p. 4). First, text form is the dominant presentation style of the experimental task. This form is said to limit the way of information gathering. Second, the information given in the experiment is selected and pre-structured, maybe unintendedly leading and influencing the cognitive process. Both points of criticism are substantial and need to be considered by the experimenter. The technical evolution of the information board method offers the possibility of presenting other content forms than text. However, arguing with an average person's horizon of experience, a text is a common form of information, which thus finds its utilization in daily life decision-making. Especially digits also have its representation in the brain, as a study of Damarla and Just (2013, p. 2624) revealed. Text alone as input information hence seems appropriate.

Nutt (1998) studies the framing effect in strategic environments. This effect describes the impact of information pre-structuring on decision-making.[70] Following the evidence presented, the framing effect needs to be considered. Hence, task and content shall have a neutral expression to minimize the effect. This indication is especially relevant for the information structuring and implicates the presentation, too.

[70] For a detailed description of the framing effect in decision-making see Tversky and Kahneman (1981) and Nutt (1998) or in the present treatise Section 9.3.3.4.

In conclusion, the information board method shows poor performance when the experimental intention is to examine the information gathering process in the sense of a free and creative cognitive act. If it is acceptable or even favorable for experimental information to be determined and structured a priori, the opposite is true. In that case, the method performs with an excellent ability to control and assess all experiment related information.

8.2.2 Eye Tracking

Purpose

This technique tracks eye movements of experimental subjects while they solve predefined decision-making tasks. Measurements of the pupil size sometimes accompany this method.[71] The underlying and widely accepted idea is that eye movement, and pupil size are highly correlated to a person's attention and thus to its information processing. Google Scholar lists about 486,000 search results when asking for "eye track* decision process*"[72], revealing a broad application base. This technique predominantly finds application in information gathering and processing research with predefined information location, rather than active information search tasks, because of technical constraints.

Functionality and Points of criticism

In a basic setup, a subject is asked to fulfill a specific (experimentally induced) task including a predefined set of information located in the viewing range. Cameras focus on the eyes of a subject while fulfilling tasks. Eye movement, gaze, and pupil size are recorded and afterward evaluated. It is often of particular interest for researchers how long gazes remain in specific ‘areas of interest’ and which kind of patterns the eye movements show. The data interpretation then usually is based on a predefined behavioral model. In goals and partly in implementation and data analysis lie substantial similarities to the mouse tracking method. The favorable consequence is that models and implementation experiences could be relevant for both methods.

[71] The idea of measuring the pupil size to conclude on cognitive processes or at least on cognition-related phenomena is even older than the eye tracking method itself. In a mainly neurophysiological approach, yet not strictly medically described, Kahneman et al. (1969, p. 164) examine heart rate, pupillary, and electrical skin resistance changes during a mental task. In their study, subjects are asked to solve tasks of assumingly different complexity. As a result, the authors found a direct (functional) relationship between the severity of a task (complexity) and the featured measurements. They conclude that these features constitute reliable indicators of mental effort. The effective connection between pupil size and the state of mind is already known from Hess (1965, pp. 53 f.), in which he referred to the pupil as a 'window to the soul'.

[72] Search dates from 27/4/2016.

Technical variations exist in the number of cameras deployed and their placement during the experiment as well as the information presentation within the task process. Numerous suppliers offer such technique including experiment and analysis software for convenient use and various occasions. Following basic market rules, the still growing number of systems is likely to lead to a drop in prices sooner or later. However, depending on the quality and sophistication of the systems, a single unit at least costs about 100 US-$ and is far from being common in digital devices yet.[73] Despite that, before using an eye tracker, it is urgently necessary to adjust the tracking technique individually to a subject and the visual information of the experiment to ensure accurate and precise data. Adjustment is this technique's severest obstacle, even limiting its potentials. Latest developments in that field suggest an upcoming solution to this problem by automatic adjustment and wearable camera technique.[74]

For the goals of this study, eye trackers surely have the potential to deliver relevant data in both type and accuracy to describe decision-making under time pressure. Nonetheless, decision-making in the sense of articulating one's choice would be best implemented by mouse use. Thus, a mouse (or similar input device) remains necessary. Furthermore, the online experiment approach would be somewhat tricky to implement, because of a substantial lack of technical equipment and possibilities to adjust it. Criticisms regarding presentation form and pre-structuring of information are similar to mouse tracking.

8.2.3 Verbal Protocols

Purpose, Functionality, and Points of criticism

The process tracing method 'verbal protocols' works by asking a subject, either during the process or afterward, about its inner reflections of the cognitive process. A subject then formulates the process with own words or a given set of vocabulary. The experimenter uses so derived information to infer on the cognitive process of information evaluation and integration. This technique thus belongs to the class of information evaluation and integration methods.

---

[73] The cost of some high-end systems is two orders of magnitude above the lowest offer (results of a Google search the author of this treatise conducted in December 2015).

[74] Especially for virtual reality applications, the quality of dynamic orientation components is increasing, which in turn are relevant for adjusting the environment of a camera. Such an adjustment offers the possibility to examine spatially free active information search.

Verbal protocols are in use long before the term process tracing became popular. Already in the beginning of the 20th-century cognitive scientists, searching for specific processes of thought employed self-questioning. This method still finds application nowadays among cognitive research in the decision-making context.[75] Significant points of criticism concern subjectivity and unreliability making it almost impossible to reconstruct results of this method (Nisbett and Wilson 1977a, p. 231; Kühberger et al. 2011, p. 5). The authors further report that verbal protocols explicitly influence the cognitive process. The influence is evident in approaches where a participant is asked to describe thoughts while fulfilling tasks. The verbalization surely binds resources which are no longer available for the cognitive task-solving process. That argumentation is in proper compliance with the Cost-Benefit Model of Beach and Mitchell (1978). The points of criticism seem substantial.

However, the method has its benefits in revealing problem-solving strategies that directly hint to the cognitive process as Montgomery and Svenson (1976, p. 283) and Payne (1976, p. 366) among others show. Ranyard and Svenson (2011) offer an integrated approach where verbal protocols or think aloud data find application among other process tracing methods (pp. 117 ff.). They present a model that includes converging analysis and interpretation from each assessed method. Here, the general principle holds, the more information about the cognitive process is obtained, the richer the interpretation can be. Of course, since each method applied requires its model for analysis, the integration becomes a somewhat complicated task.

Evaluating the potentials of verbal protocols in the actual application environment of decision-making under time pressure, it is primarily the latter aspect that hinders its use here. With the resource competing cognitive tasks of thinking and verbalization, the negative influence of the method on the underlying problem-solving process is evident. This negative influence is tolerable in decision-making environments where time plays a minor role and (parts of the) processes might be solved sequentially rather than in parallel. However, under time limitation conditions this seems to be a severe counterargument. Verbal protocols are hence usually rejected as an adequate method (Ericsson and Simon 1980, pp. 215 f.; Russo et al. 1989, pp. 759 ff.; Ranyard and Svenson 2011, pp. 132 f.). An alternative could be post-decision verbal protocols, as Ranyard and Svenson (2011, p. 125) suggest. They employ recordings of the decision-making process which are shown to a participant to support the memory process. That indeed would solve the resource

[75] See Ericsson and Moxley (2011) for an introduction.

competition problem. The authors do not intend to inform the participants about the post-decision questionnaire before completing their tasks in order to generate authentic and initial data and to avoid "justificatory errors" (Ranyard and Svenson 2011, p. 134, note 1). This procedure, in turn, has its limits in reoccurring or repeated tasks, where post-decision verbal protocol potentially leads to impairment and distortion (Ranyard and Svenson 2011, p. 125). Lyle (2003), among others, describes a possible solution to this dilemma called 'stimulated recall' by video-documentation and video-based debriefing.[76]

All in all, the effort of implementing verbal protocols does not seem to weigh out the expected benefit of so generated data. It is thus not considered for technical realization.

8.2.4 Functional Magnetic Resonance Imaging

Purpose

Functional Magnetic Resonance Imaging (fMRI) is a medical-technique imaging method, mainly employed to visualize physiological functions such as brain activity. This technique also enables to observe dynamic processes within the brain or brain regions. In (neurologically induced) behavioral and decision-making research, images obtained in this way are usually used to relate brain activity to simultaneous behavior. These results are used to infer on the underlying cognitive processes in process tracing research (Coricelli and Rusconi 2011). The fMRI is representative for the class of process tracing methods which are applied to examine physiological and neurological aspects.

Functionality and Points of criticism

The functionality of fMRI and points of criticism in a PTM context are reported by Coricelli and Rusconi (2011), providing most of the following information. The fMRI enables the visualization of changes in blood flow within brain areas, which hint at the current metabolism. The metabolism is an indicator of neuronal activity. The visualization is based on the magnetic characteristics of oxygenated and de-oxygenated blood. Scientists interpret the relative rise of oxygenated blood concentration as a need for oxygen of the cells within a specified/observed area. The likely reason for that is an increasing (cognitive) activity. This method is noninvasive.

[76] This approach increases the technical and organizational efforts significantly. The effort requires a significant reduction in the number of decision tasks, making this method attractive for studying few, but more complex tasks.

Noninvasive brain stimulation requires comparatively less experimental participants since treatment and control group can be identical, enabling intra-subject examinations. The experimental setup is technically quite complex. However, the measurement itself is local and can be done within few minutes making it acceptable in time duration in comparison to the other three presented process tracing methods. Moreover, this method offers the possibility of providing spatial pictures of anticipated brain activity. Two super-national research programs, funded by the US and the EU some two decades and one decade ago, lead to intensified neurological studies, especially improving imaging technologies and interpretation of brain activity. Still, the mapping of brain activity to underlying cognitive processes is to call incomplete yet.

One primary reason for this might be intensive intellectual and financial hurdles (Payne and Venkatraman 2011, p. 232).[77] This fact might also be the reason the group of researchers and scientists able to applicate fMRI being comparatively small. The technique is indeed by far the most expensive one of the four presented process tracing techniques. Handling aspects seem to be further obstacles, since applying this technique requires both neurological education and at least a technical instruction that surely is not an issue of self-study. Another issue is its comparatively slow dynamic behavior. The visualization process requires some seconds to adapt to the changing circumstances, because of the measurement method. That makes its solution somewhat inadequate for examining decision-making under time pressure.

In conclusion, employing fMRI in process tracing studies seems a promising way to examine brain activity in a much more comprising manner. However, its application in the present study seems inappropriate, because of the technical and financial hurdles, as well as the methodological limitations described above.

8.2.5 A Comparison of Relevant PTM

In the following, the four PTMs are compared concerning purpose, functionalities, applicability in online experiment, participant's equipment for online experiment and criticism. Those aspects are complemented by the earlier introduced selection aspects generating proper data, cost-benefit (relating proper data generation to technical and financial effort), being established and providing proper documentation. Table 18 contains the results of the comparison. Entries which do not fol-

[77] Some internet sources put the cost of a (new) fMRI system at up to US-$ 4,000,000 in 2012/2013 (https://www.quora.com/How-much-does-an-fMRI-machine-cost, queried on 2018/09/28).

low immediately from the PTM-related introductions in the preceding sections, are developed and explained in more detail in the following remarks.

TABLE 18 – COMPARISON OF PTM

Aspect	**Mouse tracking**	**Eye tracking**	**Verbal Protocols**	**fMRI**
Purpose	Information gathering and processing	Information gathering and processing	Information evaluation and integration	Physiological and neurological aspects
Applicability	Online / stationary	Online / stationary	Online / stationary	Stationary
Participant's equipment for online exp.	PC, GUI, input device (mouse)	Eye tracking system, GUI, input device	PC, GUI, input device (microphone)	Not applicable
Criticism	Pre-structured information; text-form presentation	Pre-structured information; text-form presentation; requires input device	Subjectivity; impairment and distortion in post-recordings; memory resource competition in parallel recordings	Lack of knowledge concerning brain functionalities; technological and financial obstacles
Proper data generation?	Yes	Yes	No	Yes
Cost-benefit rank	1	2	4	3
Established?	Yes	Yes	Yes	Yes
Well documented?	Yes	Yes	Yes	Yes, requires neurological knowledge

Of all process tracing techniques, eye tracking shows the most similarities to mouse tracking.[78] Techniques of other process tracing categories, which potentially substitute mouse tracking here, are verbal protocols and fMRI. The latter is representative for the set of neurological techniques for evaluating the applicability for the current research goals. The verbal protocol is a rather classical technique that had already been used long before the term 'process tracing' was coined and

[78] In Costa-Gomes et al. (2001, p. 1195, footnote 7) mouse tracking is even referred to as "[…] automated way of doing eye-movement studies […]".

is still in use. Despite its comprehensive history of application, it faces the substantial criticism of being too subjective to be a valid scientific method.

The use of mouse movement data to derive thinking (patterns) from observable behavior is applied with growing frequency in behavioral studies (Willemsen and Johnson 2011, p. 22).[79] Its advantages over the also commonly used eye-tracking technique are simple but convincing for specific experimental approaches: availability, familiarity, and calibration (Payne et al. 1988, pp. 28 f.; Willemsen and Johnson 2011, p. 23). Even though others describe the eye tracking technique as “relatively low cost” (Reutskaja et al. 2011, p. 924), the equipment costs lie at least about two orders of magnitude above comparable mouse tracking systems. Other than eye trackers, the mouse is the standard I/O device of many computers. It is common in handling and availability to most people and thus to potential participants in experiments, too. There is especially no need to calibrate the cursor’s movement on a screen. Altogether that keeps an experimental setup low in preparation effort and costs. Online experiments are possible with mouse tracking, eye tracking, and verbal protocols since the subjects can provide the experimental equipment. However, implementing cameras (eye tracking) and microphones (verbal protocols) in an online experiment with the remote experimenter and few possibilities of technical support for participants is thus less common than mouse usage.

On the other hand, eye trackers directly measure eye fixation and eye movement during a task. Even though tracking mouse movement was developed in analogy to the eye tracking technique, its underlying goal is not necessarily to predict eye fixation. That, of course, would make it less accurate compared to the eye tracker. However, this circumstance is not a disadvantage as Russo (1978, 105) states: whereas a fixation is “a natural observable unit of behavior”, it may neither correspond with the “cognitive unit” nor with the “computational unit”.[80] Just and Carpenter (1976, p. 461) favor gaze time (total viewing time of one or more single eye fixations) instead. Gaze times can be interpreted successfully, whereas single fixations are meaningless. The regularly observed effect of people closing their eyes or looking upwards while processing a task supports this fact. Russo assumes that subjects avoid interruption of the ongoing computing process through new, but irrelevant stimuli (Russo 1978, p. 101). This assumption corresponds perfectly to a single halt of a mouse cursor. The cursor is expected to stand still while the user is pro-

[79] See Willemsen and Johnson (2010) for a recent introduction and application overview in studies.
[80] In that context Russo falsified the often applied assumption of “eye fixation equals computing time” (Russo 1978 p. 104), making its use inappropriate to infer on the cognitive process.

cessing information, as several studies, investigating the relationship between eye gaze and cursor position, implicate (Huang et al. 2011, p. 1229; Huang et al. 2012, p. 1346).

Both process-tracing techniques – mouse tracking and eye tracking – focus on obtaining appropriate data to derive insights into the cognitive process of information acquisition and processing (Willemsen and Johnson 2011, pp. 3 f.). They, therefore, produce the same type of dataset: the spatial location of attention over time.

The demanding challenge both techniques face is the interpretation of the obtained data to link them to the cognitive process. The interpretation could not be based on other than logically derived and at best verifiable assumptions about the connection following a precise model of the process (Costa-Gomes et al. 2001, p. 1196; Kühberger et al. 2011, p. 15; Willemsen and Johnson 2011, p. 32). Moreover, there is no doubt about the interacting coexistence of the cognitive and the motor system as Freeman et al. (2011, p. 59) stated. Due to the similarities mentioned above between mouse tracking and eye tracking, methods and approaches of interpretation for either of the techniques developed so far could work with both of them, too.

Polonio et al. (2015) study eye movements in two-person normal-form games using an eye tracker and classify their experimental subjects' behavior according to their pattern of visual information acquisition. Their interpretation of eye movement patterns as part of the information acquisition process follows the idea of underlying strategies, but relaxes the restriction of strategy matching, since they do not focus on specific heuristics. They classify single (eye) moves with respect to their potential use in a strategy with a certain level of sophistication (Polonio et al. 2015, pp. 93 f.). This perspective-changing approach is similar to the one proposed in this study: its interpretation of data is the interpretation of single moves, without the necessity to put it in an overall picture of a complete strategy.

Also, studying the relationship between eye gaze and cursor position is an active field of research so far. Huang et al. (2012, p. 1341) give a brief introduction to recent studies. Several studies concerning information search on websites mention a strong correlation between eye gaze and cursor position (Huang et al. 2012, p. 1349). This fact is especially valid when the mouse is interacting with elements on the screen by moving and clicking since this is the purpose of an input device. Not surprisingly this field's research concentrates much on predicting the attention from cursor movement and eye gazing in text reading tasks. Various mouse move patterns, such as reading and interacting are reported in this context (Huang et al. 2011, pp. 1228 ff.; Huang et al. 2012, p. 1346). These findings confirm the results of other approaches conducted in decision-making studies, which assume a reliable link between observable behavioral patterns and the cognitive process. Consequently, metric frameworks are developed to operationalize this connection (Costa-Gomes et al. 2001, pp. 1204 ff.; Freeman and Ambady 2010, pp. 229 f.; Huang et al. 2012, pp. 1347 ff.; Polonio et al. 2015, pp. 92 ff.; Devetag et al. 2016, pp. 186 ff.).

Payne et al. (1988, pp. 596 f.) and Costa-Gomes et al. (2001, pp. 1200 f.) employ a particular setup for their experiment to deal with the ‘attention/cursor position’ problem: the information set of a task is presented in table form. Except for the head of the table (containing general information such as the name of an alternative or attribute), all information is made invisible by covering the cells' content. It only gets visible when the cursor is moved over a cell (in some experimental variations a click on the cell is needed). This method should ensure a high correlation between attention and mouse position. However, the possibility of positioning the cursor over a cell, but not paying attention to the content, yet is presently not taken into account. This uncertainty is generally seen as a “small defect of [the mouse tracking] system” (Wang 2011, p. 187). Both studies record the order of cells visited by the subjects and the time spent in a cell. Assume a subject halts the cursor in a cell but draws attention anywhere else. Even though this case cannot be observed directly the halt duration could be a considerable hint. Outside the cell, no information is presented. So, a likely thing a subject is doing when exceeding the expectable reading time is thinking about the content of the cell just visited, possibly in context of the content gathered before.[81]

The studies of Payne et al. (1988, p. 562) and Costa-Gomes et al. (2001, pp. 1200 ff.) use the process tracing software *Mouselab* in the version of 1986. The system is described in a paper published by Johnson et al. (1989). A software update in 2010 extends both name and functionality. Firming under its new name *MouselabWeb*[82], it features the capability to conduct experiments online over the internet. The experiment files are programmed in HTML, JS, and PHP,

[81] Of course, it is possible that a subject leaves the cursor in a cell and alternates the focus between the content and somewhere else. Some authors from studies about the relationship between pupillary size and concentration regard this behavior as an action to concentrate on one's mind (Hess 1965; Kahneman et al. 1969; Wang 2011).

[82] *MouselabWeb* is published under the GNU General Public License and hence free to modify. For further instructions see the projects website http://www.mouselabweb.org/ (visited on 2018/06/08) and the designers’ paper (Willemsen and Johnson 2011).

what makes them easy to modify.[83] Same as Mouselab, the latest version supplies the recording of mouse clicks over time and cell, employing the HTML mouse event handler *onclick* (the event is a click on an HTML object[84]), *onmouseover* (the event is a cursor move to an object; it responds when the cursor passes the object's border) and *onmouseout* (the event is a cursor move away from an object; it responds when the cursor passes the object's border). After 2001, with HTML 4.01 another mouse event handler was introduced: *onmousemove*. It responds to a position change of the cursor in an object. Thus, it provides the ability to track every single mouse movement over time and pixels in an object, resulting in increased resolution of the motion.[85]

Implementing the mouse event handler *onmousemove* in *MouselabWeb* experiment files can enrich the set of process tracing data.[86] Finally, data of the duration, the clicking and moving in all objects could be analyzed and combined to give a detailed picture of mouse movement behavior. Huang et al. (2012, p. 1346) used this enriched dataset to identify subjects' behavioral patterns they called types of mouse usage. The applied method to identify those patterns is at least questionable from a process analytical standpoint. The authors make the identification manually since they have not implemented automated data analysis. The movements are judged and categorized by the experimenters while revising a record of a participant fulfilling an online text search task. This method is of insufficient use when having behavior data on a large scale. Furthermore, it is subjective up to a certain degree and thus negatively impacts the explanatory power of the derived results.

Johnson and Payne (1985), Payne et al. (1988), Johnson et al. (1993), and Costa-Gomes et al. (2001) had to rely on data that only provide response times and element changes. Adding *onmousemove* to the tracking system dramatically improves the database to study cognitive processes in more detail compared to the technique proposed by Johnson et al. (1989). Of course,

[83] All three abbreviations stand for programming languages: HTML: Hypertext Markup Language, JS: JavaScript, and PHP: Hypertext Preprocessor.
[84] The object term in HTML has various implications, describing entities that can be referenced. In the context here, this explicitly comprises visualizations of information such as table cells and buttons.
[85] The website http://www.w3schools.com/tags/ref_eventattributes.asp (viewed 2018/06/08) provides an actual and complete overview on technically implementable mouse event handlers in HTML.
[86] There is a growing number of process tracing software based on mouse movement. Besides *MouselabWeb*, scholars use *z-Tree* of Urs Fischbacher (Fischbacher 2007) and *MouseTracker* (Freeman and Ambady 2010). However, none of the latter two mentioned explicitly supports strategic decision-making experiments in normal-form and HTML to the knowledge of the author. Hence, *MouselabWeb* is meeting the application's needs best. The different mouse tracking software systems are discussed in more detail in Section 9.3.1.1.

this advantage can only be realized with a comprehensive framework to translate mouse movements into EIPs.

With the arguments stated here, mouse tracking seems advantageous concerning experimental costs, applicability and ease of technical implementation compared to eye tracker and fMRI. The intended functionalities are pretty alike, even though eye tracker and even more fMRI assumingly offer a more direct link to the cognitive process. This advantage can be widely compensated by proper adjustment when using a mouse tracker. Thus, the decisive aspect is lastly the availability of the techniques, making mouse tracking the favored choice for further consideration in experiment. All three methods still face the same challenges in interpretation and explanation of generated data regarding its meaning for the cognitive process. Therefore, an interpretation framework must be developed for the selected PTM that takes into account the underlying decision-making tasks. This is the subject of the following section.

8.3 Identifying Elementary Information Processes

8.3.1 Elementary Manual Motor Acts

Even though some EIPs leave direct hints in the behavioral data, others must be traced through interpreting the context.[87] As stated earlier, mouse movement data offers a rich set of behavioral information. In the context of process tracing techniques, Freeman and Ambady (2010, p. 226) use the term 'motor dynamics' to comprise any data derived from hand movement. However, to the knowledge of the author no precise categorization of such data is published so far. In a different context, the documentation of HTML lists basic patterns of mouse movement behavior under the notion of 'event'. Amosov (1967, p. 108) used the term 'elementary motor acts' to refer to all human motor acts, not limited to the manual ones. The author of this treatise follows this notion and suggests designating relevant events 'Elementary Manual Motor Acts' (EMMAs).[88] Since they are directly observable, those EMMAs form the starting point of data interpretation. The procedure is as follows:

[87] The application of mouse tracking as PTM requires interpreting observed behavior in the context of the underlying problem task (Costa-Gomes et al. 2001, p. 1198; Johnson et al. 2008, p. 264; Willemsen and Johnson 2011, p. 24; Schulte Mecklenbeck et al. 2011a, p. 735).
[88] Of course, in this context EMMAs are limited to mouse movement by hand.

1. Identify EMMAs in mouse movement data.
2. Use properties of EMMAs (like timestamps, duration, relative and absolute position on screen) for interpretation, set movements in task context and create a movement history.
3. Derive EIPs (based on underlying movement) from data interpretation.

The programing language HTML provides a number of mouse event handlers which form technical categories. This framework can be used to categorize movement data. The event handler enable automatic recording of mouse movements together with data needed for interpretation, such as time stamp and duration, relative and absolute position on screen and direction of movement. HTML is thus chosen as implementation environment to experiment.[89] The set of EMMAs can then be derived directly from the applied mouse event handlers.[90]

Which kind of mouse event handlers make sense to be applied depends mostly on task and presentation of the task in an experiment. The following contemplation focuses on a payoff matrix within a strategic game task in normal-form as presentation form.

A subject reads the information given in the payoff matrix and later on decides for a (pure) strategy. The time span between the decision task starts and a subject ends the task by announcing a decision provides the relevant information about the mental process of strategic decision-making. A designer's objective here is to cause as much mouse usage as possible without disturbing or diverting the mental process. Of course, that implies the presence of a more profound knowledge or a fitting model of the process which is to be studied itself. Hence, design related biases are likely problems which need to be discussed together with the results of an experiment. Costa-Gomes et al. (2001) in a strategic and Payne et al. (1988) in a nonstrategic task environment showed a possible way to positively influence the mouse movement of participants in their experiments. In both studies, the payoff matrix information is covered at the beginning. By clicking on a cell, it uncovers the hidden information. It should be noted that the authors of both studies use a strict version of that incentive, covering the information again immediately after the cursor left the cell. Other incentives or useful obstacles are conceivable.

[89] The basic idea in applying HTML is to record mouse events to link them to actions on a webpage. The process tracing method of mouse tracking makes use of it by tracing the interactions of a user with the webpage-based experiment.
[90] Of course, that also works the other way around when designing an experiment and a fixed set of EMMAs needs to be tested or deployed.

HTML in its latest version 5.0 provides as many as seventeen mouse event handlers with nine alone added in the last update. The version is in use since the end of 2014 and therefore too new to be supported as standard in every browser at the time of the experiment. In addition, the new mouse event handlers primarily target the increasingly popular 'drag and drop' feature. This feature is not common in normal-form games experiments yet. The focus here is therefore on the HTML 4.01 standard with the event handlers specified in Table 19.

Speaking of functionality in normal-form game tasks, some of the mentioned event handlers are more qualified for use than others. The event handler *onmousewheel* could only contribute if the full information intended to be presented does not fit on a single screen. This is unlikely in the above environment (unless the experimenter designs a highly complex task with a very large number of alternatives).

TABLE 19 – MOUSE EVENT HANDLERS ACCORDING TO HTML 4.01 STANDARD

Event handler	**Description**
onclick	Responds to a mouse click on an element.
ondblclick	Responds to a mouse double-click on an element.
onmousedown	Responds to a mouse button pressed down on an element.
onmousemove	Responds to the mouse pointer moving while it is over an element.
onmouseout	Responds to the mouse pointer moving out of an element.
onmouseover	Responds to the mouse pointer hovering over an element.
onmouseup	Responds to releasing a mouse button over an element.
onmousewheel	Responds to the mouse wheel rolling up or down over a frame element.

The underlying motions of *onmousedown* and *onmouseup* in combination are similar to the motions that trigger the *onclick* event handler. A mouse click down as a 'half click' along with a mouse movement causes the element pointed to by the cursor to be dragged (or at least attempted to drag it). If there is no drag function intended this event handler is of no use. A mouse down without a mouse move needs to result in a click unless a subject is willing to hold it down and do no more interaction until the experiment stops. It is the same with mouse up which usually completes a mouse click. Hence, the mouse click handler can substitute this event.

The last event handler to mention here is *ondblclick*. It responds to a double-click expecting to trigger a different function compared to a single click. This event could be valuable in combination with appropriate experimental design, even though it enlarges the set of possible actions for a subject and thus would add to the complexity of the task handling.[91] One could argue that double-clicking is a natural motion for opening an object concerning (Windows-based) computer applications, whereas single clicks link to selecting objects. Nonetheless, *onclick* can fully substitute *ondblclick* to implement opening and selecting tasks. A necessary precondition is a strict separation of tasks in different elements (e.g., clicking a cell in a payoff matrix corresponds to opening this cell and clicking a strategy button corresponds to selecting this strategy).

The remaining event handlers contribute to the information process tracing and hence are recommended to use. With *onmouseover* and *onmouseout*, one could easily measure the time a subject spends on an element (to read or think about the presented information). Also, one could control the visibility of a cell's information that way. Of course, onmousemove can substitute both *event handlers. Howeve*r, that requires permanently calculating the absolute position of the cursor on the screen over time and comparing this with the absolute positions of the elements. Hence, by employing *onmouseover* together with *onmouseout*, coding effort and CPU's calculating time could be decreased.

The event handler *onmousemove* could act as the crucial process-tracing element since it offers the possibility to record information about the whole cursor movement during an experimental task. The data density one could achieve with it depends mainly on the data speed of the internet connection between client and host. However, recording a time/position stamp every two to ten milliseconds is technically possible. Manor and Gordon (2003) experimentally derived minimal human reaction times of 100 ms to 200 ms, depending on the task and the employed PTM (pp. 90 ff.). The author determined a maximum cursor speed (in pixels per second) of five to ten pixels per milliseconds in a pre-experiment.[92] The author measured smaller velocities in tasks where

[91] The complexity of task handling in experimental terms is a serious point of criticism concerning the data quality since it potentially affects the mental process of decision-making. This issue is a major critique of the process tracing technique of verbal protocols, and a disadvantage of mouse movement tracking (Kühberger et al. 2011; Schulte Mecklenbeck et al. 2011a).

[92] Assuming a maximum of ten milliseconds between two measurements and a cursor velocity of ten pixels per milliseconds the cursor could 'jump' about 100 px or 3.2 cm (at a 72dpi-solution) between two measuring points. It is most likely that a cursor takes nearly the straight way between the two measuring points. A circle or curved way seems almost impossible at this speed (Huang et al. 2012, p. 1343). With this assumption, one can interpolate the cursor track accurately.

cursor movement precision is relevant.[93] The recording frequency of the event handler *onmouse-move* thus seems to be sufficient to follow the mouse trajectory.

A mouse trajectory general consists of a stream of moving, clicking and halting for a predefined time span (working as a threshold to separate a stop from movement). Following this, the deduction of EMMAs from the above-discussed mouse event handlers results in an elementary set as shown in Table 20.

TABLE 20 – SET OF EMMAS AND DEFINITION

EMMA	**HTML mouse event handler**	**Definition**
CLICK	*onmouseclick*	Mouse click (left or right) on a HTML object.
MOVE	*onmousemove*	Mouse moves at least one pixel in x- or y-direction.
HALT	any	Mouse halt; i.e., the duration above a certain (pre-defined) threshold.

8.3.2 Interpretation Metric

Apparently, the set of EMMAs is too small compared to the set of EIPs to ensure a bijective mapping. Therefore, it is necessary to employ additional mouse movement data. As suggested above, the movement in context of the information presented on the screen is essential to be taken into account (Payne et al. 1988, p. 563; Costa-Gomes et al. 2001, pp. 1227 f.; Devetag et al. 2016, p. 190; Polonio et al. 2015, pp. 82 ff.).

Costa-Gomes et al. (2001, p. 1210) and Willemsen and Johnson (2011, p. 24) propose two testable assumptions that greatly help to interpret a subject's mental processing based on the information acquired. Those are the paradigms 'occurrence' and 'adjacency'. The former postulates that information must be visible before one can use it. 'Adjacency' postulates that information gathering is highly correlated to information use. This postulation implies two aspects. First, a subject almost has no personal information such as pre-knowledge concerning the particular task that finds application in decision-making. Second, subjects process the information that is recently acquired. Information is not stored for long for later use. Bits of information that are no longer

[93] Fitts' Law is dealing with the speed-accuracy phenomenon in motor system tasks, linking the speed to the size of and the distance towards the object (Fitts 1954, pp. 267 ff.).

present in memory are acquired again. However, adjacency does not affect the aspect of knowledge about problem-solving methods such as heuristics (Costa-Gomes et al. 2001, p. 1210). Both paradigms find application in the analysis of many process-tracing studies.

Polonio et al. (2015, p. 86), for example, identify twelve eye movement patterns within specific areas (they call areas of interest) of normal-form game payoff matrices. They mainly record movements within a cell – henceforth labeled as the inner-cell movement – and movements among cells – henceforth labeled intercell movements. Gathering and analyzing data from those two categories, they derive potential strategic interests of their subjects. It adds to the interpretation to log the movements from one cell to another (intercell movement) and from own payoff to the opponent's or the other way around (inner-cell movement). The interest lies in the number of defined movements as well as in the absolute time in which a subject concentrates on certain information (Payne et al. 1988, p. 562; Costa-Gomes et al. 2001, p. 1202; Johnson et al. 2008, p. 264; Willemsen and Johnson 2011, p. 27; Polonio et al. 2015, p. 84).

In a 'Private Information' game[94], Brocas et al. (2014, p. 965) identify a specific payoff per game they assume to be highly relevant to distinguish a strategic from a nonstrategic lookup behavior. The authors relate this to the duration and frequency this specific payout is observed and correlate that data with selection data.

The current process tracing literature concerning the information board technique provides two prominent approaches to measure time and frequency of looking up information in normal-form games. In the first approach, the experimental setup separates own and opponent's payoff information into different cells (e.g., Devetag et al. (2016, p. 185)) or in two different tables (e.g., Costa-Gomes et al. (2001, p. 1201)). One unit of information is presented per cell. This way, it is simple to record which payoff is in focus. In separating the payoff information, a subject is guided to look through either one's own or through the opponent's payoff, especially in the design of Costa-Gomes et al. (2001, p. 1201). One could argue that this way deviates from the orthodox normal-form or even favors a certain procedure over another. In terms of process tracing, these arguments are unconvincing because it is most important to identify the focus of attention. And both methods have succeeded.

[94] In Private Information games, players have incomplete information about their opponents. For example, their respective utility functions are unknown to the opponents. This requires developing expectations about the unknown information to infer on the behavior of other players.

Eye tracking studies predominantly apply a second approach as follows. Among others, Polonio et al. (2015, p. 84) separated own and opponent's payoff in one cell through setting the information in different corners of a cell. This way, the payoff information is distant enough to be distinguished when a subject looks at one or the other.

A third way to identify the focus of information gathering is close to the second presented here. This one represents an appropriate adaption for mouse movement needs. At this point, it is essential to discuss the relationship between cursor position on screen and user's attention to detail.

One widely regards the cursor as an item that helps to focus on particular information on the screen (Huang et al. 2011, p. 1225; Huang et al. 2012, p. 1342). The head of the arrow serves as the point of attraction. So, wherever the cursor points to, it is assumed to be the point of interest. Huang et al. report a spatial gap between eye focus and cursor position in reading tasks. The size of the gap usually depends on the type of mouse user, having one type following reading horizontally, one following vertically and one not using the mouse as reading aid. Not surprisingly, the distance between cursor and eye focus is smallest for the type following reading horizontally. The reading task, of course, is different from the decision task studied here. The reading part in a cell is limited to the payoff. So, overall, the reading is not expected to last long and a cursor that supports reading the way Huang et al. (2012, p. 1346) report is not very likely.

However, marking a point of interest for other than reading, for instance, thinking is what one could expect instead. If the mouse is in use and not halting immediately after entering the cell, one could assume that the tip of the cursor points to the eye focus and hence marking the current area of interest of a subject. If the cursor halts immediately after entering the cell, the tip most likely points to an area close to the border of the cell. However, what appears at first glance to be speculation should be confirmed by repeated observable actions of a subject.[95] Therefore, the idea of Huan et al. to classify the behavior of mouse users seems to be appropriate and useful, even though their manual analysis approach is not feasible for a significant volume of data.

Besides, they found that the size of the gap between eye gaze and the mouse pointer depends on the movement speed, tending to have a smaller gap the faster the cursor is moving. That finding underlines the reliable connection between cursor, eye fixation, and the point of interest. So,

---

[95] Since most experimental tasks based on normal-form games contain a payoff matrix that consists of four cells or more, one could identify a person with just a limited use of the mouse very soon. Nonetheless, there are ways of enhancing mouse movement through incentives, unlikely to disturb or influence the thinking process.

whenever the cursor is moving, the attention is, too. Following that, the beginning of a cursor move at the same time marks the end of the attention to the point visited and halted before. This statement is confirmed by Russo (1978, p. 92).

In summary, a typical mouse trajectory consists of movements interrupted by several cursor halts.[96] If an interaction is part of the design, then a halt could also involve one or more clicks.[97] The phases of attention generally correspond with cursor halts. Hence, the duration of the mouse movement stops as well as the current location needs to be detected. That is of minor technical complexity regarding its implementation as will be shown later.

Of course, there are prominent exceptions of a typical mouse trajectory. One example is an oscillating mouse movement between the payoff of two not necessarily adjacent cells or within one cell. In that case, it is most likely that the observer is trying to deal with the information of both cells or of both players from one cell at the same time. The interpretation again depends on the moves the subject had made before that. The record of movements helps to form a proper context, open to interpretation. Not surprisingly, the EIPs COMPARE and ADD are becoming relevant in this context.

Another example of an atypical mouse trajectory is the circle movement where the attention lies in the center of the circle. One can interpret the movement here as a means of concentration. Generally, this could not be spotted in eye tracking studies, since the eyes are focussing on the information a mouse move would circle.

It should be noted that the stated exceptions of oscillation and circle movement need a rapid mouse movement on their turning points.[98] This condition is met when the turning points are left faster than the halt threshold. If the move reaches that speed, a halt is detected, and the recorded trajectory is no exception anymore.

The HTML event handler *onmousemove* tracks both mouse movement and hold times within a cell. It uses the arrowhead to locate it. Hence, tracking the coordinates of the cursor means tracking the location of the arrowhead.

[96] The time span characterizing a halt is discussed and defined in Subsection 8.3.2.3.1.
[97] Clicking while moving is regarded as an untypical behavior since it evokes a dragging effect on the clicked element, and this is usually not a function implemented in a normal-form game environment.
[98] That is the point, where at least one of the x- and y-coordinates changes direction.

Now assume a normal-form payoff matrix such as the one shown in Figure 21. A subject in an experiment faces the role of a row player with own payoffs stated first and opponent's second, separated by a semicolon and two surrounding blanks.

FIGURE 21 – PAYOFF MATRIX

The following consideration focus on a single, opened cell. Figure 22 contains an example of a magnified open cell. Two parts are visible that divide the area. The information on own and opponent's payoff covers one part. The second part represents the surrounding empty area.

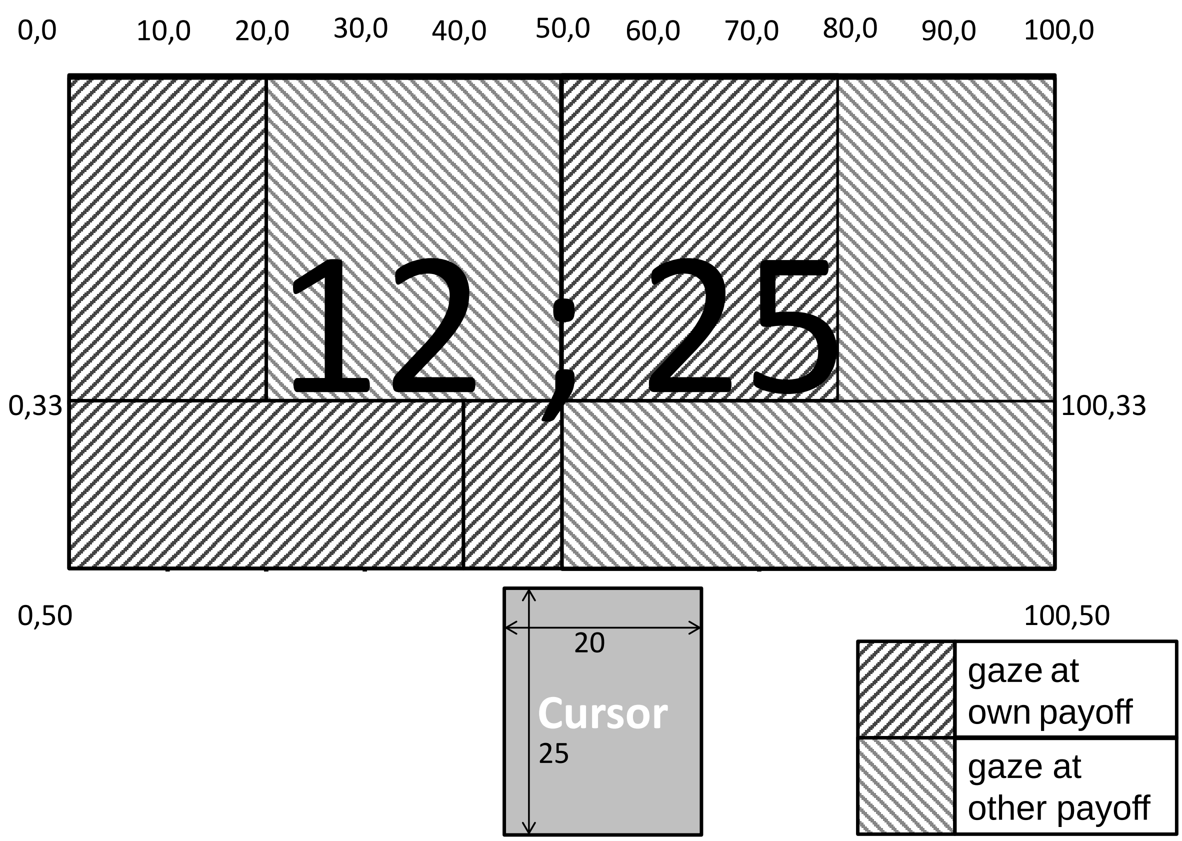
The next step is to develop a relationship between the position of the mouse cursor within the cell and the potential attention. If the body of the pointer is located on one of the two payoffs (own or opponent), it covers this very payoff information. It thus cannot be in focus. So, the size of the cursor is also vital to be taken into account. It primarily depends on the screen solution but has a common relative size of twenty by twenty-five pixels. Even though the standard cursor has a top-heavy arrow shape, the area it covers could be approximated by a rectangle (see the gray rectangle in Figure 22).

Two fundamental assumptions regarding position and size of the cursor help to define the areas of attention in a cell:

1. If the cursor area (partly) covers payoff information that payoff is not in focus, but the other is and
2. if both pieces of payoff information are completely visible (i.e., not covered), the nearest information to the tip of the arrowhead is in focus.

Both assumptions consequently follow the paradigms ‘occurrence’ and ‘adjacency’ proposed by Costa-Gomes et al. (2001, p. 1210) in case of information acquisition and processing within a single cell. Following these assumptions, the area of a cell could be divided into zones of attention as shown in Figure 22. Areas in the same hatching represent attention toward the same piece of information when the mouse pointer stops in there. The area borders are of no spatial extent, having a defined width of zero px. With this, there are no neutral zones in a cell and hence no indecisive points. The cell from the example in Figure 22 above has a size of 100 px in width and 50 px in height, the information has a font size of 12 px. The area of attention within a cell can be separated in two almost equally sized parts. Depending on the location of the mouse cursor, the interpretation is either 'a player looks at own piece of payoff information' or 'a player looks at opponent's piece of payoff information'. The share of the two parts is just slightly changing with the payoff's number of digits varying. Assuming that every pixel in a cell is almost equiprobable frequented[99], none of the two pieces of payoff information is favored by the design of information presentation regarding the supposed attention. With this schema a measuring method exists that links a cursor's position in a cell to a person's attention and hence can measure its time length and frequency.

FIGURE 22 – SINGLE CELL AND GAZE AREAS



[99] However, Polonio et al. (2015, p. 11) remarked in their eye tracking study that "the level of attention towards each AOI (area of interest) was […] influenced by the spatial location of the payoffs in the screen […]", with the own payoff being significantly longer in focus.

Decision information is the information about the subject's final choice in a decision task. The HTML event handler *onclick* records it. It is mostly expected to emerge as the final stage of the decision process. However, experiments under time pressure indicate that people make decisions at every stage of the cognitive process. This behavior is especially expectable in task environments where making no decision is somehow costly, or the subject can revise the initial decision at any time of the process (Caplin et al. 2011, p. 2900).

Polonio et al. (2015, p. 81) successfully forego choice information when classifying behavioral patterns in decision making. However, many authors of process tracing studies depend on decision information to evaluate human behavior regarding the outcome of the process. It is also used to determine the level of strategic reasoning in their studies (for instance Payne et al. (1988, p. 541), Costa-Gomes et al. (2001, pp. 1206 ff.), Devetag et al. (2016, p. 179)). Decision information thus remains a principal argument in the authors' reasoning of behavior. That is surprising since the cited studies are conducted in the spirit of process theory. This approach explicitly seeks to offer an alternative to the then well-established decision-based behavior research in decision theory (Simon 1978, p. 2; Kühberger et al. 2011, p. 2). Even those, who compare data with heuristic models assume that the final decision in the task relates to the strategic sophistication or a potential heuristic. That could be a reliable way of operating in some cases, even though it faces notable points of criticism:

1. The explanatory potential of decision information, in general, is likely to decrease with rising task complexity since a growing number of alternatives leads to a decreasing probability of choosing one option over another.

2. Almost all of the heuristics discussed earlier in this paper provide a random decision if the process does not lead to a single final option.

3. The ongoing heuristic process could be obstructed by many different internal and external factors, such as time pressure or dwindling interest, leading to an early and per definition random choice. So, even if a subject begins with a specific heuristic, it is not necessarily finished properly.[100]

[100] Proper termination of heuristics implies a completed heuristic procedure.

Of course, an experimental design with repetition of similar tasks could compensate some of the critical points up to a certain degree, but in that case, learning plays a predominant role. As stated earlier, the effects of learning are explicitly not subject of this study.

So, it is at least questionable whether a subject chooses an option with certainty. Hence, linking a single decision to a heuristic or strategy to conclude on one's behavior is doubtful. For the reasons stated above, decision information should not be the only argument to infer on the cognitive process. For a comprehensive analysis regarding the proposed concept of preparation time, it is thus considered together with other process information. This approach is in good compliance with a study in decision field theory of Busemeyer and Townsend (1993).

In the following, a systematic link between EIPs from Table 12 and EMMAs in Table 20 is presented, serving as an operational basis for interpretation of behavior. The underlying scenario is a normal-form game as stated earlier in Figure 21, p. 123. The subsequently proposed connections between EMMAs and EIPs are presented in the following form, serving as definition for EIP i:

EMMA X_1 (context information) & … & EMMA X_n (context information) =: EIP i,

with X_1 to X_n being EMMAs from the given set of EMMAs (Table 20). The definitions are introduced with a consecutive Roman number. The context information could be any information to identify the action, i.e., the move history as the history of prior moves, the cursor location on the screen or the halt duration.

8.3.2.1 MOVE, OPEN, ELIMINATE, CHOOSE (I) & END

With the assumptions above, moving the cursor to a cell is equivalent to moving the attention and hence simple regarding a connection:

I MOVE (cursor) =: MOVE (attention).

In the beginning, all cells are closed. They can be opened and closed by click. The cursor must be anywhere in the area of the cell for opening and closing. Herewith, the connection between the EIP OPEN and EMMAs is

II MOVE (cursor in the area of a closed cell) & CLICK =: OPEN,

and the corresponding case

III MOVE (cursor in the area of an open cell) & CLICK =: ELIMINATE (cell from consideration).

The EIP CHOOSE can be relevant in different situations, as discussed along with the development of the heuristics' production systems. In the following, choosing between different strategies is called CHOOSE (I). Choosing one strategy over another is usually part of the task in a normal-form game. It is thus explicitly considered in the experimental design by deploying buttons for each of the options. With this, one can define CHOOSE (I) as:

IV MOVE (cursor in area of a strategy button) & CLICK =: CHOOSE (I) (option/ strategy).

A mouse click on the finish button terminates the task. One can formulate the relation as follows:

V MOVE (cursor in area of finish button) & CLICK =: END (task).

Another criterion to end a task in time pressure studies could be exceeding a time limit. The program stops the working process automatically. Here, no explicit action of the user is required. One cannot observe a certain EMMA in this case.

8.3.2.2 FOCUS & CHOOSE (II)

The focus of attention generally lies on the arrowhead of the mouse pointer, as elaborated earlier in this paper. Nevertheless, for the individual case, the validity of this statement highly depends on the intensity of a subject's mouse usage. That includes both movement (e.g., length and duration) and limitation to specific functionalities of mouse usage or interaction. Some authors address this individual behavior as 'type of mouse user' or 'mouse use patterns'. Costa-Gomes et al. (2001, p. 1194) assume that the type is at least constant over a period of one full task when identifying their mouse usage types. Others relax that assumption by introducing behavioral patterns which characterize some aspects of the mouse trajectory (Huang et al. 2012, p. 1346). Hereby, the classification of a mouse usage type can be realized by identifying such patterns within the movement trajectory and determining their frequency.

For the intentions of this research, it is necessary to evaluate subjects’ mouse activity. Hence, distinguishing activity from inactivity is crucial. Huang et al. (2012) study mouse behavioral patterns in reading tasks. Following findings of Claypool et al. (2001, p. 6), they examined inactivity as a behavioral pattern. Huang et al. (2012, p. 1346) define inactive as lack of mouse movement over a specified period ($\geq 1\ s$). Their experimental task was different from the one studied in the current paper. Thus, behavioral patterns are expected to differ here. However, linking inactivity to an absence of mouse movement over time seems reasonable. Halting times could be interpreted as part of the cognitive process and hence are of information, especially when its duration exceeds just a single second. Thus, the circumstances of an information-less halt need to be specified. One applicable definition of inactivity, further used here, reads as follows:

VI Mouse INACTIVITY is the absolute absence of any mouse moves and interactions, except for the minimum number of moves needed to get access to information.

Definition VI implies that mouse moves decrease towards an absolute minimum.[101] Almost no movements occur except the ones to open a closed cell and finally to choose an option. Cursor movement gives no hint for any other cognitive action. The location of the cursor could be either in a traceable cell, without any movement or somewhere out of the interactive part of the screen where the movement is not traced. In that case, the mouse movement dataset is reduced to the one used by Costa-Gomes et al. (2001) for a strategic and by Payne et al. (1988) for a nonstrategic task environment. Even though deriving EIPs through the analysis of EMMAs seems impossible then, both studies mentioned above offer alternative opportunities for behavioral data analysis.

The definition is valid for both, pattern and type: if the cursor leaves the traceable interactive part of the screen, it is recorded as inactive. If it is in a traceable part, the status is depending on its previous and subsequent moves. If the movement is close to the minimum set of moves, the halt counts as inactivity. The halt is likely to result from little motivation to use the mouse as an aid to support the cognitive process. A long-lasting thought concerning the piece of information the mouse cursor currently points is rather unlikely. Prior experiments teach that inactivity usually occurs with very few, but very long halts relative to the whole task processing time. Herewith, the halt time is defined relative to the whole task activity and compared to benchmarks resulting from simulation or empirical value. The overall individual performance finally determines the interpretation of this particular pattern.

Following this definition, every other form of movement represents an activity. EMMA patterns can describe those movements. With this, the connection between EIP and EMMA for active mouse usage reads as follows:

[101] That minimum set of moves depends on the particular game design. It is generally larger if one can only visualize the content of a single cell at a time, compared to the case where cells can be opened independently.

VII (current cursor location) =: FOCUS.

Some of the examined heuristics' production systems, mainly the strategic ones, suggest a form of choice beside the selection of strategies: the procedure requires an implicit first decision for the focus on either of the two possible payoff information perspectives, own or opponent. As stated earlier, choice results could be inspected considering the repeated FOCUS on one of the two player's payoff. The more frequent a subject focusses on one player's payoff in its proceedings, the more substantiated the assumption about a subject's decision regarding the focus. Expected patterns thus show periods of consecutive focussing of same player's payoff. In theory, such periods should contain visits of all cells from one choice alternative at least. A testable hypothesis for such periods would be:

H_1: "The focus of attention in period x lies more often on player y's payoff." against

H_0: "The focus is equally split on both players' payoff."

One need to analyze mouse coordinates with respect to this hypothesis. Following this, the link between EIP and EMMAs reads as follows:

VIII (repeated (current cursor location)) =: CHOOSE (II).

8.3.2.3 Complex Pattern-based EIPs

The identification of the remaining EIPs READ, COMPARE, ADD and STORE is based on the three key elements 'halt duration', 'movement patterns' and 'context information'. In the following, the implications of the three key elements are discussed in more detail.

8.3.2.3.1 Halt Durations and their Interpretation

Every reaction on a stimulus, such as newly presented information, consumes time within the cognitive process (for recognition) as well as in its psycho-motor extensions (i.e., eye–hand system). During that period of information processing and reaction preparing[102], usually no other action is recorded. This behavior is equivalent to a halt, or a fixation[103].

[102] For a detailed psychophysical description of an eye-brain-hand interaction in a reaction task see Braun et al. (2003, p. 662).

[103] Note that the term 'fixation' in this context is used to describe a nonmoving cursor over a specified period. The points of criticism regarding eye fixation that Just and Carpenter (1976, p. 476) report thus do not account here (also see Russo (1978, p. 105)).

The thresholds introduced in the succeeding paragraphs are used to discriminate fixations from movements and classify fixations on the base of halt durations. The classification represents the intent suggested by such a pattern. All of the process-tracing techniques mentioned above employ time discrete recordings of movement. The measuring accuracy of the PTM data scan corresponds with the frequency single data values are recorded. Their technical documentations specify such frequencies as follows: fMRI: ~2 kHz, eye tracker: ~1 kHz, and mouse tracker: ~500 – 250 Hz. Moran et al. (1983) report a minimal perception time with >=180 ms. Hoffrage (1999, p. 151) determined a minimal average perception time of 325 ms per box view under high time pressure in his Mouselab-based information board experiment. Reutskaja et al. (2011, p. 914) report a mean eye fixation duration of 373 ms for their experiment. According to these experimental results, a scan requires a rate of at least 6 Hz. The PTM's measurement accuracy is hence at least one order of magnitude below the expected minimum perception time, playing a minor role in the interpretation of halts.

Besides recording precision, short-term halts of different durations can be detected very frequently in a subject's behavior. The classification of fixation and fixation durations vary markedly for applied process tracing techniques, the experimental task given to the subjects of the study (Kleinmuntz and Schkade 1990, pp. 14 ff.), and individual abilities.[104] This fact is not surprising at all. It can be interpreted as support for the EIP approach and the processing theory, as will be shown in the following. First, assume that the brain is the starting point of a reaction with the physiological reaction system being brain-eye-hand, in this order. The process tracing techniques then measure reactions at different stages of the reaction system. Spatially closest to the brain is the fMRI that measures brain activity.[105] Eye tracking measures eye moves and mouse tracking hand moves. With growing distance to the brain, the time delay of the observable reaction increases, too. Furthermore, different tasks evoke different reaction times because different EIPs are involved.

[104] Processing speed is commonly linked to the ability of an individual's short-term memory, as reported by Bull and Johnston (1997, p. 5) who analyzed numerous relevant studies. From this can be followed that processing speed differs among individuals. The individual diversity is confirmed in experiments by Dansereau and Gregg (1966, p. 71) for mental calculation skills and by Kaplan et al. (1993, pp. 263 ff.) for judgment tasks.
[105] Gonzalez et al. (2005, p. 10) for example study reaction times for a visual stimulus using the process tracing method of fMRI and scale their findings at around 10 ms.

The commonly used threshold for eye fixation studies is about 200 ms. The time value is confirmed by reading experiments, as Salthouse and Ellis (1980, p. 213) inform.[106] Reutskaja et al. (2011, p. 904) use a threshold of 50 ms for fixation times of pictures in their eye tracking study but give no further explanation on how they choose this level. Manor and Gordon (2003, p. 92) suggest a threshold of 100 ms after conducting an experimental study on visual fixation times in free-viewing tasks of geometric objects and human faces. They find that this threshold time is optimal to “discriminate fixation from other oculomotor activity effectively […]” (Manor and Gordon 2003, p. 85). Moreover, some findings suggest that people especially recognize common single words or few-numbered digits as a single pattern.[107] Hence, they can be processed as one unit of information in a single step as Damarla and Just (2013, p. 2630) report. Since the 100 ms threshold is measured for single information units, it should be adequate for the visual recognition of common single words or few-numbered digit patterns, too.

The threshold levels are derived from measuring either brain activity or eye fixations. Following the proposed brain-eye-hand-system, mouse move based studies additionally need to take into account the time of activating the hand to move and thus have generally larger thresholds. A study of VanRullen and Thorpe (2001b, p. 665) indicates a duration of about 100 ms and less for the motor reaction part (of pressing a button). In general, the other part of the threshold is dedicated to the minimum perception time (Payne et al. 1988, p. 563), with a duration comparable to findings of other process tracing techniques’ studies. Efron (1970, p. 62) determined this time span as being between 120 ms and 240 ms for visual tasks. Adding the two parts of the threshold leads to a total of about 200 ms – 300 ms. As the studied experimental tasks are consisting of less demanding reading tasks such as cells filled with just a single word or few-figured digits, a lower bound of about 200 ms is preferred here.

Some mouse move based studies claim a second way of determining an applicable threshold. The studies frequently referred to in this treatise (Moran et al. 1983; Payne et al. 1988; Costa-Gomes et al. 2001; Willemsen and Johnson 2011) cannot discriminate between move and inactivity of a

[106] Probably not the first to report, Russo (1978, p. 105) confirms a typical eye fixation of about 230 ms.
[107] Salthouse and Ellis (1980, p. 207) analyze determinants of eye fixations and give a simple but convincing physiological reason for eye movement and fixations: "In reading and most other visual search activities the eyes typically move between two and five times per second in order to bring environmental information into the foveal region of clearest vision". Following that, a pattern in the understanding of the present study is the information, which is in the foveal region of the most precise vision at the same time.

cursor within a cell because of their experimental design.[108] Instead, they measure the time a mouse pointer spends within a cell. To determine a threshold, the authors of the studies estimate the minimal moving time from one cell to another, arguing with Fitts' Law (see footnote 93). Payne et al. (1988, p. 563) determined a threshold of less than 100 ms for their setup. That duration holds for the other mentioned studies in this context too, since they applied the same standardized software (*MouseLab*). The minimal perception time forms the other part of the sum representing the threshold. The proposed threshold then adds to about 180 ms to 200 ms – a duration comparable to eye fixation studies and the one developed above. Note that in comparison to the procedure described above, the theoretically indicated part of moving the cursor out of a cell substitutes the motor activation part to estimate a reasonable threshold here.

Literature about motor reaction times related to the present study concept is comparatively rare. However, a threshold of 180 ms to 200 ms suggested for subject's minimal perception time including the motoric activation part is confirmed by different methods and thus seemingly robust. Besides the neurophysiological argumentation for this level of time and the different methodology to determine the threshold, the most persuasive argument here is namely the comparability with other mouse move based studies. The reading tasks proposed in this paper (EIP READ) hence apply this threshold level.

Using the terminology of Kahneman's model of cognition (Kahneman 2012, p. 33), perception as well as the EIP READ seems to be executed by the fast and seemingly effortlessly working System I.[109] Findings of Russo (1978, p. 100), VanRullen and Thorpe (2001a, p. 655), and VanRullen and Thorpe (2001b, p. 459) tend to support this suggestion. However, Just and Carpenter (1976, p. 476) point out that the act of reading in combination with perception highly depends on the task. It is thus quite possible that the slowly and effortful System II is applied when reading, say, a legal contract to understand its implications. In the case here, digits are read and perceived as patterns. No such contextual constraints like in the example mentioned above exist. Hereby, the minimum perception time is a pure neurophysiological matter and should not differ much between subjects of sound mental health.

[108] Those studies apply versions of *Mouselab* that have not embedded the event handler *onmousemove* or a similar function, and thus movement within a cell cannot be traced.

[109] See the introduction of Kahneman's model of cognition in Chapter 3.1, p. 14. He postulates two systems of cognition: a fast and seemingly effortlessly working 'System I' that is less accurate and a comparably slowly and effortful working 'System II' that is more accurate.

On the contrary, the EIPs COMPARE, ADD and STORE need cognitive attention and hence mental resources while processing[110] – properties associated with Kahneman's System II (compare Kahneman (2012, p. 33)). If this holds and this very part of the cognitive system Kahneman refers to as System II process those EIPs they consume resources which in turn takes time. Some studies deal with estimating duration times for several types of EIPs, intending to characterize EIPs by their processing time. That would make identification easy. Their calculations predominantly rely on eye fixation data from within-subject studies or simulation.[111] However, other studies suggest that processing times of those EIPs need to be related to individual abilities (Bull and Johnston 1997, p. 19; VanRullen and Thorpe 2001a, pp. 657 f., 2001b, p. 456). Hence, the expected halt durations are likely to differ from subject to subject for the same type of EIP. It is not even sure whether a subject's processing time for the same type of EIP is constant with differing tasks or task complexities (Johnson 1990, p. 142). Thus, determining the three EIPs mentioned above by solely measuring processing times at least raises methodological concerns and is not followed in the present research approach. The following section proposes an alternative.

8.3.2.3.2 Movement Patterns and Context Information

The determination of the EIPs COMPARE, ADD and STORE from the mouse move data stream can be achieved by primarily identifying relevant mouse movement patterns, supported by estimating processing times. This approach has its origin in Payne (1976, p. 369) who counts horizontal and vertical moves[112] and interprets them as either inter-attribute or inter-alternative comparisons. Several other studies so far follow this idea, just adapting their interpretation to the examined tasks (Payne et al. 1988; Bettman et al. 1990; Payne et al. 1992; Costa-Gomes et al. 2001). For normal-form game tasks, two studies are worth citing here:

---

110 All three of them use STM to store (STORE) or recall information for the next step in the process (ADD, COMPARE). Further insights into the interaction of STM and arithmetic provide the study of Bull and Johnston (1997) for instance.

111 Dansereau (1969) estimates the duration of single-digit additions of university students by 0.8 to 1.1 seconds (s), reading and comparing of single digits to 0.3 s, giving reference times for the EIPs ADD, READ and COMPARE (as reported in Johnson and Payne 1985, p. 405). He, therefore, employed a mixed approach of eye fixations and verbal protocols for his model. These values are used as reference within theory building and data analysis in several related studies (Simon 1974; Ericsson and Simon 1980; Johnson and Payne 1985; Payne et al. 1988; Russo et al. 1989; Bettman et al. 1990; Johnson 1990; Ericsson and Kintsch 1995). Russo (1978, p. 108) determined the duration of MOVE in his eye movement experiment with 0.23 s. In a simulation, Johnson et al. (1989, p. 50) could confirm this value. Johnson (1990, p. 18) derived reference times for several EIPs from simulation which he emphasizes are "mostly in line with prior cognitive studies": ADD 0.84 s; READ (encoding information + motor activity) 1.19 s.; COMPARE: 0.08 s; ELIMINATION: 1.8 s. Reutskaja et al. (2011, p. 923) give a threshold for fixation times in consumer search tasks with 0.35 s which might fit in the same category as the later presented EIP COMPARE II.

112 Note that a move in that context means a cursor change between cells.

Costa-Gomes et al. (2001, pp. 1204 ff.) dedicate a substantial part of their paper to the analysis of the information acquisition process within their econometric framework. They assume that subjects fully apply particular heuristics (they call 'types') on similar tasks. This type determines information acquisition behavior and due to that a subject's decision. In their framework, the authors propose thirteen variables measured to derive the type's implications. When deriving potential behavior from type definitions, Costa-Gomes et al. (2001, pp. 1204 ff.) follow a strictly normative path. This approach is impressive because of both the elaborateness in the heuristics' procedure and the numerous assumptions they need for classifying subject's behavior. All of them are plausible.

Some essential process variables Costa-Gomes et al. (2001, p. 1196) measure concern payoff lookups and the length of lookup sequences. Lookups are categorized regarding own and opponent's payoff information. Lookup sequences they define as consecutive lookups of one player's payoff. Within this sequence, the authors separately measure the number of horizontal and vertical moves from cell to cell.

The analysis relies on detailed knowledge of the heuristics' problem-solving procedure. Consequently, one can obtain benchmark levels for process variables from the heuristics performance data. Starting with the normative descriptions of the heuristics, Costa-Gomes et al. (2001, pp. 1222 ff.) count the minimum necessary lookups (own, opponent) and moves (horizontal, vertical) for each heuristic applied in each inspected game task (varying in size from 2 by 2 to 4x3 and payoff structure). This data offers a criterion to decide whether the number of lookups and moves a subject has made in an experimental task is enough for a specific heuristic. So, if the number is lower than the normatively derived one, a subject has not gathered all information necessary for a particular heuristic. Hence, this specific heuristic is unlikely to be employed by the subject. It serves as a lower bound. However, one cannot infer anything specific from this.[113]

Concerning the lookup sequences, Costa-Gomes et al. (2001, pp. 1210 ff.) assume that all of their inspected heuristics will show adjacent lookup pairs.[114] Those are distinguishable from random lookups since these heuristics primarily focus on one player's payoff simultaneously which is located in separate tables.

[113] This fact becomes evident when accepting that subjects are usually not stringently utilizing a heuristic, if at all and thus showing unexpected lookups.
[114] Note that this follows an assumption arguing that comparisons are usually made pairwise. Even if it is no regularity, this effect is observed very frequently by several decision-making studies (Saaty 2008, pp. 95 ff.).

Costa-Gomes et al. (2001, p. 1233) mainly work with averages. Time durations are represented by the assumption that "a type's relevant look-ups have longer average gaze times than other look-ups". The average value shall be appropriate to distinguish between heuristics. For identifying EIPs, which are said to be connected by mouse moves and to differ in time, this approach seems somewhat inapplicable. Still, some parts are – as will be shown in the following.

Devetag et al. (2016) study eye movement patterns in normal-form game tasks. In contrast to Costa-Gomes et al. (2001), they used a task design where both players' payoffs are situated in one table, paired in one cell each (Devetag et al. 2016, pp. 184 ff.). For their behavioral analysis, the authors define areas of interest (AOI) around the payoffs in each cell (Devetag et al. 2016, p. 185). They measure time spent in an AOI (fixation time), number of lookups of an AOI (fixation count), number of returns to an AOI during one trial (number of runs) as well as number and types of saccades (i.e., eye move from one to another AOI). Cluster analysis identifies patterns. The authors employed the CH-model of Camerer et al. (2004) to classify behavior.

Devetag et al. (2016, p. 198) report a correlation between eye movement patterns of visual information acquisition and subject choices. This correlation implies that behavior types that differ in their ability to develop a sophisticated decision use different information acquisition patterns. If a direct link between patterns and underlying information processing exists, this correlation hints to the use of different EIPs or sequences of EIPs within the process.

Both studies show that a move from one cell to another can be interpreted as part of patterns regarding information acquisition and information processing. One can conclude on the particular pattern by considering context information, such as the payoff inspected and next cell visited. The remaining EIPs listed in Table 12 can now be derived from EMMAs combined with movement patterns and context information.

8.3.2.3.3 READ, COMPARE, ADD & STORE

READ is per definition used to store information into STM. There is a certain redundancy with STORE that also saves information in STM. The difference here is the origin of the information which is to store. Whereas READ acquires information from outside the cognitive system (i.e., from the screen showing payoff matrix), STORE saves meta-information resulting from prior operations (e.g., cumulating a sum, result from a comparison). Nevertheless, both ways lead to data, which are obtainable in STM for subsequent operations. Obviously, READ offers the information which is to be processed and hence needs to begin the cognitive information process. Following this, READ is always a predecessor of ADD, COMPERE, and STORE. Herewith, it is arguable that READ is always part of a sequence of EIPs causing a halt which lasts longer than the above-discussed threshold (about 180 ms to 200 ms). With this, the EIP links to EMMAs through the following statement:

IX HALT (200 ms + x > halt >= 180 ms) =: READ (focussed cell content into STM).

To make a comparison, one needs at least two units of comparable information. Such a comparison is usually made pairwise (Russo and Dosher 1983).[115] The sequence of visited information within the mouse move data stream provides the context relevant for interpretation. If a subject wants to compare two of the own strategies, one could test for dominance. In this case, the subject is likely to compare the payoffs of the two own strategies for each (at least two) of the opponent's strategies. However, this approach assumes the somewhat rational behavior of the subject.

Here, the expected move pattern would exhibit a sequence of horizontal moves between two cells of two strategies. A halt that lasts longer than the READ threshold follows each move. The focus lies predominantly on the own payoff. Vertical or diagonal moves to the next pair of cells from the same two strategies succeed the pairwise comparison.

A subject needs to store the results from each of the pairwise payoff comparisons for all of the opponent’s alternatives in STM. The probability of forgetting one such meta-information is thus growing with a rising number of alternatives, but decreasing with growing costs of information acquisition. Hence, repeated comparisons are likely to occur in the former case, which in turn supports the interpretation of the movement pattern. In case of comparing the opponent's strategies, one could expect a pattern derived analog.

Now take the case of searching for the highest payoff, respectively the lowest payoff of a strategy. The pattern changes since a subject only inspects a single strategy, probably systematically. People with a western dominated direction of reading and writing would very likely begin with a systematic inspection of the first cell on the top left and follow to the right (Abed 1991, p. 531). However, one can identify such a pattern without strictly obeying this systematic. Again, the probability to inspect systematic search behavior grows with rising number of alternatives: it is generally easier to remember which cell is monitored already and which remains unseen when

[115] Also compare footnote 114, p. 131 in this treatise.

applying a systematic information search concept in comparison to an unstructured search. With a growing number of alternatives, the number of cells increases exponentially. In an unstructured search, this also increases the effort to store the information of the cells already monitored. To avoid high efforts and to spend fewer resources, subjects likely tend to apply a systematic search where the storage of the last visited cell and the search direction is sufficient. Even if there is no such systematics, a subject will look up all cells one after another before coming to a solution or revisiting another cell.[116] So, in that case, identifying the pattern solely requires a visitation of all cells, one after another and with revisiting just a few (theoretically zero) cells. Observing a systematic behavior would clearly support this interpretation. Hereby, COMPARE is expected to show two distinctive patterns which read as follows:

X.I READ (one cell's focussed payoff)
+ MOVE (vertically (resp. horizontally) to next cell of different strategy)
+ HALT (> HALT $_{\text{READ}}$, same payoff focus)
+ MOVE (to next cell of one of the inspected strategies (could be either horizontally or vertically or horizontally + vertically))
AND repetition with same strategies, but different cells in adjacent move history
=: COMPARE I (two strategies (search for dominance)).

X.II READ (one cell's focussed payoff)
+ MOVE (systematically) to next unvisited cell (of strategy or of whole matrix)
+ HALT (> HALT $_{\text{READ}}$, same payoff focus)
AND Repetition in adjacent move history until all cells visited
AND almost no cells revisited in adjacent move history
=: COMPARE II (payoffs to identify maximum (resp. minimum)).

In analogy to COMPARE the EIP ADD can be described as an operation which needs two units of information for execution. The first one is the current sum, stored in STM, and the second one is the payoff from the last visited cell. Whereas COMPARE II (X.II) aims to find the maximum (resp. minimum) payoff for one player within the whole payoff matrix, ADD is applied in two cases.

[116] Note that a subject needs to store two units of information in her STM: the current maximum (resp. current minimum) and the payoff from the last cell visited. Hence, it is unlikely to forget the maximum (minimum), and so there is no need to go back and revisit cells.

In the first case, a subject might play the *Altruism* heuristic and hence searches the combined maximum of own and opponent's payoff. At this point, the focus changes within a cell after reading one player's focus. With the same argumentation as above subjects with a westerly reading and writing education start with reading the own payoff information since it is the first entry from left (compare Figure 21). Although the reading direction plays a subordinate role in this interpretation, one might expect it to be considered repeatedly when visiting cells. Context information is thus necessary for its identification.

It is worth noting that a mouse move within a cell is occurring more likely when the payoff sums in the matrix are more homogenous. The more the payoff sums differ, the easier it is for a subject to identify the maximum. In that case, a mouse move is expected to be kept short to save resources. However, movement patterns are expected to show two reading halts per visited cell. The first halt is shorter (just READ) than the second (READ + ADD). One move between the players' payoff attention areas accompanies both halts (see Figure 22, p. 124). Additionally, the movement trajectory exhibits that this mouse movement behavior is repeated for all cells of the payoff matrix.

In the second case subjects (implicitly or unconsciously) equally weight the entry possibility of an opponent's strategy choice. The best answer is identified by calculating the maximum payoff sum of own alternatives (see description of heuristic *Naïve*, Section 4.2). Corresponding movement trajectories show a more or less systematic search for each of the subject's alternatives. When reading a cell's entry, the own payoff is in focus exclusively. Also, halts are longer than the reading threshold, since adding is combined with information acquisition here. The movement pattern then shows reading of own payoff in the first visited cell and a horizontal move to a next cell in the same strategy with same focus. Movement trajectories show that this mouse movement behavior is repeated for a whole strategy and probably for all of a subject's strategies. If a subject applies heuristic *L2*, the focus lies on opponent's payoffs and alternatives first, since *L2* is the best answer to *Naïve*. The movement pattern for ADD then can be derived in analogy to the aforementioned procedure. Herewith, the two patterns read as follows:

XI.I READ (one cell's focussed payoff)
+ MOVE (horizontally in same cell)
+ HALT (> HALT $_{\text{READ}}$, other payoff focus)
AND repetition with different cells in adjacent move history

AND repetition with all cells in whole move history

=: ADD I (payoffs information of one cell)

XI.II READ (one cell's focussed payoff)

+ MOVE (horizontally (resp. vertically) to next cell of same strategy)

+ HALT (> HALT $_{\text{READ}}$, same payoff focus)

AND repetition with same strategy, but different cells in adjacent move history

AND repetition with all strategies in whole move history

=: ADD II (payoff information of one player's strategy)

The last EIP to be discussed here is STORE. As mentioned above, it is expected to occur together with ADD and COMPARE – operations which need their results to be stored for later use. Since this is the case when the calculation part ends, STORE finishes the halt process (as last part of the eye fixation) and usually a mouse move follows. Experimental studies done to determine the time needed to store information in STM or 'working memory' are rare (see Russo (1978) for a prominent exception). As a result, no reference value can be found in literature for the time needed to store meta-information of the kind mentioned earlier. However, it can be argued that STORE is always a successor of READ, ADD and COMPARE since it stores the results of the operations COMPARE and ADD. Hence, it could be concluded that the relation 'after ADD or COMPARE follows STORE' is reasonable. With this, the connection between EIP and EMMAs reads as follows:

XII (after) ADD or (after) COMPARE =: STORE

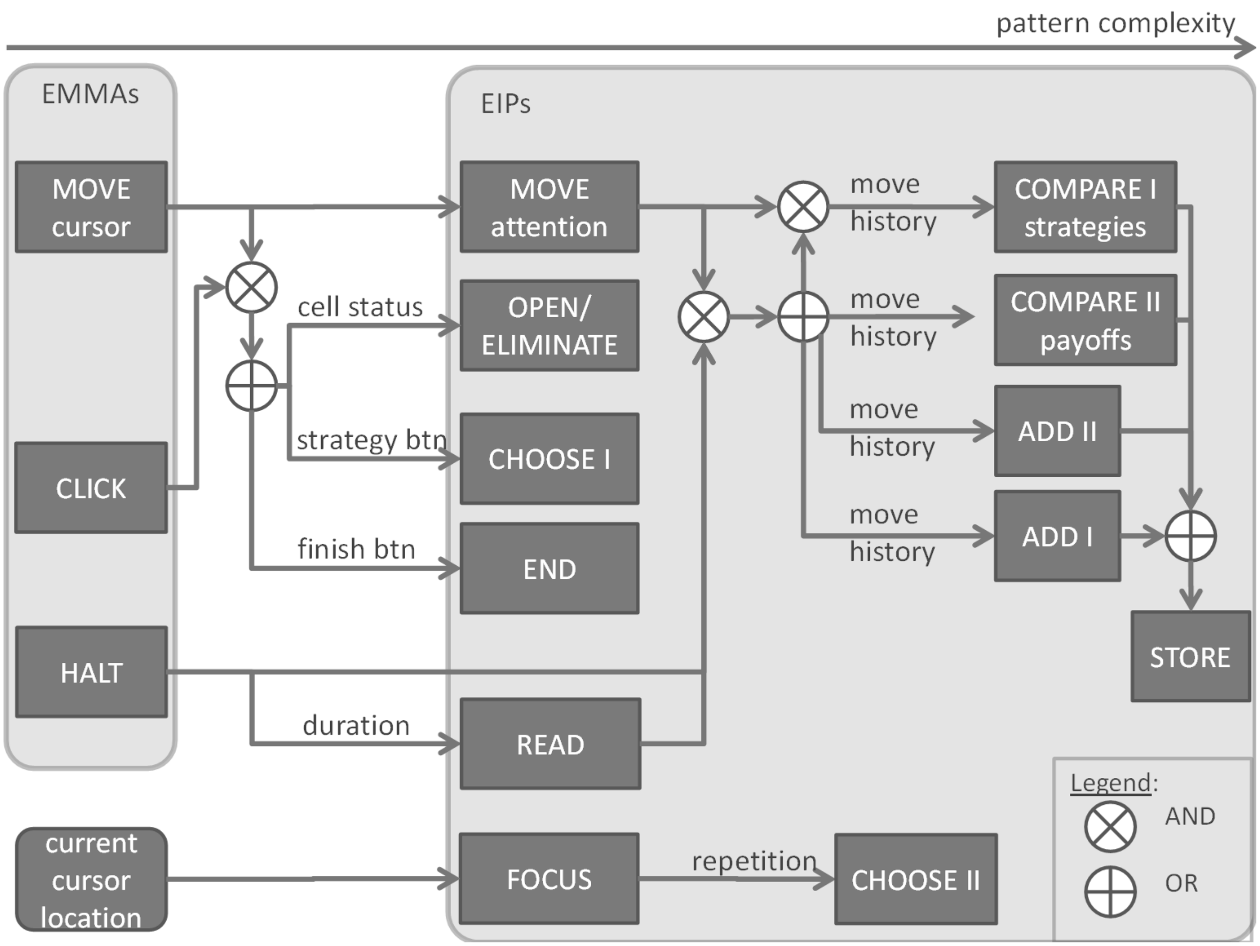
Definition XII concludes the interpretation metric. The minimum set of EIPs is completely matched with EMMAs. It can henceforth serve the analysis of mouse movement data regarding a subject's strategic behavior.

The matching process works without any decision information so far. It is worth to mention that even though the interpretation matrix does not consider this information, it can principally add to the analysis. Every decision has the potential to be the outcome of reasonable information acquisition and processing. Of course, this requires normal-form games where the choice of alternatives can be related to a normative level of strategic sophistication. In cases where subjects apply behavioral patterns of a particular strategic level decisions that match that level contribute to the interpretation of behavior (in analogy to Costa-Gomes et al. (2001)). Those patterns are similar to the (part of the) corresponding production systems.

Some of the assumptions presented to match EIPs with EMMAs are not formulated precisely so far but rather general. That concerns especially duration times and the number of reoccurring behavioral patterns. However, with little effort, this uncertainty can be solved by parameterizing the open aspects.

Figure 23 gives an overview of the interpretation metric's systematics on the base of the definitions (I–XII) given above. EMMAs, depicted in the left of the graphic, are used in specific combinations to identify EIPs (depicted right). Arrows and logical operations (AND, OR) represent such combinations. The arrows inscriptions contain the most critical context information, required to identify the EIP.

FIGURE 23 – SYSTEMATICS OF INTERPRETATION METRIC



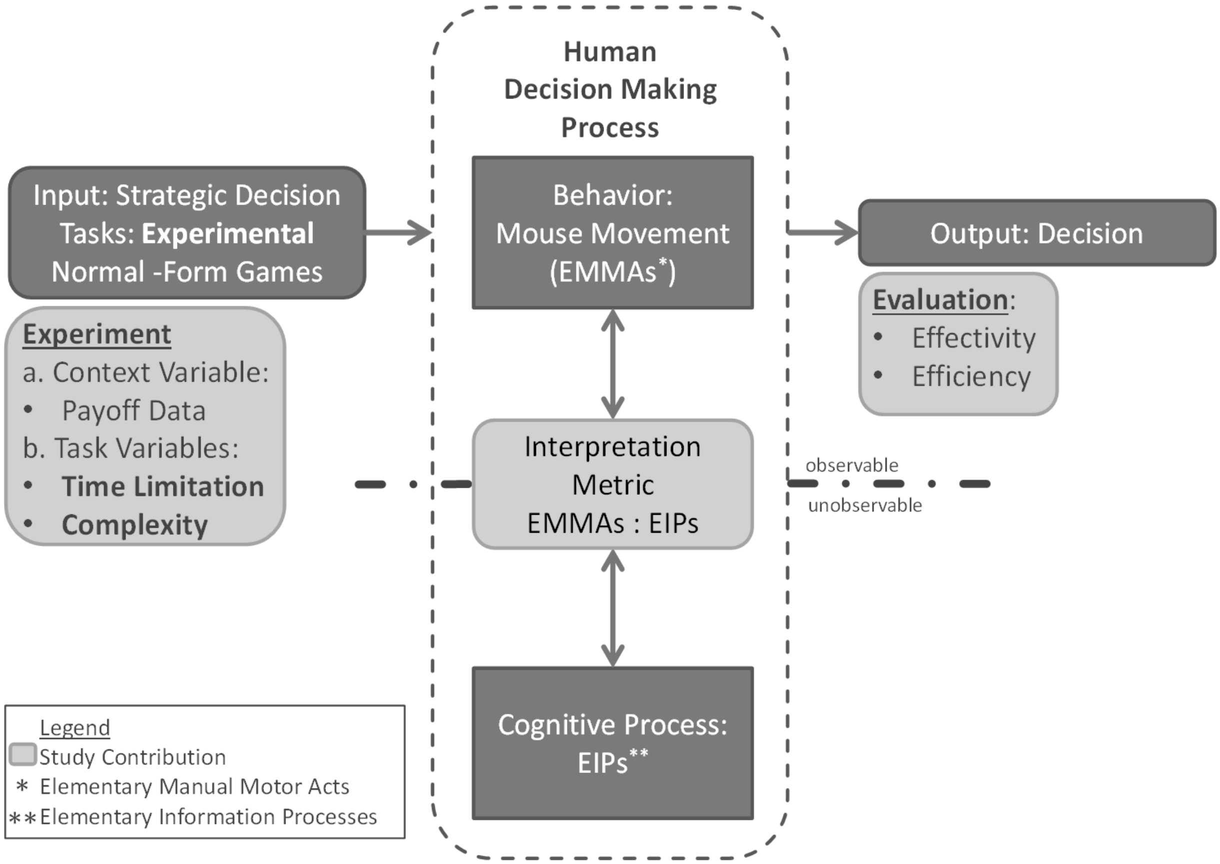
Some of the EIPs require a combination of other EIPs and EMMAs for identification. They thus reveal more complex patterns of behavior. The pattern complexity is increasing from left to right within Figure 23. The 'current cursor location' (lower left corner in the graphic) is not an EMMA in the sense of a manual motor act, but necessary to determine the focus of attention. It is thus included here. With this interpretation metric on hand, it is possible to code the proposed systematics and hence automate data analysis.

In this chapter, the author selected the PTM ‘Mouse tracking' as the technological base for developing the interpretation metric. Both, technique and metric are part of the experimental approach. They are employed to acquire information and to interpret it. The results of Chapter 8 answer at the same time how one can describe, explain and determine the influence of time pressure on the cognitive decision-making process (see the first research question of this treatise, Section 1.2).

9 Experimental Procedure

This chapter is dedicated to the design and execution of an online controlled behavioral experiment. Figure 24 illustrates the experimental approach.

FIGURE 24 – EXPERIMENTAL APPROACH



The primary goal of this decision-making experiment is to acquire comprehensive data to evaluate the implications of the preparation time model. This requires studying cognitive processes in decision-making under time pressure in general and identifying behavioral patterns in particular. To gather this kind of information experimentally and link it to the preparation time model, several conditions necessarily need to be met. It starts with the analysis concept that specifies which kind of data (quality and quantity) is needed to answer the main questions of this study and in which form the data can be processed. This has practical implications for the experimental design, including tasks, subjects, incentives and presentation. That in turn influences the technical implementation of the experimental concept. The design of the experiment thus follows consistent and logically sequencing arguments, inducing a structured approach.

Henceforth the following schedule is suggested:

1. Summarize key questions of this study and formulate relating hypotheses.

2. Specify data necessary to answer key questions, including statistical hypothesis testing.

3. Derive experimental design from specified data.

4. Determine technical implementation from experimental design.

With this approach, the author adapts the general design of (statistical) experiments which emphasizes the three consecutive steps of planning, execution, and analysis.[117] Time limitation and complexity serve as independent task variables. Their influence on the decision-making process in strategic tasks is the subject of the ongoing study.

9.1 Key Questions and Hypotheses

The data acquired within the experiment shall answer the remaining research questions of this treatise. One can specify for this experiment (the bold roman number in parenthesis refers to the notion of Section 1.2):

- What behavioral patterns can be observed during the process of decision-making and how can they be classified **(III)**?
- What is the quantitative effect of time limitation and task complexity on the effectiveness and efficiency of the decision-making process in a strategic decision environment **(IV)**?
- How does time pressure affect behavior on a process-component-level and concerning patterns (**IV**)?

9.1.1 Assumptions

The approach to answering the fundamental questions is based on four assumptions that directly influence the method of examination. They arise from particular process-oriented studies and already find consideration within the earlier introduced preparation time model (Chapter 3). The results of the preparation time model developed in the simulation part of this treatise also find consideration. The assumptions read as follows:

[117] The design of experiments is a standard procedure in various fields of science and hence surely belongs to the standard repertoire of empirical research. A simple introduction offer, for example, Vogt (1988) or Fisz (1989), the first mentioned for a more technical purpose. Kagel and Roth (1995) follows a more problem-centered orientation in their introduction with emphasis on the field of experimental economics.

1. Decision-making under time pressure in strategic tasks follows a cognitive process – conscious or unconscious. That process links to observable behavior as a sequence of motor acts (Miller 1960, p. 697; Simon and Newell 1971, p. 146; Kühberger et al. 2011, p. 2).

2. The cognitive process can be sequenced employing EIPs (as proposed by Johnson and Payne (1985, p. 398) and Huber (1980, p. 188) and confirmed by Payne et al. (1988), Gertzen (1992), Costa-Gomes et al. (2001), among others).

3. EIPs can be traced by the EMMA metric as presented in Subsection 8.3.2.

4. The performance concept proposed in Section 5.1, is a suitable method for estimating effectiveness and efficiency of choice behavior.

Those assumptions are widely confirmed in the related studies named above, or at least seem reasonable and appropriate in the context of this work. The following testable hypotheses, which determine the required dataset, specify the fundamental questions.

9.1.2 Hypotheses

Relevant hypotheses can be developed based on the research questions stated above and in connection with results of the preparation time model. Their classification is twofold: the first part of hypotheses deal with the general performance of behavior under time pressure, using the evaluation concept introduced in Section 5.1. The hypotheses' descriptions are listed in the order of their contribution to an individual goal. This part is followed by hypotheses that generally relate to the application of problem-solving methods (PSMs, e.g. heuristics) and to patterns of behavior. Table 21 provides an overview of the reviewed hypotheses.

In compliance with standard statistical procedures, the following hypotheses and their formulations are translated into a form of null hypothesis H_0 and alternative hypothesis H_1, referring to a particular population parameter that is in the center of the claim. The hypothesis H_0 in this context is stating the assumingly true status quo in behavioral research whereas H_1 depicts the contradictory ideas that follow the preliminary analyses. As the cutoff condition to reject an H_0-hypothesis (‘α-level’), the frequently employed value of α = 0.05 is used. The p-value of the test statistic regarding the underlying sample distribution is given additionally.

TABLE 21 – HYPOTHESES AND CONTRIBUTIONS

Contribution→ **Hypotheses↓**	**Goal** **1**	**2**	**3**	**4**	**Heuristics/ patterns**
I General sensitivity	x	x	x	x	
II Qualitative sensitivity effectiveness		x	x		
III Reduction of alternatives		x			
IV Payoff strategic PSM			x		x
V Payoff nonstrategic PSM			x		x
VI Qualitative sensitivity efficiency				x	
VII Application of heuristics					x
VIII Heuristics' completeness					x
IX Application of *Random*					x
X *Equilibrium* choice					x

9.1.2.1 Hypothesis I: General Sensitivity to Time Pressure

Several decision-making studies that contain a time dimension of limiting the task fulfilling procedure identify a specific reaction in behavior. Similar effects are expectable in a strategic task environment. Especially with studies by Sutter et al. (2003) and Lindner and Sutter (2013) this sensibility is proved for strategic tasks other than normal-form games. Since normal-form games are not subject to time pressure investigations yet, the conclusion seems reasonable but needs an explicit verification. This consideration leads to the first and fundamental hypothesis for the goals of the present research:

I. Decision-making in the strategic task environments of normal-form games is affected by time pressure[118].

Decision-making is evaluated by the performance concept presented in Section 4.1.[119] This concept includes providing a choice (Goal 1), eliminating alternatives (Goal 2), payoff generation (Goal 3), and the sum of EIPs (Goal 4). Goals 1 to 3 contribute to effectiveness while Goal 4 rep-

[118] The presented hypotheses all focus on time pressure conditions in strategic task environments of normal-form games. Time pressure is consequently modeled by employing time limitation and complexity – as suggested in Part II and described on various other occasions in this treatise. Both facts – task environment and time pressure – are thus not stated again explicitly in the following hypotheses' formulation, but find implicit consideration.
[119] Table 13 provides an overview of the evaluation concept in use.

resents the efficiency aspect. It is sufficient to show the sensitivity of at least one component to prove overall sensitivity. The four aspects are subject to following hypotheses that concretize the general sensitivity assumption by adding an orientation to the sensitivity (positive or negative). A conclusion is drawn inductively: if the effectiveness of time pressure can be determined for one of the four goals, it would prove the general sensitivity at the same time. This general sensibility would be independent of any interference between the effects of the goals. Constellations where effects cancel each other out would, however, require further consideration. Statistically, one can interpret this statement as follows:

The decision-making behavior with respect to a specific goal g_i in a task with a certain parameter constellation satisfies a probability distribution with the mean value μ. Given a parameter constellation with experimental design V_e and with a fixed complexity c, a change in the time limit t should then significantly change the distribution mean according to the following hypothesis for at least one of the four goals g_1 to g_4, as long as a time-related sensitivity exists:

H_1: "$\mu(V_e, g_i, t_1, c) \neq \mu(V_e, g_i, t_2, c)$", testing the corresponding null hypothesis

(I.a) H_0: "$\mu(V_e, g_i, t_1, c) = \mu(V_e, g_i, t_2, c)$" with $t_1 > t_2$ and $c = const.$

The similar procedure is applied to formulate the hypothesis when evaluating the influence of complexity c, with the value of c varying:

H_1: "$\mu(V_e, g_i, t_1, c_1) \neq \mu(V_e, g_i, t_2, c_2)$" and

(I.b) H_0: "$\mu(V_e, g_i, t_1, c_1) = \mu(V_e, g_i, t_2, c_2)$" with $c_1 < c_2$ and $\frac{t_i}{c_i^2} \approx const.$

Note that the time limit t is not constant for tasks where the complexity is varying. Instead, when modeling the time pressure an attempt was made to keep the 'time per payoff matrix ment' t_i/c_i^2 constant for a particular stage of time limitation.[120] This circumstance holds for all of the following hypotheses even though it is not explicitly mentioned again.

9.1.2.2 Hypothesis II: Qualitative Sensitivity of Effectiveness to Time Pressure

The first hypothesis can be specified adding a qualitative direction and quantitative aspects which are suggested by various studies on time pressure (Zakay and Wooler 1984, p. 279; Payne et al.

[120] As will be argued later, the task structure is similar to the simulation approach in Part II. The task values of time limits are determined considering multiple influences. One of them is the intention to keep that rate constant within the four rounds and among the four stages of time limitation.

1988, p. 579; Zakay 1993, pp. 68 f.; Johnson et al. 1993, p. 115; Ariely and Zakay 2001, p. 204). At this point, performances concerning effectiveness and efficiency are expected to differ in their quality of sensitivity. Hence, separate hypotheses represent their performance sensitivity. Effectiveness is the subject of the second to fifth hypotheses, while the sixth addresses efficiency. Looking at the counterintuitive results of Lindner and Sutter (2013, p. 544), the second hypothesis carries a realistic chance of being falsified. The simulation results presented in this treatise heads in the opposite direction. Hereby, the second hypothesis reads as follows:

II. Decision-making is negatively affected by time pressure regarding effectiveness.

The related hypotheses are fairly similar to the ones of Hypothesis I but now gain a singular orientation, also affecting the null hypothesis H_0:

H_1: "$\mu(V_e, g_i, t_1, c) > \mu(V_e, g_i, t_2, c)$", testing the corresponding null hypothesis

(II.a) H_0: "$\mu(V_e, g_i, t_1, c) \leq \mu(V_e, g_i, t_2, c)$" with $t_1 > t_2$ and $c = const.$

Again, μ is the mean of the underlying distribution and $t_1 > t_2$. The corresponding hypothesis of complexity reads as follows:

H_1: "$\mu(V_e, g_i, t_1, c_1) > \mu(V_e, g_i, t_2, c_2)$" and

(II.b) H_0: "$\mu(V_e, g_i, t_1, c_1) \leq \mu(V_e, g_i, t_2, c_2)$" with $c_1 < c_2$ and $\frac{t_i}{c_i^2} \approx const.$

There is no simple line of argumentation to state the connection between growing complexity and payoff development or the ability to reduce alternatives. Complexity is thus assumed to show an ambivalent influence on effectiveness.

9.1.2.3 Hypothesis III: Reduction of Alternatives under Time Pressure

The following three hypotheses are directly deduced from the simulation results stated in Part II. They refer to quantitative effects of time pressure regarding effectiveness and efficiency. Both terms are addressed in the simulation to measure the performance of the heuristics.[121] The orientation of the time limit's influence tends to depend on the underlying heuristic, making this issue an additional indicator for characterizing decision-making behavior itself. Focussing on the indi-

[121] As will be discussed later, the same concept of performance is applied in the experimental part of this treatise, too.

vidual goals that contribute to effectiveness, at least the second one, reduce alternatives, seems to allow a general hypothesis according to findings from simulation:

III. Effectiveness, regarding the ability to reduce alternatives, is decreasing with shrinking time restrictions and increasing with growing complexity.

Let $\mu(V_e, g_2, t_i, c_i)$ be the mean of the underlying distribution of the realization of Goal 2 (Reduce alternatives) under experimental setting V_e and with time limit t_i and complexity c_i in task i. Hypothesis VI can then be evaluated with:

H_1: "$\mu(V_e, g_2, t_1, c) > \mu(V_e, g_2, t_2, c)$", testing the corresponding null hypothesis

(III.a) H_0: "$\mu(V_e, g_2, t_1, c) \leq \mu(V_e, g_2, t_2, c)$" with $t_1 > t_2$ and $c = const$ for the tasks 1 and 2.

The corresponding hypothesis regarding complexity reads as:

H_1: "$\mu(V_e, g_2, t_1, c_1) < \mu(V_e, g_2, t_2, c_2)$" and

(III.b) H_0: "$\mu(V_e, g_2, t_1, c_1) \geq \mu(V_e, g_2, t_2, c_2)$" with $c_1 < c_2$ and $\frac{t_i}{c_i^2} \approx const.$

9.1.2.4 Hypotheses IV and V: Payoff-Generation under Time Pressure

The following two hypotheses focus on the ability to generate payoff. EIP-intensive problem-solving strategies in human decision-making seem to collectively suffer from time pressure (Hypothesis IV), while the less-EIP-intensive ones perform stable or even better under such conditions (Hypothesis V):

IV. For EIP-intensive solution concepts (such as heuristics *D1*, *D2*, and *Equilibrium*) effectiveness concerning generating payoff is decreasing.

Significant for those EIP-intensive heuristics is the occurrence of the strategic EIP 'COMPARE I' that hints to the comparison of adjacent matrix cells of different alternatives. This pattern is unique in the solution concept of identifying dominated alternatives. No other heuristics from the examined set use it in their production systems. It is now of particular interest to relate the numbers of this EIP to the generated payoff. Hypothesis IV implies that with growing time pressure the ratio of generated payoff per number of occurring COMPARE I in a player's EIP sequence is decreasing. This consideration is hence limited to datasets of participants where the number of the EIP COMPARE I is larger zero. Since the payoff plays an essential role in the effectiveness calculation, it seems adequate to transform it into a non-negative range. This standard range allows for an inter-subject comparison of the results (Coombs et al. 1975, p. 35). The normalization is realized by putting each player's absolute result per task in context to the minimal and maximal achievable outcome of this task using the following formula (Eq. (6)):

$$u_{rel,i} = \frac{u_i - u_{min}}{u_{max} - u_{min}}, \tag{6}$$

with u_i as the realized outcome of player i in a given task, u_{min} and u_{max} as minimum and maximum payoff of the task and $u_{rel,i}$ as the relative outcome of player i. $u_{rel,i}$ is a real number, bounded to the interval [0,1].

Let $\mu^*(V_e, g_3, t_i, c_i)$ be the mean of the underlying distribution of the realization of Goal 3 (Maximize payoff) per number of EIP COMPARE I under experimental setting V_e, with time limit t_i and complexity c_i in task i. The asterisk (*) depicts the limitation of the considered dataset to EIP sequences where the number of COMPARE I is larger zero, as mentioned above. One can now test Hypothesis IV with:

H_1: "$\mu^*(V_e, g_3, t_1, c) > \mu^*(V_e, g_3, t_2, c)$", testing the corresponding null hypothesis

(IV.a) H_0: "$\mu^*(V_e, g_3, t_1, c) \leq \mu^*(V_e, g_3, t_2, c)$", with $t_1 > t_2$ and $c = const$ for tasks 1 and 2.

The corresponding hypothesis regarding complexity reads as:

H_1: "$\mu^*(V_e, g_3, t_1, c_1) > \mu^*(V_e, g_3, t_2, c_2)$ and

(IV.b) H_0: "$\mu^*(V_e, g_3, t_1, c_1) \leq \mu^*(V_e, g_3, t_2, c_2)$", with $c_1 < c_2$ and $\frac{t_i}{c_i^2} \approx const.$

Hypothesis V is formulated in analogy to Hypothesis IV:

V. For non-EIP-intensive solution concepts (such as nonstrategic heuristics and *L2*) effectiveness in terms of generating payoff is at least constant or increasing and they hence find more application under growing time pressure.

The identification of those non-EIP-intensive solution concepts is straightforward, using the complement strategies to the ones of Hypothesis IV. Herewith, all valid datasets where the EIP COMPARE I does not occur in the EIP sequence are incorporated (depicted by a tilde sign (~)). In analogy to Hypothesis IV, the test definition is:

H_1: "$\mu^\sim(V_e, g_3, t_1, c) \leq \mu^\sim(V_e, g_3, t_2, c)$, testing the corresponding null hypothesis

(V.a) H_0: "$\mu^{\sim}(V_e, g_3, t_1, c) > \mu^{\sim}(V_e, g_3, t_2, c)$", with $t_1 > t_2$ and $c = const$ for the tasks 1 and 2.

The corresponding hypothesis regarding complexity reads as:

H_1: "$\mu^{\sim}(V_e, g_3, t_1, c_1) \leq \mu^{\sim}(V_e, g_3, t_2, c_2)$ and

(V.b) H_0: "$\mu^{\sim}(V_e, g_3, t_1, c_1) > \mu^{\sim}(V_e, g_3, t_2, c_2)$", with $c_1 < c_2$ and $\frac{t_i}{c_i^2} \approx const$.

9.1.2.5 Hypothesis VI: Qualitative Sensitivity of Efficiency to Time Pressure

Following the sequence of the performance concept presented in Section 5.1, Goal 4 (Minimize the number of EIPs) contributes to efficiency. A general and intuitive assumption about the relationship with time pressure is that the shorter the time limit, the lower the number of EIPs. Increasing the complexity and thus increasing the amount of potential information to gather could principally have the opposite effect on the number of EIPs. A hypothesis needs to consider both aspects:

VI. Efficiency is increasing with decreasing time limit and decreasing with increasing complexity.

Efficiency is evaluated as numbers of applied EIPs while fulfilling the experimental task. The more EIPs are applied, the less efficient is the task fulfillment. Demanding a choice as a necessary condition secures the comparability between datasets of players at this point. This condition ensures concordance with the underlying goal hierarchy as presented in Section 5.1. The hypothesis is thus limited to the group of subjects, who draw a decision. To test this hypothesis, let $\mu(V_e, g_4, t_i, c_i)$ be the mean of the underlying distribution of the numbers of EIPs under experimental setting V_e, with time limit t_i and complexity c_i in task i. One can test Hypothesis VI with:

H_1: "$\mu(V_e, g_4, t_1, c) > \mu(V_e, g_4, t_2, c)$", testing the corresponding null hypothesis

(VI.a) H_0: "$\mu(V_e, g_4, t_1, c) \leq \mu(V_e, g_4, t_2, c)$", with $t_1 > t_2$ and $c = const$ for the tasks 1 and 2.

The corresponding hypothesis regarding complexity reads as:

H_1: "$\mu(V_e, g_4, t_1, c_1) < \mu(V_e, g_4, t_2, c_2)$" and

(VI.b) H_0: "$\mu(V_e, g_4, t_1, c_1) \geq \mu(V_e, g_4, t_2, c_2)$", with $c_1 < c_2$ and $\frac{t_i}{c_i^2} \approx const$.

Especially in this case of hypothesis testing, the influence of time limitation and complexity are supposedly clearly identifiable and hence thoroughly to be examined. When the time limit is decreasing, and complexity is increasing at the same time, the contrarily working, reciprocal impact of both can be evaluated directly.[122]

9.1.2.6 Hypothesis VII: Application of Heuristics

This hypothesis follows proposals of Bettman (1979, p. 176) emphasizing the general use of heuristics – whatever quantity and comprehension of application look like in the concrete case. Bettman principally questions the use of complete heuristics during the decision-making process and recommends parts to be applied instead (Bettman 1979, p. 33). Others, namely Payne et al. (1988) for a nonstrategic and Costa-Gomes et al. (2001) for a strategic task environment, find evidence confirming the application of heuristics. Both lines of argumentation assume that at least parts of heuristics find application in the decision-making process. Because of their analysis method, the studies mentioned implicitly postulate that heuristics are permanently in use.[123] The mere existence can be assumed mainly since some people know strategies of decision-making from their education and are thus likely to apply them to problem tasks. This fact generally comprises common heuristics, too. However, Johnson et al. (2008, p. 270) report from their gambling experiment that subjects tend to employ heuristics that differ across individuals and tasks. Whether the actual frequency of decision-making with the aid of such heuristics or selected parts among the population is almost one hundred percent in the task context of this experiment, shall be answered when testing the following hypothesis:

VII. Decision makers employ one or more heuristics, or one or more parts of one or more heuristics while fulfilling a predefined strategic task – regardless of time pressure conditions.[124]

[122] The other way around would also be working, but the described one is the anticipated working direction of the experimental task (from easy (i.e., low time pressure) to hard (i.e., high time pressure)).
[123] Costa-Gomes et al. (2001, p. 1204) for example compare participants' behavior with production systems of heuristics and give a compliance rate. Since there is at minimum one heuristic per participant showing a rate larger than zero percent, the authors infer that every decision maker employs at least one heuristic or part of it. With even a smaller set of heuristics considered in their study, Payne et al. (1988, pp. 14 f.) categorize all participants' behavior in alternative-focussed and attribute-focussed.
[124] This hypothesis might be falsified because of one or more participants do not employ the set of common heuristics. However, the minimum set of EIPs (Chapter 3) is not falsified at the same time, since this model employs single EIPs rather than heuristics or part of heuristics.

The evaluation of the concrete frequency is of further interest for researchers in general and especially in this study. It is thus presented along with a confidence interval in the analysis part of this paper.

With the interpretation metric presented in Chapter 8, it is possible to identify complete heuristics as well as significant parts of their patterns within the EIP sequence. Of course, the smaller the parts of the patterns, the more difficult is a precise mapping to the potentially used heuristics. It hence seems appropriate to study and interpret the realized EIP sequence completely in such cases.

When testing this hypothesis, it is moreover necessary to define the term ‘part of a heuristic pattern’. Whenever one heuristic is fully employed, also parts are deployed. That is intuitive. If one defines ‘part of a heuristic pattern’ on a scale of a single EIP, this hypothesis will be confirmed every time a participant at least moves the mouse once by definition.[125] In that case, it would not be very discriminatory.

A frequently used alternative is to estimate the percentage of concordance between a subject’s behavior and heuristics’ procedures (Payne et al. 1988; Johnson and Payne 1985; Costa-Gomes et al. 2001). The authors record the number of predefined types of mouse movements. Estimation is thus based on plain numbers rather than specific sequences, although especially Costa-Gomes et al. (2001) also record and interpret the repetition of specific moves to argumentatively strengthen their classification of behavior.

The simulation part of this treatise provides the required heuristics' data. Since the simulation comprises data of 466 time limits per level of complexity, it is suggested to use representations of the four-time stages ‘early’, ‘middle’, ‘late’ and ‘constant’ of each heuristic instead. However, one must determine a measure to verify the use of at least parts of the heuristics during decision-making. The above studies do not provide such a measure.[126]

A proper way is to link the definition of ‘part of heuristic pattern’ to certain sequences of EIPs that characterize a pattern. Some of the EIPs already represent such complex patterns themselves. Namely, COMPARE I and II and ADD I and II form information search and processing patterns that induce goal-orientation and organized proceeding. They are characteristic for certain heuris-

[125] If the mouse is moving, at least the EIP MOVE is recorded.
[126] This circumstance is not a drawback for these studies, since determining such criterion is not part of their concepts.

tics, even if not discriminating a single one in every case. Of course, other patterns from the heuristics' production systems are thinkable, too. Same as the four EIPs mentioned above, sequences of EIPs would form those patterns. However, they are both lacking the interpretative value in identifying systematic information search and processing.[127] Alternatively, they contain the four EIPs mentioned above themselves, unnecessarily increasing the complexity of a pattern here. Following this, their interpretative value in the context of the present examination is to be interpreted at least as critical. Relying on the identified EIPs hence seems appropriate.

The mere existence of the four pattern-EIPs can further serve as a criterion. It would then be sufficient to record them in the behavioral data stream to evaluate Hypothesis VII.[128] The convincing advantage of employing the four EIPs mentioned above is that they find application in almost all heuristics of the set examined in this study. The only exception is *Random*, which itself is rather easy to identify from the decision-making data.[129]

Now, one can study every decision by checking for the existence of either a random choice or of one of the four EIPs mentioned above. With this, a promising and comprehensible approach is taken to evaluate Hypothesis VII with the following proportion test:

$H_{1,i}$: "$p_1(V_e) < 1$", testing the corresponding null hypothesis

(VII.a) $H_{0,i}$: "$p_1(V_e) >= 1$",

with p being the proportion of people that use at least parts of the given set of heuristics in experimental setting V_e (among all tasks). Note that participants who do not use the mouse pointer at all are excluded in this consideration since their behavior data provide no analyzable information.

9.1.2.7 Hypothesis VIII: Completeness of Heuristics' Application

The following hypothesis links the probability of using at least parts of common heuristics to time pressure. If this hypothesis holds, it is reasonable to suggest that more time to think about a

[127] The underlying mouse movement could either belong to irrational and unintended behavior, too. Alternatively, the information search is slightly unorganized and chaotic in the meaning here and hence is the processing no longer directly relatable to the recorded EMMAs.
[128] Note that the identification of those EIPs is susceptible to unintended random mouse movement and clicking, even if its rather complex patterns generally minimize, not to say prevent such recordings. By the way, the same problem exists with the concepts of the cited studies where movement patterns are even less complex.
[129] Whenever a subject makes a decision without having included at all information given, it is certain that the choice is based on chance. That explicitly includes the case a player starts to gather (and process) information but cancels task processing underway for any reason.

task could result in an increased likeliness to invest some thoughts in problem-solving methodology. One can formulate the corresponding hypothesis as follows:

VIII. Heuristics are likely to be deployed more entirely under less time pressure.

This hypothesis assumes that the more time a participant has to fulfill a task, the more patterns of one or more heuristics are used and hence can be detected in the data stream. Hereby, the seventh hypothesis gains a quantitative aspect when considering the time pressure effect. The quantity is evaluated by the number of the four EIP-patterns that are relevant in Hypothesis VII, too. Similar to Hypotheses I and II, this effect is analyzed via its components time limitation and complexity.

Let $\mu(V_e, t_i, c_i)$ be again the mean of the underlying distribution of the sum of the four EIPs ADD I, ADD II, COMPARE I and COMPARE II in experimental setting V_e, time limit t_i and complexity c_i in task i. One can test Hypothesis VIII with:

H_1: "$\mu(V_e, t_1, c) > \mu(V_e, t_2, c)$", testing the corresponding null hypothesis

(VIII.a) H_0: "$\mu(V_e, t_1, c) \leq \mu(V_e, t_2, c)$", with $t_1 > t_2$ and $c = const$ for the tasks 1 and 2.

The corresponding hypothesis regarding complexity reads as:

H_1: "$\mu(V_e, t_1, c_1) > \mu(V_e, t_2, c_2)$" and

(VIII.b) H_0: "$\mu(V_e, t_1, c_1) \leq \mu(V_e, t_2, c_2)$", with $c_1 < c_2$ and $\frac{t_i}{c_i^2} \approx const.$

9.1.2.8 Hypothesis IX: Application of Heuristic *Random*

As shown in the second part of this treatise, the *Random* heuristic requires very few EIPs to come up with a decision. This small amount of EIPs makes its use especially beneficial when time pressure is high. Interpreting Hypothesis VIII the other way around, a random decision might be the 'last action' when a decision is needed at all and time pressure is growing. The related hypothesis is thus:

IX. With increasing time pressure, the deployment of the heuristic *Random* is more likely.

This increasing likeliness should be directly measurable by the share of random choices among participants $p(V_e, t_i, c_i)$ with experimental setting V_e, time limit t_i and complexity c_i, enabling the evaluation of Hypothesis IX by:

H_1: "$p(V_e, t_1, c) < p(V_e, t_2, c)$", testing the corresponding null hypothesis

(IX.a) H_0: “$p(V_e, t_1, c) \geq p(V_e, t_2, c)$”, with $t_1 > t_2$ and $c = const$ for the tasks 1 and 2.

The corresponding hypothesis regarding complexity reads as:

H_1: “$p(V_e, t_1, c_1) < p(V_e, t_2, c_2)$” and

(IX.b) H_0: “$p(V_e, t_1, c_1) \geq p(V_e, t_2, c_2)$”, with $c_1 < c_2$ and $\frac{t_i}{c_i^2} \approx const.$

9.1.2.9 Hypothesis X: Equilibrium Choice Behavior

Lindner and Sutter (2013, p. 544) report that subjects in their experiment tend to make more often equilibrium decisions under growing time pressure. As mentioned earlier, they explain these findings by chance, being unable to give a further interpretation of the underlying reasons. Following the interpretation metric presented in Chapter 8 of this treatise, the *Equilibrium* heuristic is resource intensive, leading to the conclusion that under effective time limitation this heuristic cannot be thoroughly conducted and is hence rarely applied. If this holds, the number of equilibrium choices must decrease too. The hypothesis then reads as follows:

X. The proportion of equilibrium choices decreases with growing time pressure.

This hypothesis is related to Hypothesis IV which discusses the use of strategic heuristic patterns under increasing time pressure. Nonetheless, the equilibrium decision has a probability greater than zero when the number of alternatives is finite. It is thus important to separate the chance in choice from determination, even when the decision process does not follow the model suggested in this paper. If the equilibrium alternative is chosen by chance, one could expect a proportion p_{Eq} equivalent to the following ratio:

$$p_{Eq} = \frac{n_{Eq}}{n}, \tag{7}$$

with n_{Eq} as the number of alternatives per task that would result after playing the *Equilibrium* heuristic and n as the number of all alternatives per task. This equilibrium ratio serves as the reference value for the share of equilibrium choices among participants $p(V_e, t_i, c_i)$ with experimental setting V_e, time limit t_i and complexity c_i in the following two-sided hypothesis test:

H_1: “$p(V_e, t_i, c_j) \neq p_{Eq}$”, testing the corresponding null hypothesis

(X.a) H_0: "$p(V_e, t_i, c_j) = p_{Eq}$", for all tasks, specified by its time limit i and complexity j.

How the proportion p of equilibrium decisions per task is evolving over changing time pressure conditions can then be tested with:

H_1: "$p(V_e, t_1, c) > p(V_e, t_2, c)$", testing the corresponding null hypothesis

(X.b) H_0: "$p(V_e, t_1, c) \leq p(V_e, t_2, c)$", with $t_1 > t_2$ and $c = const$ for the tasks 1 and 2.

The corresponding hypothesis regarding complexity reads as:

H_1: "$p(V_e, t_1, c_1) > p(V_e, t_2, c_2)$ " and

(X.c) H_0: "$p(V_e, t_1, c_1) \leq p(V_e, t_2, c_2)$ ", with $c_1 < c_2$ and $\frac{t_i}{c_i^2} \approx const$.

The set of hypotheses further requires an appropriate database to be (statistically) validated. It is thus necessary to acquire data with suitable quality and quantity. The level of significance of the statistical hypothesis tests target determines the latter. The former involves at least two aspects. On the one hand, this term addresses the test variables' relevance for explaining the results. On the other hand, it subsumes general issues such as measurability of test variables, careful data acquisition and minimization of error sources and data noise. Based on this set of hypotheses and the subsequent considerations, the data characteristics can now be specified.

9.2 Experiment's Data Types

The first two hypotheses explicitly refer to behavior in strategic decision environments, stating the need for an adequate variety of tasks. One possible way is to use specific strategic tasks that define the decision environment. In the present treatise, this is realized by a predefined subset of the class of normal-form games. Studying the influence of the task variables 'time limitation' and 'complexity' on the heuristics' performance requires their parameterization.[130] Additionally, context variables such as values of payoff matrices are varied.

The data created by subjects in the experiment also face certain necessities resulting from the analysis concept that is applied. As stated earlier, EIPs represent a subject's behavior. The interpretation metric (Section 8.3) is used to derive EIPs from observable behavior (i.e., EMMAs). The EMMAs need data according to Table 22. In addition, status information must be gathered

[130] Therefore, the performance concept of Section 5.1 is applied.

from the elements of the tasks (matrix cells and buttons) as well as the individual mouse movement path. With this, mouse trajectories and movement histories can be determined.

TABLE 22 – EMMAS AND DATA

EMMA	**DATA**
MOVE	Mouse position coordinates (relative: in the object and absolute: in the frame) per time, time stamps, object information,
CLICK	Mouse event: click, mouse position coordinates, object information[131], timestamps
HALT	Mouse position coordinates, object information, halt duration: timestamps

Summarizing those needs, the experimentally derived behavioral data from mouse movement shall comprise mouse position, inherent time stamps in narrow time steps (to determine the path of mouse movement) and interaction data such as clicks and moves over elements of the interface to reconstruct the information acquisition and processing context.

These data form the basis on which EMMAs can be identified and translated into EIPs, along with their historical occurrence during decision-making. This translation, in turn, is necessary to identify patterns. The results help to test Hypotheses IV, V, and VII to IX. With the amount of identified EIPs per player and task also Hypothesis VI can be tested, evaluating Goal 4. Adding information about the generated payoff to identified EIPs enables evaluation of effectiveness regarding Goal 3 and thus contributes to test Hypotheses IV, and V. Choice information finally helps to evaluate Hypothesis X.

This data seems sufficient to test the set of hypotheses directly or indirectly after some predefined transformations.

9.3 Experiment Design

A primary purpose of experimentation is to monitor and evaluate all relevant information associated with this experiment. The information gathering includes the behavior of participants during a predefined task and potentially influencing sources. In an online experiment this is more difficult than in laboratory studies. However, measures are advised from other web-based experi-

[131] Object information mainly contains object ID and object content. The object ID can be used to reconstruct a mouse pointer's absolute position within the frame. The object content gives information about the chosen alternatives in case of a button object or about payoff information in case of a matrix cell object.

ments to support this objective, ensuring control of information widely (Musch and Reips 2000, pp. 70 ff.):

- Formulate tasks such that problem task and space of results are sufficiently known to the experimenter ex-ante.
- Directly or indirectly examine every action of a subject that is regarded as relevant and record it for analysis.
- Make sure that the experimental data can be structured and analyzed qualitatively as well as quantitatively, especially with statistical methods.
- Control environmental influences. If this is impossible, identify and minimize sources that compromise the validity of the experimental data.

Those general recommendations need to be further specified for this treatise's purpose. One can derive several aspects from the recommendations which directly influence the experimental design. Other important sources for design-relevant questions are the experiment-based model and the simulation, the hypotheses and related experiments. The so derived aspects are discussed in the following. Table 23 provides an overview based on a rough structural attempt.

TABLE 23 – ASPECTS OF EXPERIMENTAL DESIGN

Technical aspects	**Content aspects**	**Design aspects**
Development and execution environment	Games and payoff structure	Subjects
Input / output relations and optimization	Task variables time and complexity	Incentives
Data storage	Time display	Questionnaire
Data preparation and analysis	/	Variation of task setup
/	/	Supporting aspects

9.3.1 Technical Aspects

9.3.1.1 Development and Execution Environment

The metric to interpret EMMAs as EIPs developed in Chapter 8, requires mouse tracking as process tracing techniques in a particular technical realization. For this realization again several technical opportunities exist: *MouselabWeb* (Willemsen and Johnson 2011), *MouseTracker* (Freeman and Ambady 2010), *z-Tree* (Fischbacher 2007) and *ORSEE* (Greiner 2004). All of them represent more or less well-established open-source software used to develop and conduct behavioral experiments. *MouselabWeb* and *MouseTracker* are directly conceived as process tracer. The other two focus on more general applicability, even though providing blueprints for decision-based behavioral experiments in their application library. *z-Tree* and *MouseTracker* have implemented software-specific programming languages and are principally bound to local networks as long as they are not embedded in a Java application. None of them offer the flexibility in modification and frequency in an application at the same time.

In opposition to them, *MouselabWeb* offers both a development and an experiment framework in HTML and JavaScript (JS). It is thus easy to be modified for experimental needs and compatible to use in standard browsers. *MouselabWeb* also provides a technical framework to control the experiment online, multiplying the number of potential participants. With this, an experiment is no longer bound to stationary laboratory PCs. For this research *MouselabWeb* thus seems advantageous compared to the other three mentioned software tools and further serves as a development and experiment framework. The following technical aspects deal with certain parts of the data accumulation process.

9.3.1.2 Input / Output Relations and Optimization

The experimental concept is about the interaction between a subject and a Personal Computer (PC). This setup requires a suitable Input-/Output-device (I/O-device) enabling necessary data gathering. The first-choice medium in use to participate in the experiment is a PC with a mouse as the input device and a monitor as the output device. The graphical user interface (GUI) uses both I/O-devices to allow direct interaction between a user and the PC. Since the technical implementation is tailored to mouse usage, other input devices – such as touchpads[132] – are not supported. HTML favors mouse usage by providing a series of exclusive event handlers that find no equivalent for other input devices. Among these is *onmousemove* which enables permanently tracking movement coordinates. The handling concept of touchpads rather relies on discrete pointing than on permanent input. An equivalent event handler is thus not provided and movement data not obtainable. Using fingers as pointer leads to another serious problem. The size of fingertips makes it more difficult to determine the coordinates of the pointing compared to a mouse-based solution. This precision, in turn, is required by the interpretation metric.

9.3.1.3 Data Storage

Game tasks with different time limits and complexity must be implemented in an experimental MouselabWeb design. Corresponding files that contain the programmed game tasks must be connected to a data storage system. For such applications, databases are preferably applied since they enable structured data storage with multiple parallel access. The browser executing the experiment thus represents the front-end, and the database on a network server functions as a back-end regarding the classical architecture of client-server applications. The technical implementation of this concept for the experiment will be presented in Section 9.4, Figure 25 (p. 182). At this point, numerous technical solutions exist. *MouselabWeb* provides an in-built PHP-based SQL-interface which is convenient to apply for the objectives of this technical approach.

9.3.1.4 Data Preparation and Analysis

The data collected from the experiment strongly depends on the feedback design. Table 22 already sets out the data for accurate analysis of experimental performance. Following this, the mouse event handlers provide this information in particular. It is therefore important to implement these event handlers and connect them to the data stream through appropriate functions. Table 19 (p. 117) lists the identified event-handlers. The following Section 9.4 discusses further technical issues regarding their implementation when concretizing the modification of *MouselabWeb* as framework software and the resulting experiment files.

---

[132] Touchpads are regarded as another common class of input devices that find application especially in mobile computing devices such as smartphones and tablets. Other input devices are not expected to occur within the experiment (where participants provide the experimental equipment). It is thus the only input device besides the computer mouse that is discussed here.

One can principally execute the analysis of the experimental data with any data processing software that provides capabilities to handle the amount of data and of course the type of data. Because the information collected is in numbers and string formats, this is not too difficult for most conventional software programs. It is thus the availability and convenience in use that determines the choice. For the objectives of this research, *Microsoft Excel* (version 2010) and *Matlab* (in version R2014b) are applied.

The targeted competition style including rankings and success related incentives (see Subsection 9.3.3.2 for a detailed consideration), might cause an unwanted but for data analysis necessary to be considered side effect: repeated playing. Some participants want to improve their rankings because of ambition, prestige or similar reasons. Repeated games, especially when facing the same game content, automatically lead to learning effects and thus could increase performance. Since repeated play is not a subject of investigation, this (technically feasible) possibility is not explicitly referenced in the experiment instructions. For the same reason, specific participation data such as player name, IP address and time stamp of participation is recorded to identify repeated attendance. Such data is then excluded from the analysis.

9.3.2 Content Aspects

The experimental approach considers the results of Part I and II of this treatise. Therefore, the game tasks used in the simulation are largely adopted for the experiment. Those imply varying complexity and time limitation. The latter needs a revision since in simulation it was modeled as number of EIPs. For the experiment, time limits measure in seconds. Other, mostly individual aspects regarded as being influential in human decision-making, such as the state of emotion, mood, and beliefs, are not studied. For simplification, those aspects are assumed to remain individually constant throughout the experiment when applying this setup.

9.3.2.1 Games and Payoff Structure

The game types in use here as well as their payoff sets at different levels of complexity are selected analogously to the simulation part of this research. In Appendix B, all game types are depicted with their payoff sets. The basic game types are ‘Hawk-Dove’, ‘Chicken’, ‘Stackelberg’s Leadership’ and ‘Prisoners’ Dilemma’. This set is complemented by four 2 by 2–games suggested by Costa-Gomes et al. (2001, p. 1203).

9.3.2.2 Task Variables of the Experiment

Time pressure is designed varying time limitation and complexity of the tasks as results from the first parts of this treatise show. Table 24 depicts the selection of parameters. The games in normal-form vary in size between 2 by 2 and 5 by 5 alternatives. Each of the four numbers of alternatives needs to be played by a subject at least once to gain a comprehensive dataset.

TABLE 24 – TREATMENT VARIABLES: TIME LIMIT AND COMPLEXITY IN THE THREE EXPERIMENT SETUPS

Task no.	**Round**	**Complexity**	**Time limit**	
			V_1/V_3 **[s]**	V_2 **[s]**
1	1	2x2	30	30
2	1	2x2	13	19
3	1	2x2	7	9
4	1	2x2	6	8
5	2	3x3	75	75
6	2	3x3	35	56
7	2	3x3	20	13
8	2	3x3	10	9
9	3	4x4	150	150
10	3	4x4	54	104
11	3	4x4	22	24
12	3	4x4	12	8
13	4	5x5	200	200
14	4	5x5	53	120
15	4	5x5	31	26
16	4	5x5	15	7

The simulation presented in the second part of this treatise comprises 466 time limitation instances. EIPs are used as time limit parameters. For practical reasons, this is not possible in an experiment: 466 time limit instances would soon exceed any quantitative constraint of both experimenter and participants. Also, setting a time limit in number of EIPs requires a simultaneous measure of applied EIPs. That, in turn, is difficult to realize in technical terms and thus hard to control.[133] For this reason, the time limitation is designed classically as time measured in seconds.

The parameters are determined analogously to the simulative approach. The categorization of the heuristics' processing in the four stages 'early', 'middle', 'late' and 'constant' is adopted here, leading to four rounds of different time limits. The corresponding restrictions are henceforth labeled as 'severe', 'strong', 'moderate' and 'ineffective', respectively.

Very few studies offer references for time limitations. As discussed in Part I and in good compliance with other studies[134], time requirement is further directly related to the concrete task. Hence, it is recommendable to determine the reference time limitations by pre-experiments. Among those studies following this method is Lindner and Sutter (2013, p. 543, footnote 4). Without further discussion, they chose as their limit as the time that 15 % of the participants of the pre-experiments needed to come up with a decision for a particular task. This seems arbitrary[135] and not differentiated enough for the current study and is hence not followed.

The approach selected in this treatise follows the concept of timely stages derived from simulation. Those are the 'constant phase' with no effective time limit and the 'transition phase' with the stages 'early', 'middle' and 'late' representing effective time limitation.[136] The time limit values for those stages are determined from pre-experiments. About 80 subjects, students of macroeconomics and already partly introduced to game-theoretic principles, participated in the pre-experiments. The procedure to determine the time limit values is as follows:

- In the 'late stage', the *Equilibrium* heuristic usually terminates. Thus, the mean time of selected participants from pre-experiments, who are instructed to play according to the *Equilibrium* heuristic, is taken.
- A time limit for the 'middle stage' is derived from the mean time of all participants from pre-experiments.

[133] The analysis would necessarily be implemented in the front-end to be executed in time. The source code for this is quite complex and hence of notable size, markedly requiring working memory and processing time. That, in turn, would make the application slow, imprecise (since the participant is moving too fast for the execution of the script) and hence inadequate for use.
[134] Also see Reutskaja et al. (2011, p. 916) for argumentation.
[135] Note that Lindner and Sutter (2013) are aiming to recognize a qualitative impact of time limitation on decision-making. The underlying cognitive mechanism is not subject to their study.
[136] Compare Figure 6, p. 64.

- In the ‘early stage’, predominantly nonstrategic heuristics finish their procedures. Among those are the heuristics *Optimist* and *Altruist*. Five participants from pre-experiments were instructed to play one of these two heuristics. The mean time they needed denotes the time limit in the ‘early stage’.

The complete procedure to determine the time limits was conducted for every game type and level of complexity.

9.3.2.3 Time Display

People who become aware of a time limitation split their concentration between fulfilling their actual task and monitoring the remaining time, as Zakay (1993, p. 61) reports. In this context he assumes an underlying cognitive process being responsible since no biological element could be detected so far that monitors or "senses" time. If the time phase spans about seconds or a few minutes, the estimation is quite precise. Any larger time phases need clocks to adequately estimate the remains (Zakay 1993, pp. 61 ff.). In opposition to seeing and hearing, this requires actual concentration. Moreover, as far as the remaining time is not made visible (e.g., by a timer), it applies cognitive resources to anticipate the time left (Zakay 1993, p. 65). This anticipation in turn reduces the resources for the actual problem-solving process.

Summarizing those arguments, not showing the remaining time has a negative impact on the decision-making process. Hence, implementing a time display seems reasonable, even though it might be found less frequent in comparable real problem situations. Of course, a time display may share attention and hence resources, too. Since one requires no time estimation, but clear recognition here, this fact should be of minor influence (Wickens 1981, pp. 36 f.; Payne et al. 1988, pp. 561 f.).[137]

9.3.3 Design Aspects

Designing experiments in behavior research always leave the experimenter with numerous options, addressing basic choices such as the form of conduct, subjects, and incentives to name a few of the most prominent aspects. The particular choice then usually is a matter of two general

[137] Both aspects – showing or hiding the remaining time in a time-limited task – have its reasonable implications in real problem situations. Thus, it would be interesting to examine both variants in the same task environment and compare results to evaluate the respective impact on decision-making here. Nonetheless, for the reasons stated above the author decided to show the remaining time.

considerations: effectiveness and efficiency. In other words: which experiment methods serve best the goals of the current examination under the constraint of economic reasons? Especially the latter fact seems essential enough that available solution concepts, conveniently embedded in existing structures, become attractive and thus are frequently selected (among others, see Gnambs and Strassnig (2007) and Henrich et al. (2010) for a discussion of selected aspects). Speaking of the effectiveness, the required dataset seems the most cogent argument to choose one option over another. This data set in turn was discussed already in Section 9.2. As an experimental environment, the software *MouselabWeb* is selected, on the one hand linking this research approach to methods and results of relevant research studies. On the other hand, this software provides all necessary means to conduct such a behavioral experiment: since the software is quite flexible to modify, the game tasks and tasks variables can be embedded and varied as well as the correct data types collected. Subsection 9.3.1 addressed the technical issues. The reported concretizations are based on a pragmatic and successive improvement of design and embedded parameter ranges by three pre-experiment sessions. Those sessions especially helped to determine time pressure parameters and experimental data that ought to be examined. Other aspects, mostly of supportive nature such as contents of the introductory part, could be specified as well.

9.3.3.1 Subjects

The selection of subjects principally has a significant influence on the validity range of an experiment's results. Findings obtained by examining a homogeneous group of participants in relation to particular social aspects may not be universal but must be limited to their peer group. However, many experimental behavioral studies are still confronted with time or financial constraints and therefore rely on the easily accessible and largely homogeneous subject pool of university students. That could profoundly influence the representativeness of the results (Henrich et al. 2010, pp. 83 ff.).[138] Again, this fact is not often considered by the authors of the experimental studies mentioned in this paper. Besides employing the own (university-based) subject pool, online-based recruiting networks such as *ORSEE* have been developed to support organization of

[138] The authors give a problem-centered introduction into the phenomenon of standard subjects in experimental behavioral science and its potential impacts on various presented results. After analyzing more than a hundred psychological studies, they conclude that the typical conception of human behavior is basically the image of a white American psychology student's behavior. However, the authors generally acknowledge the difficulty to find appropriate, diverse subjects for experiment.

a proper subject pool, as well as experimental scheduling, data providing, and payment for participants (Greiner 2004, p. 63).

This experiment seeks general findings towards decision-making under time pressure. Conducting the experiment only in German could be critical to the goals set. However, to maintain a decent cross-section of the population, an attempt is made to attract participants from very different online sources for the experiment. Whether groups are large enough to determine significance in a statistic sense cannot be answered reliably until the end of the experiment. It is thus necessary to determine the generality of the findings after the experiment since nothing can be said about participants beforehand.

MouselabWeb offers the possibility of conducting behavioral experiments with a standard internet browser. With a connection to the internet, the participants could take part from anywhere. In fact, this potentially expands the number of participants and their diversity significantly. Even if the number of participants is potentially higher than in laboratory-based experiments, it is a matter of a proper acquisition concept to realize this advantage. Nevertheless, the particular quantity remains unclear ex-ante as long as one cannot provide a (realistic) limit. Hereby, the statistical analysis – especially the confidence level – is determined ex-post, in contrary to common laboratory experiments where the argumentation usually is the other way around.[139]

A subject qualifies to participate in the experiment if it understands the tasks and interacts with the experiment environment. Motivation and attitude of a participant towards the experiment is not tested. However, one can widely allege them as fundamentally sympathetic, confident, and eager to be productive since participation in the experiment is voluntary. The object of the investigation, as well as compensation arrangements for the participation expenses, is explained as early as in the introduction part of the experiment. In this way the author wants to exclude a disappointment of the expectations regarding the experiment. Hence, participation is intrinsic motivated. However, the actual reasons for attending remain unclear for the experimenter. Moreover, this derivation is in line with the findings of Göritz (2006, pp. 65 f.). That, in turn, would reduce the need for monetary incentives which are discussed in the following subsection.

[139] Even though the scheduled quantity varies in reality, the number of participants in laboratory-based experiments is principally plannable. With this, the targeted confidence level determines mostly the number of subjects for an experiment.

Of course, one has to consider appropriate actions to compensate for the accompanying loss of control in an online experiment. This consideration includes that data analysis must be designed appropriately to identify undesired behavior and distinguish between the variety of players who meet the requirements as mentioned earlier and those who do not.

One can meet the concerns mentioned above for instance by providing a proper introduction section including a questionnaire to collect social data from the participants and a training session (see Subsection 9.3.3.4). Individualization of information and consequent data monitoring can encounter unintended repeated participation (see Subsection 9.3.1.4). After identifying those unusable datasets, one can easily exclude them from consideration.

9.3.3.2 Incentives

As for any experimenter who needs to rely on empirical data for his investigation, the declared goal is to

- attract the highest number possible of subjects,
- who optimally complete the whole experiment, and
- who generate useful data.

These three goals can be achieved with or at least supported by incentives (Tversky and Kahneman 1986, p. 5274). When planning an online experiment, designing participation incentives for potential volunteers is an essential issue that the experimenter must consider. The answer to that question fundamentally impacts the acquisition and motivation of participants in the experiment. Appropriate means accounted in this study are specified in the following.

Perseveration rate

The number of participants and the perseveration rate[140] in experiments which expect more than one action from a participant need to be considered separately. Based on experiences, the number of volunteers in online experiments is expected to decrease with the ongoing procedure. The reasons for that are varied but regarded as influenceable up to a certain degree by incentives. In the present multi-round experiment, data is necessary from every stage to draw a proper conclusion concerning hypothesis testing and classifying behavior.

[140] The perseveration rate in the meaning here is the ratio of the number of participants ending one round of experiment and the number of participants beginning the same round. The result is a real number from the interval [0,1], with a ratio of 1.0 being the optimal case of all participants finishing.

Since these rounds are applied sequentially, with principally increasing difficulty from game to game and from round to round, the interest in completing the experiment might decline. If the experiment is terminated prematurely, the data for the subsequent tasks are missing for the evaluation.

The experiences from two pre-experiment sessions with tasks comparable in design and difficulty at Helmut-Schmidt-University in late 2013 and early 2014 are helpful to evaluate the expectable perseveration rate in the context of the present experiment. In the first sessions, students and academic staff were asked via email (university-wide mailing list) to participate in a study. No incentives were given. Of those roughly 3000 potential subjects 114 started the first round. Of those, 86 also began the second and 49 the third round. Finally, 41 finished the whole procedure. That gives a total perseveration rate of about 36 % measured from the beginning to the end with a mean perseveration rate of about 72 % from round to round.

In the second session, about 50 undergraduates of a macroeconomics course were advised to fulfill the tasks as an exercise beforehand an exam. As in the first session, no incentives were given. Ten students participated and began the first round. Of those, eight ended the whole session resulting in a mean perseveration rate of 80 % over the four rounds, losing one subject after the first and one after the second round. Even though especially the second session has too few participants to evaluate the numbers statistically, two aspects are to be mentioned: first, one can find the highest perseveration rate for both sessions between the third and the last round. In other words, whoever endures up to the third round is likely to keep up. Second, the resonance regarding this topic seems not overwhelming. Even at a course that principally deals with the topic 'game-theoretic decision-making', very few participants could be attracted. This circumstance emphasizes the question for a proper incentive to at least begin the experimental session.

In the case of online experiment, no or just a few personal ties to persons or institutions exist that could motivate participants through a feeling of personal obligation. With this, the personal commitment to ending the tasks is supposably lower than in the university-based session. Incentives are thus proposed to compensate for this aspect. The net result of those two competing influential aspects is difficult to evaluate. It is thus the more conservative case expected, i.e., a lower perseveration rate for the online experiment than experienced in the university-based case. That, in turn, has a direct impact on both incentives design and the number of internet sources that need to be acquired to activate a proper number of participants.

Data Quality

The collected data features are generally determined ex-ante through the technical procedure and the requirements of the experimenter. Their concretization then is based on the actions of the subjects. Defining a data error as an undesirable deviation from a (previously defined) state or value, both of the areas stated above are potential reasons for errors. Hence, the experimental design must be adapted to both, to avoid mistakes and to recognize flaws. While errors in the technical procedure can be minimized to noise by the experimenter, this possibility is limited in the field of a subject's behavior. Since the participants' behavior is treated as variable under investigation, total control is impossible and not appropriate. At this point, the author identifies the incentive design as a significant factor.

Following the other reflections, incentives could support generating appropriate data to a large extent. Hence, its design for the current experiment needs careful consideration. Concerning an appropriate incentive structure, economic and psychological experimenters follow diverse, sometimes even opposing approaches based on a fundamentally different concept of man. Economists see monetary incentives, or at least equivalent valuable benefits, as the main motivation for participating in an experiment and continuing perseverance (Musch and Reips 2000; Göritz 2006).[141] By contrast, psychologists often assume that subjects are cooperatively and intrinsically motivated and therefore prefer curiosity as a motive (Camerer 1995, p. 635).

In case of online experiment, the participation obstacle is significantly lower: neither location nor scheduling is a critical problem here. Due to high availability and low costs of the internet, the remaining relevant factor of effort for a participant is time. This effort, of course, needs to be rewarded. By conducting online experiments, one can expect more significant sample sizes of volunteers. A purely monetary incentive as usually paid in laboratory-based experiments could

[141] To attract a satisfying number of volunteers for their laboratory tests, economic experimenters need to offer a sufficiently high participation incentive. As an incentive, they usually pay an allowance as base fee and a different success-related part potentially ensuring motivation and adequate behavior (i.e., behaving in good compliance to the experiment's targets and obeying the rules given by the experimenter). Church (1993, p. 62) reports a 24 %-effect of monetary incentives on the response rate in email-based surveys. Yu and Cooper (1983, pp. 39 ff.) found similar figures in an earlier survey of studies dealing with questionnaire tasks. Additionally, the relationship, even though not always linear, is observable between rising amount of incentives and rising response rate (Yu and Cooper 1983; Church 1993; Singer et al. 1999). Note that even if tasks are entirely different between studies stated here (questionnaire/survey) and this treatise's purpose (behavioral experiment), the preconditions for acquiring participants are quite alike: all studies need to find a way to attract people to respond in either answering a questionnaire or taking part in an online experiment. As will be shown later, the timely effort when participating in the experiment is comparable to usual online surveys. The distinctions in tasks are hence regarded as irrelevant and the results of above-cited studies useable here, too.

thus become quite expensive, easily exceeding the author's research budget. Of course, the payment procedure could be a technical and organizational hurdle as Göritz (2006) knows, but in time of increasing web payment applications, this seems a minor problem. Another finding of her is even more worth to consider: she reports that monetary incentives in online experiment are not as convincingly effective as in "offline" studies. Göritz examines the effect of monetary incentives on binding participants and comes to a sobering result. The willingness to participate when monetary incentives are rewarded is just slightly higher compared to the one without (+ 2.8 %). The perseveration rate is nevertheless increased by 4.2 % (Göritz 2006, p. 65). Considering the expectable benefit of providing monetary incentives in an experiment, the cost-benefit ratio seems unattractively low. However, since it is a supporting element, it should not be ruled out per se.

Other approaches of increasing participation numbers in web-based studies are proposed by Dillman (2011) who focusses on individualized requests or calls for entries as an appeal to altruistic motives. Unfortunately, he gives no evidence of effectiveness or any quantitative effects here. Again, the expected effort is in no reasonable relation to the benefit. A general call for participation an experimenter could state in different networks and on websites is less effort intensive and hence seems advantageous here.

A more promising alternative to the monetary-based incentive located between the psychological and the economic approach represents the creation of a competitive situation between participants. The strive for prestige and rank positions when comparing with others forms – especially for many game situations – a strong incentive. It is thus a frequently applied element in all sorts of games.[142] The prospect of openly published performance-based rankings, when communicated prior the experiment, can be a simple but effective means of binding participants. The publication has another beneficial side-effect: Russo et al. (1989, p. 765) report that subjects generally behave more according to the experimenter's intention when the actions they take face publicity. It should, therefore, be applied here. Its application seems inherent since the game-theoretic tasks examined are of noncooperative characteristic. Of course, curiosity still triggers attendance and might be important throughout the whole experiment. A proper description can support this issue

[142] Schell (2008, p. 186) for example characterizes competition as a proper way to satisfy the human urge of determining who is most skilled in something.

in turn but mainly depends on numerous other hard-to-influence aspects on the intrinsic motivation (Camerer 1995, p. 635).

Summarizing the points stated, participation, perseveration and data quality in the present experiment is supported by creating and communicating a competition situation where the players earn a prize depending on their overall performance. Finishing the experiment is also granted to strengthen the will to terminate. A lottery represents this grant, where all subjects who finish the whole experiment procedure participate. The prizes are presented as vouchers of different values for an Internet-based bookstore.[143]

9.3.3.3 Questionnaire

The questionnaire is used to investigate possible correlations between social attributes of the participants and their performance in the experiment. Its form is divided into three sections, asking the participants for some person-related details, education and current occupation. Table 25 provides an overview of directly and indirectly addressed criteria.

TABLE 25 – QUESTIONNAIRE CRITERIA

Directly asked	**Technically assessed/ automatically supervised**
User name (player name)	ID (unique)
Age	IP address
Sex	Browser user agent information
(highest) educational degree	Experiment version
Educational profession	Player type (row/ column)
Education in game theory (yes/no)	
Job	
Job status (for soldiers only)	
Foreign assignment (in days)	

Besides standard interests such as age, sex, and educational details some questions are stated concerning the participants' experiences with strategic tasks in general and with the solution concepts of game theory in particular. Due to its specific characteristic, the experimental tasks are likely to be new to many of the participants. It is further unclear whether they possess 'genuine' solution concepts in their problem-solving portfolio. Hence, the participants' experiences with

143 The Association of Friends and Supporters of the Helmut Schmidt University / University of the German Armed Forces Hamburg e. V., as the sponsoring company of the university, kindly funded the book vouchers.

such tasks are addressed, too by asking for visited courses in game theory during education. Besides, the questionnaire considers the military background including possible experiences in foreign assignments as one social aspect to be studied. Appendix D provides a screenshot of the questionnaire. In addition to the questionnaire, the experimenter monitors some technical data. The participant is asked to choose an individual name to link experiment data to a specific test person. At the same time, a user's IP address is recorded to identify unintended repeated participation effectively. For reasons in context with data interpretation the experiment software also collects browser and operating system information. This way, it can be determined whether a participant employed a mobile device or a PC. Furthermore, the online sources from which participants call the experiment are evaluated by determining the link address.

9.3.3.4 Variation of Task Setup

The appropriate processing of acquired information is based on two verifiable assumptions as Costa-Gomes et al. (2001, p. 1210) and Willemsen and Johnson (2011, p. 24) state.[144] Those are 'occurrence' and 'adjacency'. The former postulates that information must be visible before one can use it. Costa-Gomes et al. (2001) hence cover up all data at the beginning of their experimental design.[145] A player could reveal it by clicking on the cover. The experimenter monitors the information's occurrence by recording this click and the time the information is visible before the cell covers up again. Adjacency postulates that information gathering is almost equal to its use. The experimenter can record adjacency along with occurrence. Both principles are logical and also relevant here. They are thus considered in the discussion of the experimental design to ensure the validity of the acquired information and their interpretation.

The application mentioned above in the study of Costa-Gomes et al. (2001) is just one of multiple possibilities here to link information's visibility to a subject's attention. In the spirit of their study, the current experiment applies more versions of task presentation. Herewith, one can adequately evaluate the influence of the experimental design. At the beginning of the experiment, the ver-

[144] Both paradigms already play an essential role in the interpretation metric developed in this treatise (Subsection 8.3.2).
[145] For comparison of results, they applied a second experimental setting in which they let their control group play with uncovered payoff information. Costa-Gomes et al. (2001, p. 1202) report in this context of certain difficulties to monitor occurrence with this setting and hence discard this version for their analysis of information search and processing patterns. However, they used the decision information obtained from this setting for studying decision-making.

sions are randomly assigned to the subjects. A single participant sticks to one version during the whole session. This treatment shall prevent subjects' confusion.

The adjacency assumption is weaker than the occurrence assumption, requiring just *almost* equality between acquired and processed information. Indeed, the subjects might use more information than is stated in the experiment by employing experiences and knowledge when processing the task. The use of supplementary information is difficult to control. Some authors thus try to evaluate this influence by an intensive questioning of the participants prior, post and even parallel to the task solving procedure (Ericsson and Moxley 2011). Since this method is likely to have a negative impact on the performance by disturbing the cognitive process, it is not followed here. Besides, this method raises some technical concerns regarding its realization, too.

Another argumentative approach is to minimize the influence by employing tasks that are likely to be solved by methodology and problem-solving skills rather than purely remembering the correct answer. Numerical problems such as the class of normal-form games indeed are of this characteristic. Especially when changing the payoff values (by linear transformation for example), the probability of reconnaissance for the concrete problem is expected to be very low. Its solution is thus relying on the skills as mentioned earlier and the stated information.

Taking those arguments together, it is worth defining different experimental setups. This way, the influence of the design on decision-making can be identified by studying various forms of recording occurrence.

Besides this, one can generally assume adjacency when normal-form games are in use. The experimental design supports adjacency by suppressing learning effects by two means. First, game types (represented by different structures of the payoff matrix) are varied. Second, in tasks of identical game types, the underlying basic payoff matrix is multiplied with a positive integer. In the following, three setups are presented which are applied in the experiment.

As stated above, the experiment contains three different setups, in the following, referred to as versions V_1, V_2, and V_3. A subject plays one of the three versions when participating. As mentioned earlier, an automated process randomly assigns subjects to a specific version after receiving the questionnaire. With this, each version is played by about one-third of all subjects. Experimental tasks, as well as instructions and information pages, are version-specifically adapted.

The main difference between the versions is the way the subjects can technically access the payoff information. The version V_3 works in analogy to the version of Costa-Gomes et al. (2001) and with a few deviations to Johnson et al. (1989) and Willemsen and Johnson (2011). All information is covered at the beginning. The participant can reveal one cell after another. At a time only a single cell of the payoff matrix can be uncovered.

The versions V_1 and V_2 are variations of this approach. The version V_1 also starts with hidden information. However, it relaxes the policy of strict singularly opened information to a level that each cell of the payoff matrix can be covered and uncovered independently. The version V_2 instead starts with uncovered information and is thus similar to the 'open box' version of Costa-Gomes et al. (2001).[146] The main difference additionally implemented here is that cells can be closed and reopened independently.

Nevertheless, two aspects are to be mentioned to support the approaches V_1 and V_2 by pointing to the differences in comparison to the discarded approach of Costa-Gomes et al. (2001): first, these authors use matrix sizes of up to 2 by 4, whereas in this experiment complexities of 5 by 5 are reached. Pre-experiments show that the mouse usage intensifies with growing complexity. The author assumes that the pointer becomes a critical asset as a viewing and concentration aid when the number of cells is increasing. This circumstance has a positive effect on both, time spent in matrix area and duration time per cell. Second, Costa-Gomes et al. (2001) have not asked their participants in their introduction to use the mouse for information acquisition and processing. As a result, subjects reduce mouse movements to choice selection and task finishing exclusively in the 'open box' treatment (Costa-Gomes et al. 2001, p. 1202). In this treatise, the experimenter explicitly recommends mouse use for information acquisition and processing in the introductory part of the experiment.

For all three versions, 'occurrence' and 'adjacency' are monitored employing the mouse movement tracking concept proposed in Part II of this paper. This approach enables comparison with the approach of Costa-Gomes et al. (2001), by comparing the findings of V_1 and V_2 with those of V_3. Besides, it is possible to test the two refined versions in the context of a well-established approach. These results are applied to verify the interpretation concept mentioned above. Table 26 provides more details regarding those properties described as well as some additions.

[146] The authors reported from poor mouse usage in connection with this version Costa-Gomes et al. 2001, p. 1202 and thus could not study information search and processing with it. See footnote 145, p. 169 of this treatise.

TABLE 26 – COMPARISON OF EXPERIMENTAL VERSIONS

Version → Attribute ↓	$\boldsymbol{V_1}$	$\boldsymbol{V_2}$	$\boldsymbol{V_3}$	**Costa-Gomes et al.**
Game types (and payoff structure)	Identical between $V_1 - V_3$; at a complexity of 2x2 also identical to Costa-Gomes et al.			Partly identical
Time limitation	Identical to V_3	Generally shorter than V_1	Identical to V_1	No time limit
Complexity	Identical (matrix size of 2x2 – 5x5)			2x2 – 4x2
Task	Chose single strategy	Eliminate strategies	Chose single strategy	Chose single strategy
Payoff matrix - shape - content	 - quadratic - own & opponent's payoff combined in one matrix	 - identical to V_1 - identical to V_1	 - identical to V_1 - identical to V_1	 - n by m - own & opponent's payoff separated in two matrices
Content visibility in payoff matrix - at the beginning - at a time	 - closed - any cell	 - open - any cell	 - closed - one cell	 - closed - one cell

9.3.3.5 Supporting Aspects

The development of the experiment primarily focusses on data generation as stated earlier. Having enough qualified participants is crucial for determining significance of the findings and thus for the success of this study. It is hence necessary to attract enough volunteers and to bind them to the experiment for the whole session, if possible. Various forms of incentives were discussed earlier already. Besides, literature reports of more, mostly smaller design aspects to support the goals as mentioned earlier. Everything that increases motivation and maintenance without significantly changing the intention of this research or decreasing data quality is highly appreciated and hence considered. With the 'Framing Effect' and the 'Halo Effect', Kahneman mentions two such intention-changing design effects one needs to consider when developing an experiment (Kahneman 2012).

Tversky and Kahneman initially reported the 'Framing Effect'. It describes the influence of a task's formulation on decision-making (Tversky and Kahneman 1981, pp. 457 f.). Their description explicitly contains the concretely selected information stated together with the task. The literature suggests specific prevention means to avoid unintended consequences from this effect (Tversky and Kahneman 1986, pp. 5270 ff.). This experiment's design fully covers those means. They are presented in the following:

- formulate the goal of the task as neutrally as possible by avoiding judgmental attributes,
- balance the given information to prevent favoring a particular view or attitude, and finally
- use a non-monetary payoff unit to avoid associations with real gains and losses which could evoke risk attitude-related behavior. In case of this experiment, the payoff unit is hence declared as ‘points’. This measure also supports that the focus of the participants lies on the competition, rather than on making (or losing) money.

The term ‘Halo Effect’ is primarily coined by Nisbett and Wilson (1977b) for decision behavioral observations. It describes the attention-grabbing influence of visual stimuli that causes neglecting other information given in the context of a decision task. It is suggested to design equally valued information (regarding their importance) uniformly to avoid such unintended influence. That affects color, form, and position of the individual elements on the experiment's task webpages. In the light of those effects, one can evaluate the other design element proposals as follows.

The interaction aspect is strengthened by the dynamic, browser-based see & click concept with various actions required over the whole experiment process. The handling is designed as easy as possible and does undoubtedly meet ordinary internet users' experiences.

In the same spirit, the author implements the following interactive feature for the display of the payoff matrix. One can increase the attractiveness of mouse usage by combining visual features with it. A conventional method is to change the contents or at least the contents style when interacting with the mouse pointer. In the case of the current experiment, the content is made to appear bold and bigger every time the mouse pointer enters a matrix cell, supporting the visualization of the payoffs and stimulating the mouse usage when focussing one's interest on a particular area.

Introduction part

The introduction is a comparatively important part of the experiment. Its function is manifold here. At first, it should raise and maintain attention, already beginning with the initial page. One can achieve this intention by presenting proper contents, such as an interest-catching detail related to the experiment or prospect of incentives. Of course, the facts stated on the first page should inform about what can be expected in the experiment, too. It is regarded as helpful to increase the number of participants and reduce the early-termination rate at the same time by merely presenting the expectable duration of the experiment process (Musch and Reips 2000, p. 19). Also, it could give a motivational aspect, emphasizing a participant's contribution to the whole project.

A questionnaire succeeds the homepage. On a single screen, the visitors are asked to answer a couple of questions regarding their person, education, and job. For more details see Subsection 9.3.3.3. Instructions follow in this form. They contain an informational part and a short test of understanding. The full text of the instructions the participants have to screen before the experiment begins, is given in Appendix D of this treatise. The content of the informational part relates to the information given at the starting page but provides more details. Especially the experiment procedure is emphasized here.

Additionally, the last instruction page refers to four facultative screens subsumed under the label 'tips'. The tips contain further information about game theory in general and on the underlying idea of the fundamental solution concept, without giving a detailed procedure. Moreover, the participants are informed about what happens when no action is taken (payoff equals zero) and that they can solve all tasks with basic arithmetic operations.

A short test succeeds the informational part to guarantee a minimum of understanding. It contains two questions of comprehension which are equal for all participants. The first question addresses the payoff generation and the concept of optimality. This question ensures that the subjects understand which cell entry of the payoff matrix belongs to which player. Besides, the subjects are informed once more that the optimality criterion is equal to the maximization of payoff. The second question deals with the time constraint. The participants are asked to identify the remaining time displayed, again laying the focus on the time limitation of the tasks and addressing the time's measuring unit. Two screenshots of actual tasks visually support both questions. After going through this little test, the participants finish the instruction part, having visited four obligatory screens in total.

The instruction is succeeded by a first pause screen where the subjects can rest, and start the experiment when feeling ready. For the same reason, this pause screen appears after every four completed tasks. It appears a total of four times in the experiment.

Another design element regarded as being successful in maintaining attention but less frequently combinable with the general intention of a scientific experiment is the fun factor. People are generally attracted by games, investing time and sometimes even monetary resources to play. Thus, having a good time playing would be incentive enough. The experiment here cannot fully adopt a computer game design. However, the experiment design can consider particular aspects. Schell (2008) discusses twelve determinants to be considered in the development and design of exciting and attractive games. Although his book is aiming at computer game developers, many of the variables mentioned are generally influencing for games and might be of use in this case, too. Especially the motivational variables are regarded as particularly relevant in the context of this study. Among these, the author identifies the following seven aspects that effectively influence the experimental design to various degrees: ‘Fairness’, ‘Luck vs. Ability ‘, ‘Reward and Feedback’, ’Simplicity vs. Complexity ‘,’Relevance of Player’s Decision’ and ‘Challenge vs. Achievement’. Those factors are raised in the context of computer game design and not for experiments. They thus need to be regarded as recommendations only. The following discussion considers the aspects' implementation in more detail.

Fairness

The concept of fairness in this context reflects the equality between the players to be able to achieve success. The present study guarantees this intention by balancing the payoff structure of the games so that any participant could generate the same overall payoff, regardless of the game type combination or player perspective (row or column) played. The subjects have hence the same chances to win the rankings. Further, the level of difficulty is rising with each round played. Since this represents the concept of the experiment, every player faces the same level of difficulty when participating and is in so far treated equally.

Whether the tasks are fair in the sense of being individually solvable, needs to remain undecided. It is the objective of the experiment to examine the subjects' problem-solving ability under time pressure. Of course, one can expect that people perform differently here. Hence, the tasks' level of difficulty is not a valid aspect regarding fairness.

Luck and Ability

One can use the influence of this aspect for classifying games. On one side of the scale reside games where the outcome is heavily depending on chance. The other side represents games where skills and knowledge are critical. In between, there are games of virtually any thinkable variation. Most classic games show elements of both in varying combinations since this is of unique attraction for people. In the experiment, the payoff is also determined by both. On one hand, the player faces the hard-to-predictable, mainly unforeseeable opponent's behavior and on the other can influence the outcome by its own choice. This game-theoretic characteristic is also relevant for the aspect of effective ‘Relevance of Players Decision’. Schell (2008, p. 181) further mentions the point that low-risk decisions are associated with low income and high risks with high income. This risk-related payoff can be found in the experiment, too. Nonetheless, this function is reduced to choice (where a subject can generate positive and sometimes negative payoffs) and no choice (with a payoff equal to zero) and is hence very limited.

Feedback and Reward

Brehmer (1992, p. 234) for example mentioned that feedback delay could have a negative influence on participants performance. A consequence of immediate feedback is a potential change in behavior when the outcome in one task is not satisfactory as Mellers et al. (1992, p. 331) find for a nonstrategic, risky decision-making context. That, in turn, supports the critical suggestion that no immediate feedback potentially means no change of behavior through learning effects. The study of Costa-Gomes et al. (2001, p. 1198) explicitly addresses the feedback aspect. The authors purposely abandon direct feedback, arguing with the intention to minimize the influence of learning effects on the results. They design their experiment in a laboratory environment and hence retain the clinical control about the subjects to compensate the missing of positive impacts on motivation. The motivation is monitored and influenced by both interaction with experimental staff and the prospect of payment after the end of the experiment. Hereby, the experimenters keep their termination rate at a moderate level.

As mentioned already, the intended online form of the current experiment allows no such clinical control and thus needs to rely on alternative means. Immediate feedback is technically not possible here (see footnote 133). For the above-stated reasons it also not favored here. Related methods are thus used to inform the subjects about their goal attainment after the experimental session.[147] Already proposed in the introduction part, the subjects are informed that after a comparatively short period of data analysis[148] the performance results are published in the form of a ranking. The distribution of the incentives is presented this way, too.

The reward is both materially and immaterially organized, as discussed in Subsection 9.3.3.2. Besides bookstore vouchers, partly distributed via a lottery, a ranking represents the performance in both ranks and absolute numbers of success. The procedure of determining the ranks is made transparently by the explanation in the introductions. This transparency also adds to the aspects of 'Fairness' and 'Relevance of Players Decision'.

The experimental setup also includes a possibility for participants to give feedback to the experimenter during the whole process. The experimenter provides contact information (email address and social media links) in the footer of the experiment's webpages. This statement is combined with a request to contact the experimenter in case of problems, requests, and suggestions in context of the experiment.

Options for Manipulation and Countermeasures

Besides raising the attractiveness of the online experiment for potential participants, one needs to consider the security issue with great caution. Since private information and monetary values are handled, the online activity could principally come into the focus of bogus intentions. Thus, specific measures need to counter vulnerability. Of course, the thread level requires a realistic evaluation.

This experiment principally faces manipulation threads, since it is online available and offers valuable monetary vouchers. However, it handles minor values which are moreover unknown to a potential user before the experiment. Also, the range of online distribution is limited to a handful of websites and some social communities. A manipulation out of economic reasons seems thus not very likely. Also, the data collected in the experiment procedure is too specific to be attrac-

[147] Note that the experimental session described in this treatise is comparably short, just lasting minutes rather than hours which for example Costa-Gomes et al. (2001) assess for their experiment. The motivation is thus more likely to remain without further external control throughout the experimental process. However, the influence of incentives is still regarded as very helpful to increase the perseveration rate.

[148] The length of the period is not further concretized in the introduction. It took the author about two weeks to generate the feedback information of the rankings.

tive for any commercial use on the one hand. On the other, the private data is not specific enough for fraud-intended use. Following this argumentation, the level of the thread seems low to moderate.

Nonetheless, the experiment provides specific security measures. At first, the IP of any user is monitored, making repeated participation transparent. If someone intends to manipulate or illegally collect data from the experiment, the intruder either faces a password secured barrier to the database server in the back-end. Alternatively, a data phishing approach could try to collect data in time by providing a website similar to the experiment site with malicious intention. However, no such phishing site was observed during the experiment process. Of course, well-versed people could overcome those technical obstacles to access the private data.

The vouchers are handed out after evaluation of the test data by e-mail conversation with the winner. The server has not stored coupon data associated with the experimental data. It is therefore impossible to steal the data in this way

9.4 Technical Implementation of the Experiment

A set of design aspects has been identified and discussed in the previous section. These issues find application in a particular experiment environment. As stated earlier, *MouselabWeb* builds the technical framework. It employs the process tracing technique mouse tracking (Kühberger et al. 2011, p. 4). The software in its recent version is a digital toolkit developed by Eric J. Johnson and Martijn C. Willemsen that supports the design of decision-making experiments (Willemsen and Johnson 2011, pp. 26 ff.). It records a subject-generated mouse event data stream that helps to examine the underlying cognitive process in decision-making.

After determining the development environment with *MouselabWeb*, it seems appropriate to use the intended experimental approach as a template for further discussion of the technical implementation. This way, every a priori identified aspect is considered in a logical structure.

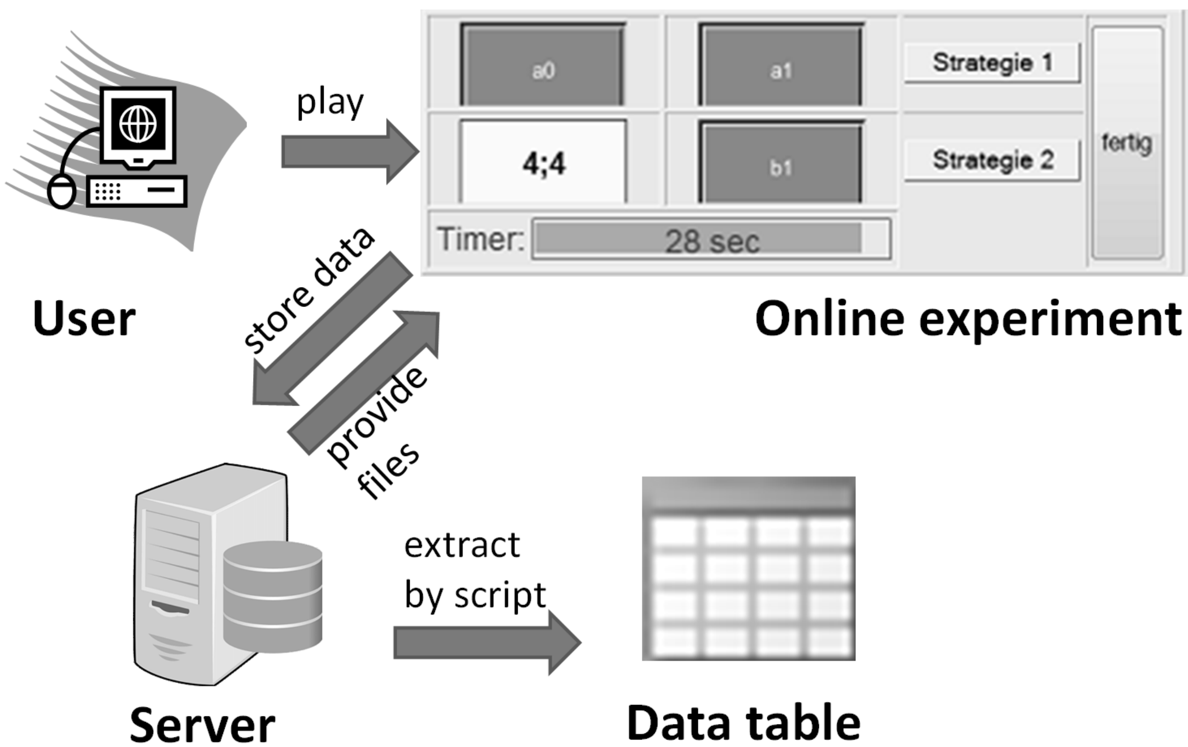
Tools of Experimentation

Access to the experiment is managed by links that are presented in different social networks and embedded in certain websites. Additionally, links are spread by specific University-based mailing lists. The links can be analyzed regarding their calling origins. Technically, this is realized by creating an experimental homepage for each starting point where a link is embedded. The calling homepage further sends its ID via cookies[149] to the next page where it is stored using HTML form elements.

As a programming language, HTML is used throughout the experiment design, allowing the use of standard web browsers. JavaScript realizes interactions between players and games. PHP scripts allow the transfer of collected information from the browser to the central data server. The server provides a SQL database for each variation of the task setup (see Subsection 8.3.2.3.3 for further considerations). Figure 25 shows the corresponding file architecture.

The introductory pages are conventionally linked together by buttons that contain the uniform resource locator (URL = web address) of the successor page. Questionnaire and information pages provide form elements and JavaScript functions to either store data for later migration to a server or to interact with a subject.

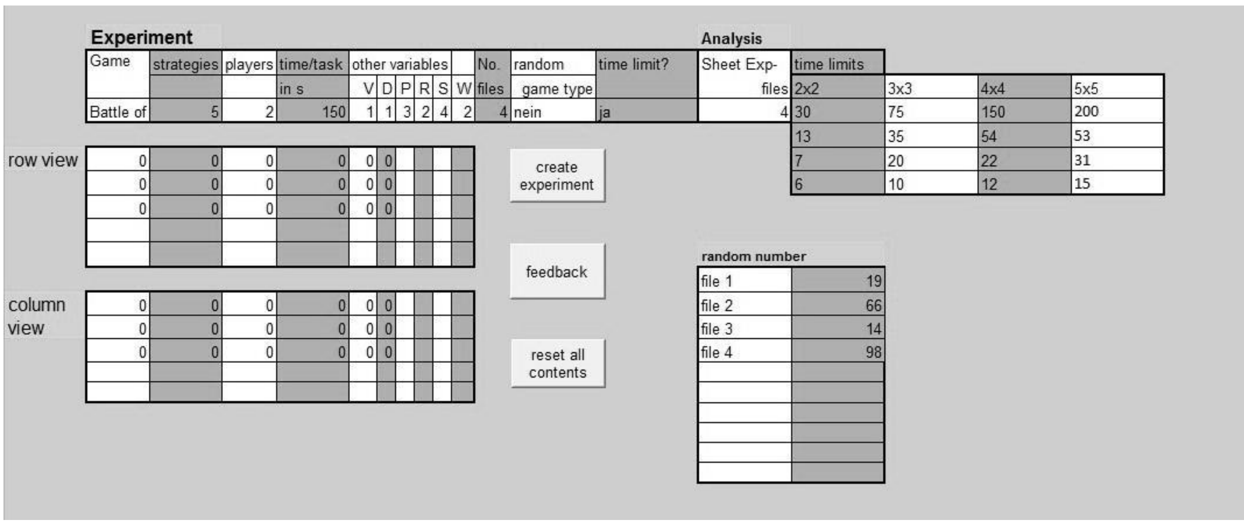
FIGURE 25 –FILE ARCHITECTURE OF THE EXPERIMENT



[149] This requires the (implicit) permission of the participants to use cookies throughout the experiment process. All subjects are informed about the use of cookies in the footnote on the homepage.

As mentioned earlier, *MouselabWeb* in its basic version does not provide all functionalities necessary to track mouse movement behavior comprehensively. The task files it creates are thus serving as templates which need to be enhanced. A program henceforth called ‘automatic game task designer’[150] adds missing functionalities. This software application is exclusively developed for this experiment and employs *Excel* as front end and VBA as back end. Figure 26 gives an impression of the front-end depicting parameter-entry fields.[151]

FIGURE 26 – AUTOMATIC GAME TASK DESIGNER: FRONT END



The front end enables the experimenter to fully determine tasks variables for a game and game type, as well as some content and organizational variables. Table 27 presents the modification range of the program.

The automatic game task designer supports the experimental design favored and applied in this research. It enables task designs that utilize a given set of game types, including their related payoff matrices. Task setup variations different from the one presented here can also be realized with the automatic game task designer. After determining the task design with its various parameters, the corresponding experimental files are coded automatically and are ready to be used in the experiment proposed here.

Employing this program has at least three significant advantages: first, all files are of the same design and same quality. If the setup changes, it can quickly be realized for all game task files,

150 Further documentation of the automatic game task designer (in German) is available with the software and part of the overall documentation of the experiment.
151 The documentation of the automatic game task designer is written in German.

making the speed of creation and adaption the second advantage. Finally, modification of the *MouselabWeb* files, as well as the enhanced ones, is transparent and hence replicable. With slight modifications, other setups with comparable goals are easy to implement and fast to use for experimentation.

TABLE 27 – MODIFICATION RANGE OF AUTOMATIC GAME TASK DESIGNER

Parameter	**Value**
Game type	Random or one of the set: Prisoner's Dilemma, Chicken Game, Throwing Fingers, Stackelberg's Leadership, Hawk-Dove Game, Battle of Sexes
Number of strategies	2–5
Number of players	1–2
Time limit	yes / no
Time limit value	positive rational
Number of tasks per round	positive natural
Number of rounds	positive natural
Specifications of game types Prisoner's Dilemma and Hawk-Dove Game	rational

The sixteen task files per experimental session are grouped into four in a predefined order, with each group forming one round out of four. A subsequent pause page separates the first three rounds; a final page follows the last round. All pages are linked together by calling the URL of the next page when the current task ends, or a subject clicks a particular 'submit' button. When arriving on the final page, a JavaScript module creates and displays a unique code with which the enduring players can take part in a lottery. Those codes are a combination of random numbers and the time stamp of the moment the player ended the last task. The player/code combination is stored via a form file into a separate table within the server database.

Besides introductory files, pause and final files, and automatically generated task files, others are coded that contain functionalities for interaction and controlling the data stream to the central experiment server. Those interaction files are embedded into all task files and serve as a library of functions. This centralization of contents keep the allocating task files in a reasonable size and thus betters the overall time performance of the application.

The experiment files contain several form elements (especially the questionnaire) and mouse event triggered functions that call the interaction files coded in JavaScript. The interaction files contain form elements to gather and store data. This data is migrated in a further step to the data-base. Here, every single task started by a subject is automatically assigned an own ID. Herewith, experimental data link to corresponding players. Together with the information from the questionnaire, a comprehensive analysis of a single subject's performance is possible.

A university-owned server hosts all experiment files mentioned as well as a SQL-database to store the experiment data.

This chapter introduced the design and the execution of the decision-making experiment that acquires behavioral data to study cognitive processes in strategic tasks. Results of this experiment, especially the classification of behavioral patterns, are presented in the following chapter.

10 Results from Experimentation

This chapter is dedicated to comprehensively illustrate the data gathered from the experiment, online conducted in the period Oct 6th to Nov 11th in 2014. The focus lies on data analysis, helping to answer the central research questions raised in Chapter 1.2:

(III) What behavioral patterns in the cognitive processes can be distinguished and how can they be classified?

(IV) How does time pressure affect behavior on a process-component level and concerning patterns?

Nonetheless, other interesting facts are stated and put in context to those research questions. The chapter is organized threefold, at first considering aspects of data acquisition before going into detailed data mining in the second section with evaluating the proposed set of hypotheses. In the following third section, participants' decision-making is classified, based on the characteristics analyzed in hypothesis testing.

10.1 Data Acquisition and Description

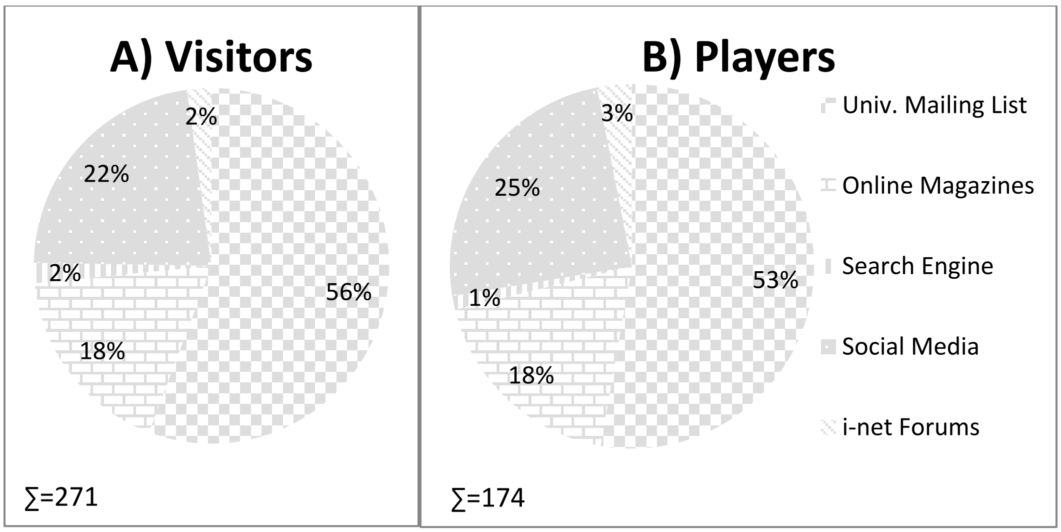
10.1.1 Attendance and Perseveration

The experiment was conducted throughout the entire month of October 2014, with announcements actively spread starting October, 6th. Sources, where potential experiment participants can be attracted to take part, were identified in several fields, as there were university-based mailing lists, online magazines, social media, internet forums and online search engines. The success of such measures is listed in Figure 27, depicting A) the originating webpages of persons who view the introductory part and B) the originating webpages of players who start the experiment.

Universities where the author of this treatise had personal contact who might be able to access the student mailing lists were selected to participate in the experiment. With that intention, five German universities were contacted, asking for access to a proper mailing list. Among those, Helmut-Schmidt-University in Hamburg and Georg-August-University in Göttingen positively responded. Others formulated doubts concerning data safety aspects or generally followed a strict and denying policy in cases of external research questionnaire requests. The former two might share that idea. However, due to personal contacts to both institutions, the proceeding was officially legitimated.

Besides universities, the online magazine “Psychologie Heute” (German for “Psychology Today”) offered the possibility of listing links to behavioral experiments. Specific criteria, such as scientific background and several formal aspects, needed to be met before the experiment was listed in their online content. Participants are said to be frequent readers of the magazine with a personal interest in the topic of psychology in a broad sense. The magazine thus attracts a variety of people, practicing scientists included. The magazine's website listed the experiment for the period of its conduct.

FIGURE 27 – SOURCES OF A) VISITORS AND B) PLAYERS



Of the ample opportunities of social media, Facebook and Twitter were used to spread information about the experiment. A dozen friends shared the link and helped to connect to a potentially extensive community. Of course, also own contacts took part in the experiment and thus contributed to the number of participants. The Twitter-based hashtags in use read "#16Spiele14Min" ("16 games, 14 minutes”) and “#ZeitdruckExperiment” (“Time Pressure Experiment”) not leading to any further reactions within the Twitter community. Looking at the numbers, Facebook indeed was more successful in attracting participants. Nonetheless, this attempt expanded the approach. It offered another way to stay connected to potential or actual participants to share information about the experiment.

Internet forums were contacted variously, in most cases by pointing to potential gains in the lottery after finishing the experiment, offering a ‘deal’. This strategy has turned out to be less successive since most of the forums that inform users about buying deals had no intention to support noncommercial incentives (three out of five). Thus, forums that focus on information sharing remained. Among those three were contacted. Only one of them had terms of use that allows publishing such requests. In total, forums contributed very little to the number of participants.

Finally, the web address of the experiment could be found by search engines. After its first week online it reached its top rank three on the first page of the Google listing when looking for the word combination "Zeitdruck Experiment Online" (English: "time pressure experiment online") and was thus at least in the German-speaking internet clearly visible among websites with such rare content. This circumstance surely supported the fact that at least a handful of participants were attracted this way. The realized participation numbers are shown in Table 28. The table also informs about the subject's perseveration within the rounds of experiment. It is quite notable that well over a third of the persons who answered the questionnaire refused their participation already after going through the introductory part (97/271). Of those who started the experimental part about 60 % kept up until the end (105/174). Among the three versions of experiment, no significance in dropping rates could be determined.[152]

Recall that all sources provided an individual starting page link. Except for the ones that were spread via email to the university-based mailing lists, the links were embedded on websites and in social media. *Google URL Shortener* service created all embedded links. This proceeding offered access to the *Google Analytics* service in its basic, free version. Here, the date-time stamp, numbers, operating system and even the country-based location of page views could be examined for any link. The experiment's questionnaire page already collected this kind of information, except for the country of origin. Since *Google Analytics* did not provide the calling IPs, the page view timestamps were a sufficiently deterministic measure to link the information from *Google Analytics* with individual data from the experiment. The webpages traced subjects' behavior beginning with the questionnaire in the experimental procedures. However, it was possible to determine how many participants refused to go beyond the questionnaire by comparing differences between number of page views and number of players. The links were called 225 times in total. Compared to that, 205 players were identified giving a ratio of about 91 %. About their termination criteria, few things can be said. Nevertheless, one aspect is to mention. Comparing the calling times of termination data with those of participating data one finds predominantly termina-

[152] Hypothesis H_0: "$p_{DR,V_{e,i}} = p_{DR,V_{e,j}}$" and H_0: "$p_{DR,V_{e,i}} \neq p_{DR,V_{e,j}}$" were tested. The mean of dropping rate among the four rounds of experiment from experiment version V_e with $e, i, j := \{1,2,3\}$ and $i \neq j$ is p_{DR,V_e}. The corresponding p-values of the t-distribution with degrees of freedom $f = 6$ are between 0.56 and 0.82. Thus, H_0 cannot be rejected in any of the test cases.

tions in the group of people who called the link in their respective night hours[153] (32/64 = 50 %) whereas day hours callers had a refusal rate of (76/207 = 37 %). That points to night-time-related reasons for termination.

TABLE 28 – FIGURES OF PARTICIPATION AND PERSERVERATION

Participants	**%**	**Σ**	V_1	V_2	V_3
Calls:		n.a.	n.a.	n.a.	n.a.
- link		225			
- mailing list		n.a.			
Questionnaire:	156	271	91	78	102
- link		203	62	65	76
- mailing list		68	29	13	26
Start round 1	100	174	64	45	65
Start round 2	68	119	41	35	43
Start round 3	64	111	36	33	42
Start round 4	61	107	35	32	40
End experiment	60	105	34	32	39

The net processing time of 16 tasks (4 rounds of 4 games) across all game variations had a maximum of 14 minutes. Adding the duration of the instructions and self-chosen breaks between rounds gave the total experiment time a subject spent on the website. A command of the subject initiated the start of the next round. The maximum duration was thus not determined a priori. Nonetheless, the intensive, task fulfilling phase of the experiment was significantly shorter than the maximum attention span Costa-Gomes et al. (2001, p. 1199, footnote 14) presumed to be about 1.5 to 2 hours.

10.1.2 Participants

Participants who answered the questionnaire gave personal information in the following categories: sex, age, level of education, education on the topic of game theory, current job status and as far as they were soldiers, days of foreign assignments. Besides, the web form of the questionnaire additionally recorded information about the browser in use to derive the device applied for the

[153] Regardless of time zones, the following is determined for this study: night hours: 10 pm – 6 am, day hours: 6 am – 10 pm.

experiment.[154] This device information was especially relevant to interpret the mouse movement data correctly.[155]

While statements concerning age and number of operating days ('μ_{days}') are metric, the other criteria are measured on a nominal scale. Table 29 gives an overview of the participants' answers. Note that analysis only considers the information of the 174 people who started the experiment. Among those were ten players who used mobile devices and pads that worked by finger touching and not by mouse pointing. The 74 tasks they worked through represent about 3.8 % of all tasks played, which are 1,936 in total. The number of tasks fulfilled with mouse pointer seemed sufficient for movement data analysis. Despite that, the ten players' tasks offered valuable decision information, including chosen alternative and processing times.

According to the gathered data, the average participant was 24 years old, either a student or a graduated employee, with no experience in game theory and decision-making under conditions of foreign assignments. About one quarter more men than women participated in the experiment (100:74). The age spans between 17 and 46 years, with a median of 24 and a standard deviation of 4.7 years. As indicated in the description of the average participant, a vast number of participants was of a higher educational level (A-levels or graduated from high school/university: 166 out off 174 participants). Asking for their current job, most picked 'student' or 'apprentice'[156] to describe their status (78) followed by 'soldier' and 'other' (both 48). Since the mailing list of Helmut-Schmidt-University was used, which is one of two universities of the German Federal Armed Forces, the majority of participating soldiers were students at the same time (47). With this, the number of students in total was 125, representing the majority of all participants.

One needs to consider this case when analyzing the performance of the three job groups. Of those soldiers, three claimed experiences from foreign assignments with 35, 100 and 180 days respectively abroad. Their number was tiny within the group of participants and insufficient for deriving significant results. However, some interesting aspects are considered in the analysis part. Altogether, the group of participants seemed quite homogeneous regarding their schooling experience. Whether sex, age, and job had a significant impact on decision-making under time pressure

[154] Table 25 gives a complete overview of information criteria gathered from the questionnaire.
[155] For detailed information, also see Subsection 9.3.1.2 in this treatise.
[156] The word 'apprentice' serves as the translation for the German technical term 'Lehrling'. It represents people who absolve the German dualistic educational system of institutionalized job training that also includes extra-occupational specific schooling.

is subject of the data evaluation in the next section. Thus, groups were formed using the mentioned categories, and their degree of heterogeneity examined.

TABLE 29 – DEMOGRAFIC CHARACTERISTICS OF PARTICIPANTS

Participants	**Σ**	V_1	V_2	V_3
Sex (m/f)	100/74	29/35	27/18	44/21
Age				
- μ	24.6	23.4	26.4	24,6
- median	24	23	26	24
- σ	4.7	4.6	4.9	4.4
Level of education[157]				
- HS	3	-	1	2
- eHS	1	-	-	1
mR	4	1	1	2
Abitur	36	14	9	13
Studies	130	49	34	47
Education in game theory (y/n)	16/ 163	5/ 59	5/ 45	6/ 59
Current job				
- student / apprentice	78	26	21	31
- soldier	48	24	8	16
(of those: currently student)	(47)	(24)	(7)	(16)
- other workers	48	14	16	18
Foreign assignment (y / n / μ_{days})	3/171/ 105	-/64/ -	3/42/ 105	-/65/ -

Among the sexes, the perseveration rate measured between answering the questionnaire and starting the experimental part differed about 12 % (male 66 % / female 54 %). Feedback from participants that could shed light on this issue was infrequent, not to say singular. Up to the end of the experiment only one participant answered the feedback request presented in the footer of the webpages. The second day the experiment was online, this inquirer reports of difficulties understanding some termini used within the introductory part. Her concerns lead to an immediate ad-

157 The subjects were asked to state their current or intended level of education, leading to the fact that students mainly name university / high school degree in this category. The data further depicted educational degrees from German education system. As far as such comparison is permissible at all, an English expression is given with the following list (Abbreviation in Table 29/ degree (German) / degree (English): HS / Hauptschulabschluss / Basic Secondary School Leaving Certificate; eHS / erweiterter Hauptschlussabschluss / Extended Basic Secondary School Qualification; mR / mittlere Reife / Secondary School Leaving Certificate; Abitur / A level.

justment of the affected webpages, improving the quality of that part. However, she did not terminate the experiment earlier because of the described problems. The reasons for differences in perseveration rate thus remained unclear. No specific hints for a proper explanation were found in the current set of gathered information. However, perseveration rates in the experimental part showed no significant differences between male and female subjects.[158]

10.1.3 General Data Statistics

The statistical analysis of experimental data is described in the following. Here, commonly known analysis methods that are applied to the hypotheses from Subsection 9.1.2, are given.[159]

Without further limitations, the data streams of the individual participants are suggested to be independent. That is especially the case when participants independently fulfill their tasks, without communicating with each other. Since the experiment was conducted online, independence in the sense that participant individually faced the tasks without knowledge of the moves of other participants and without influencing the behavior of other can be widely assumed.

Given the task variables time limitation t_i , complexity c_i , game type g_i , and the experimental design version V_e with the subscripts determining their realization in task i and experimental design version e. The integer i is running from 1 to 16 and m is an integer running from 1 to 3. Whereas time limitation varies with i, complexity is constant for every four increments of i, forming one round of tasks each with four rounds in total. The response data then is formed by the experimental mouse movement data stream. It contains information about mouse pointer coordinates together with time stamps, provided by recorded mouse events such as clicked or hovered elements. This information, in turn, is used to measure EMMA's and mouse movement history. By applying the interpretation metric to this information, the primarily focussed EIPs are determined. Also, the performance evaluation concept presented in simulation (Section 5.1) is adapted. Choice data complement this measure. This proceeding is applied in the following to evaluate the participants' decision-making patterns and the related hypotheses. The results are in a next step used to classify behavior.

[158] H_0: "Means of male and female perseveration rates are identical". H_1: "Means are different". The p-value of a two-sided t-Test with degrees of freedom $f = 6$ is 0.92. Recall, that $\propto = 0{,}95$.
[159] The underlying mathematics are described in Fisz (1989) and Bronstein et al. (2012) for example.

10.2 Data Evaluation

In this section, the experimentally gathered data is evaluated by testing the proposed hypotheses. An explanation of test procedure is given in Appendix E. In addition, other behavioral aspects are presented that were identified in the context of decision making but were not anticipated by the results of the preparation time model.

The number of experimental tasks represents the sample size of the test statistics. In case of certain tasks are relevant for the hypotheses' test procedure, they are explicitly specified. For hypothesis testing in general only tasks are considered that show mouse pointer-application to the extent that mouse movement is recorded. Three main reasons are identified why this requirement is not met. The first is the use of input devices other than a mouse pointer such as touchpads (10 of 174 players affecting 74 of 1,936 datasets). In this case, the subject's behavioral data stream provides no movement information. The second one is an early termination of the task by the subject with the effect that minimal mouse movement is observed (10 of 1,936 datasets). One can observe the third reason more frequently in experiment version V_2 than at V_1 or V_3. Subjects in that case neglect mouse use for reasoning tasks. They only use it for choice with the consequence that one cannot recognize any mouse movement within the payoff cells.

When testing Hypothesis X, all datasets – including those with an inadequate mouse usage – are used for the analysis. Results of the specific participant groups are introduced together with each hypothesis test of the comprising dataset.

All hypotheses are discussed in the following. Differences among the experiment versions occur throughout all hypotheses. They are presented along with their discussions. The analysis is organized twofold, examining the influence of task variables – time limit and complexity – separately. Figures of either proportions or means of the examined aspects are given for all hypotheses. The results of each experiment version are plotted in an individual frame. As a reference, values of decision-making without time pressure are presented as dashed lines. Values resulting from decision-making under time pressure conditions are presented as full lines. Vertical dashed gray lines section data according to categories of the underlying perspective: in the case of viewing the influence of complexity, the first section on the left side contains data from all four levels of complexity. The time limitation is ineffective at this section. It decreases (i.e., gradually becoming severer) per section from left to right. When focussing on the influence of time limitation, one section comprises data of one level of complexity. Time limitation decreases within one section from left to right. Time restrictions in figures are labeled analogously to the description in the simulation part of this treatise with ‘ineffective’ – ‘moderate’ – ‘strong’ – ‘severe’.[160] The corresponding p-value tables provide the exact limit in seconds. For all hypothesis tests, those tables are given together with other test data of all versions in Appendix F.

10.2.1 Evaluation of Hypothesis I

This hypothesis represents the general sensitivity of decision-making behavior to time pressure. This sensitivity is assumed to influence a subject’s performance in decision behavior. Analogous to the proceeding presented earlier in Section 5.1, performance is measured in the four categories ‘Decision-making’, ‘Reduction of Alternatives’, ‘Payoff Generation’ and ‘Amount of EIPs’. It is sufficient to show the sensitivity of at least one component to prove overall sensitivity. Contributing hypotheses are Hypothesis II, III, and VI testing the categories ‘Payoff Generation’, ‘Reduction of Alternatives’, and ‘Amount of EIPs’, respectively.

The aspect ‘Decision-making’ is evaluated by examining the share of decision makers in the population under varying time pressure conditions. The null hypothesis states the status quo – no significant change in proportions of decision makers within the population is observable. As the alternative, one expects any deviation. For evaluation, all datasets are examined, including those of insufficient mouse usage. The experimental version V_1 provides 664 datasets, V_2 578 and V_3 724. Plots are given in Figure 28 and Figure 29 whereas Table 60 and Table 61 (Appendix F) contain the corresponding p-values from the tests.

Influence of time limitation

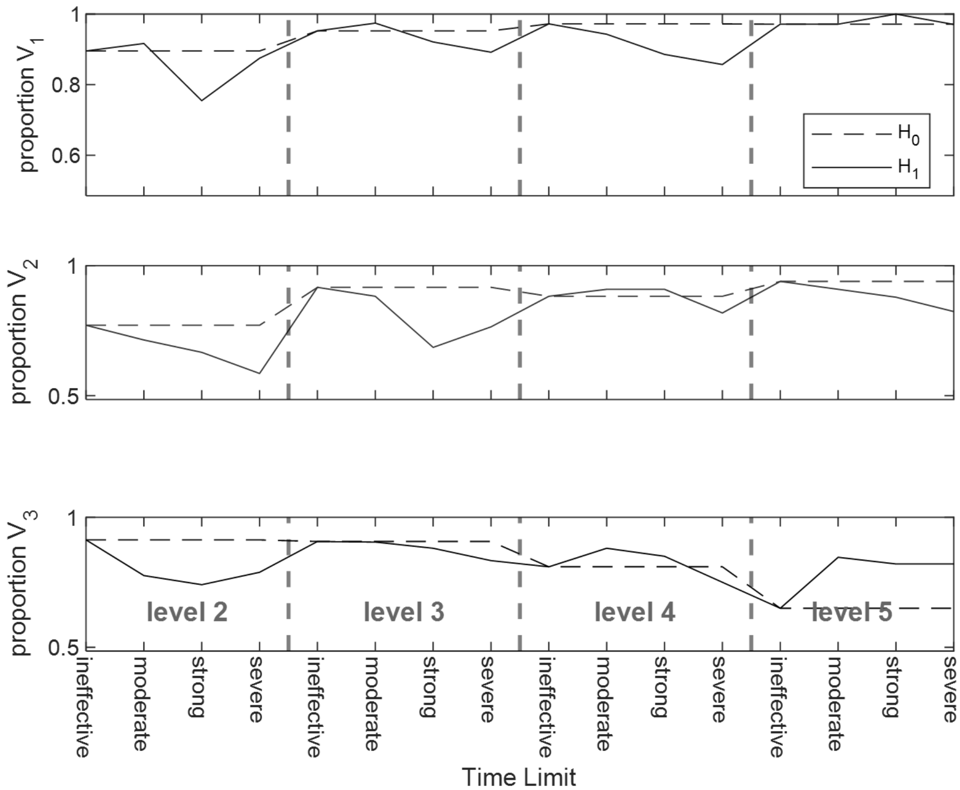
Deviations from reference values are in tendency larger when time limitation is strong or severe, as Figure 28 indicates. Even though the amount varies with the level of complexity, one can spot particular significance for such cases.

The p-values extreme enough to reject H_0 build a respectable minority among all cases (12 out of 36 cases for all experiment versions). Especially V_3 – for complexity levels two and five – is to name here, showing a significant impact of time limitation on the decision rate. This fact is illustrated by the corresponding ratios (Figure 28). However, in two-thirds of the cases, H_0 cannot be

[160] Compare designation in Subsection 9.3.2.2.

rejected. Hence, the choice rate is generally not sensitive to changing time limitation with V_3 as the exception.

FIGURE 28 – HYPOTHESIS I: INFLUENCE OF TIME LIMITATION ON CHOICE RATE

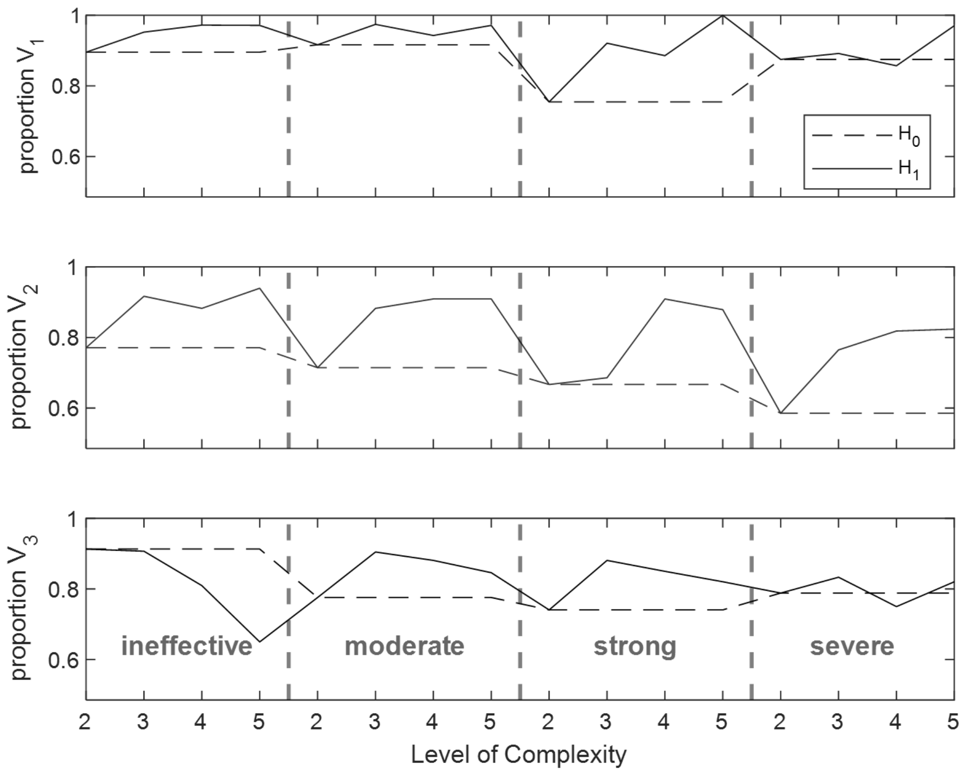


The frequently observed minimum in choice rate under strong and severe time limitation (Section 3.3) is in compliance with the preparation time model. However, in some of the game tasks, this implication is not met or not significant, raising questions of its generality. Participants' behavior thus needs to be further distinguished and classified.

Influence of complexity

The picture is slightly changing when determining the impact of complexity on the choice rate. In 14 out of 36 cases, H_0 can be rejected, with nine of them alone belonging to V_2. Figure 29 shows distinctly large deviations here. At this point, sensitivity is recognizable. The other two experiment versions also produce remarkable deviations, yet less frequently. It thus can be stated that they remain primarily insensitive. Differences in social aspects are not identified.

FIGURE 29 – HYPOTHESIS I: INFLUENCE OF COMPLEXITY ON CHOICE RATE



Together with the results of Hypotheses II, III, and VI, Hypothesis I can now be evaluated entirely. Table 30 consolidates the sensitivity results, with cells in light gray color depicting sensitivity and dark gray insensitivity. Two white colored cells can be found in the row of Goal 2 for experiment version V_3. Due to technical restraints in the experimental design, reduction of alternatives cannot be observed, and hence no data is available. Both cases are excluded from consideration. The last row contains the overall result for each experiment version regarding the observed effectiveness of time limitation and complexity, respectively. It is based on the combination of results of the contributing goals, depicted in the cells above in the same column. As intended earlier, sensitivity is determined when at least one contributing goal showed sensitivity. In this case, at least one of the cells in the column above the resulting cell is light gray. This requirement is satisfied for every column and hence general sensitivity determined.

TABLE 30 – HYPOTHESIS I: EVALUATION

Goal	**Time limitation**			**Complexity**		
(number of contributing hypothesis)	V_1	V_2	V_3	V_1	V_2	V_3
Goal 1 Choice rate (I)	insensitivity	insensitivity	sensitivity	insensitivity	sensitivity	insensitivity
Goal 2 Reduction of alternatives (III)	sensitivity	insensitivity	no data	sensitivity	insensitivity	no data
Goal 3 Payoff generation (II[161])	insensitivity	sensitivity	insensitivity	sensitivity	sensitivity	sensitivity
Goal 4 Reduction of EIPs (VI)	sensitivity	insensitivity	sensitivity	sensitivity	insensitivity	sensitivity
Hypothesis I General sensitivity	sensitivity	sensitivity	sensitivity	sensitivity	sensitivity	sensitivity

Legend:

sensitivity	insensitivity	no data

10.2.2 Evaluation of Hypothesis II

Hypothesis II postulates an adverse influence of time pressure on the effectiveness of decision-making. Following the evaluation concept proposed in Section 5.1, Goals 1 to 3 determine effectiveness. Similar to Hypothesis I, different hypotheses contribute to those goals (see Table 21). That comprises Hypothesis III[162] (reduce alternatives) as well as a combination of Hypotheses IV and V (generate payoff) with similar-weighted importance. As can be shown, Hypotheses IV and V have disjunctive datasets, which together form the dataset to evaluate whether time pressure influences the ability to generate payoff. This evaluation is combined with the outcome of Hypothesis III to evaluate Hypothesis II.

In the following, the sensitivity to generate payoff is tested. Here, all records are considered that contain a decision. This selection explicitly includes records of all input devices. In case of V_1, 608 of 664 datasets are included in the analysis (V_2: 467 of 578, V_3: 596 of 724). Charts are presented in Figure 30 and Figure 31. Corresponding p-values are given in Table 62 and Table 63 (Appendix F).

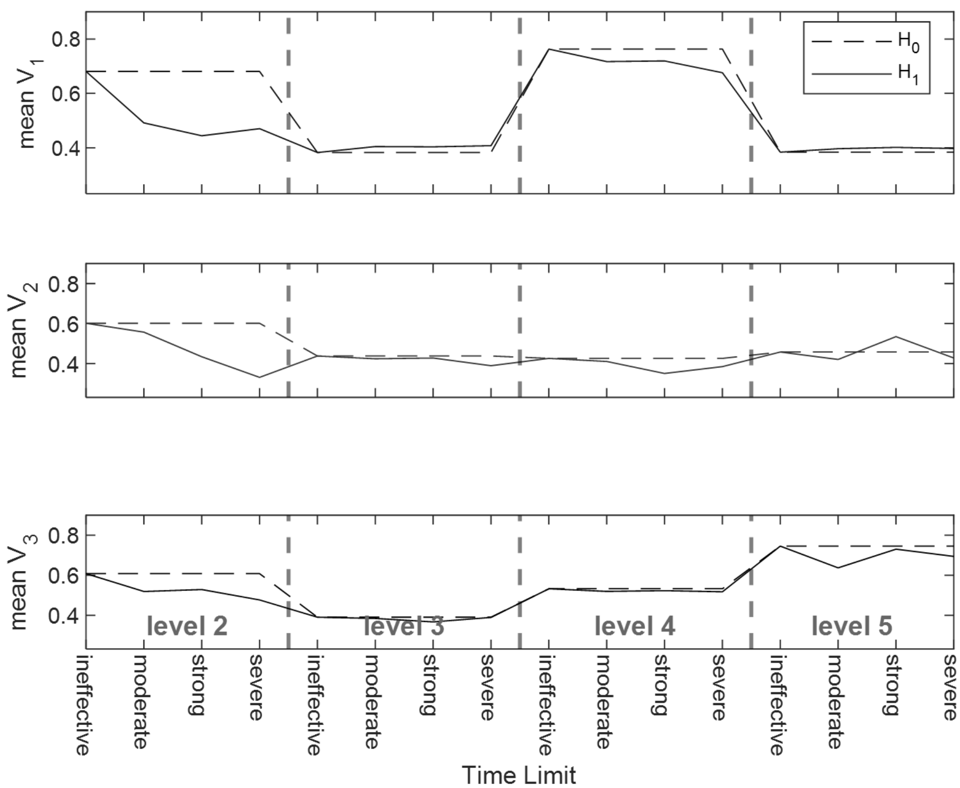
[161] Note that different hypotheses contribute to the evaluation of Hypothesis II. One of these hypotheses addresses the general sensitivity of payoff generation. The results of this hypothesis evaluation represent the evaluation of Goal 3 in this table.

[162] Only experiment versions V_1 and V_2 are concerned in Hypothesis III. Thus, the evaluation of Hypothesis II in case of V_3 is exclusively based on the corresponding results from testing the sensitivity to generate payoff.

Influence of time limitation

The graphs of the three experiment versions in Figure 30 show deviations from the reference values in the first section of low complexity. Here, the null hypothesis is rejected in eight out of nine cases, supporting the expected alternative hypothesis. In case of higher levels of complexity, such deviations diminish – the ratio of rejecting H_0 is rather small – pointing to a limited sensitivity of generating payoff towards time limitation. However, for the highest complexity level, in four out of nine cases H_0 can be rejected, even though in three cases an alternative hypothesis is supported that has an orientation as opposed to the one predicted.

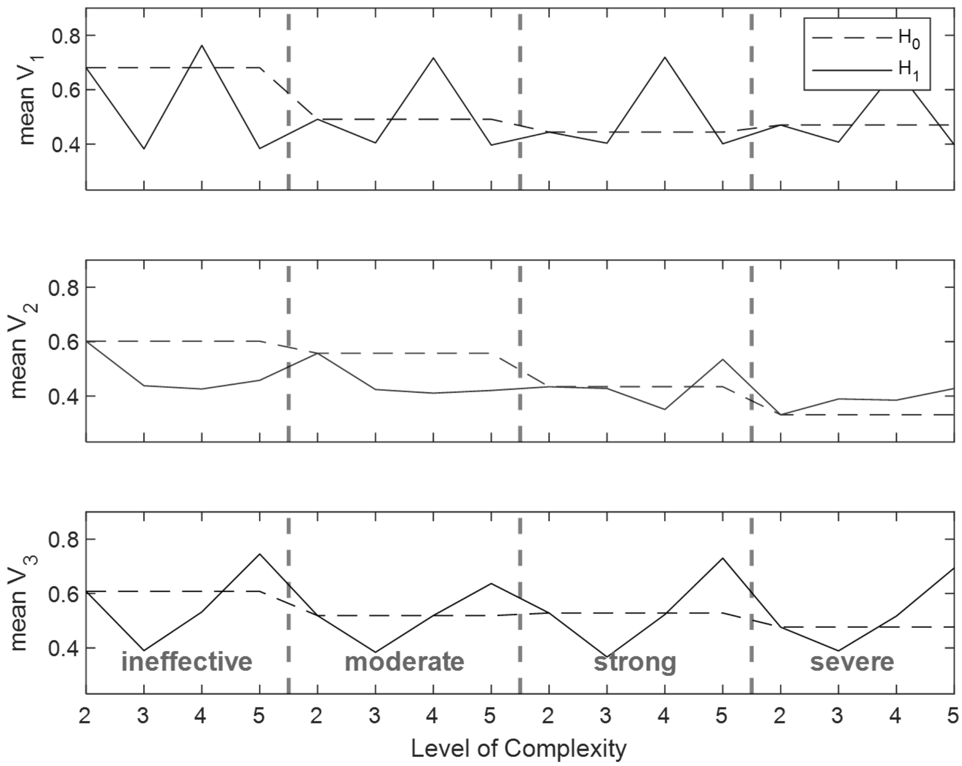
FIGURE 30 – HYPOTHESIS II: INFLUENCE OF TIME LIMITATION ON RELATIVE PAYOFF



Influence of complexity

In case of V_1, the null hypothesis is rejected in nine out of twelve cases. At this point, p-values are rather high for complexity levels of three and five and low else. The influence of the game type could be causative. Experiment versions V_2 and V_3 reject H_0 less frequently, even though their comparative values are similarly oscillating around the reference values (see Figure 31). However, their deviations seldom reach the level of significance.

FIGURE 31 – HYPOTHESIS II: INFLUENCE OF COMPLEXITY ON RELATIVE PAYOFF



Regarding social aspects, differences in behavior occur at experiment version V_3 for the lowest level of complexity: people, allegedly experienced in game-theory, significantly deviate in generating payoff from those who report no experience. For the first game – Costa 2A – the experienced players gain fewer payoffs in comparison to inexperienced ones. Costa 2A has a payoff-structure where the social payoff dramatically deviates from the solution of eliminating dominated alternatives. The latter is chosen more frequently by the unskilled player, hence generating more payoff. In case of the other three 2-by-2-games, the payoff structure is more favorable for equilibrium decisions, and the skilled players generate significantly more payoff than the others.

Also, results from V_3 indicate differences among occupational situations here. In 13 out of 16 cases, soldiers generate more payoff than students. In the remaining three cases, the opposite is true. However, only on four occasions, those deviations are significant, favoring soldiers in three and students in one case. The distribution of significant cases among the tasks shows no hint regarding sensitivity towards certain conditions.

As proposed, one can now combine these results with the one of Hypothesis III to evaluate Hypothesis II. Conclusions are based on the set of rules applied while evaluating Hypothesis I. The overview provided by Table 31 draws an almost coherent picture, indicating the correctness of this hypothesis: effectiveness, in general, is negatively affected by time pressure.

TABLE 31 – HYPOTHESIS II: EVALUATION

Goal	**Time limitation**			**Complexity**		
(number of contributing hypothesis)	V_1	V_2	V_3	V_1	V_2	V_3
Goal 2 Reduction of alternatives (III)	sensitivity	insensitivity	no data	sensitivity	insensitivity	no data
Goal 3 Payoff generation (IV+V)	insensitivity	sensitivity	insensitivity	sensitivity	sensitivity	sensitivity
Hypothesis II Negative influence on effectiveness	sensitivity	sensitivity	insensitivity	sensitivity	sensitivity	sensitivity

Legend:

sensitivity	insensitivity	no data

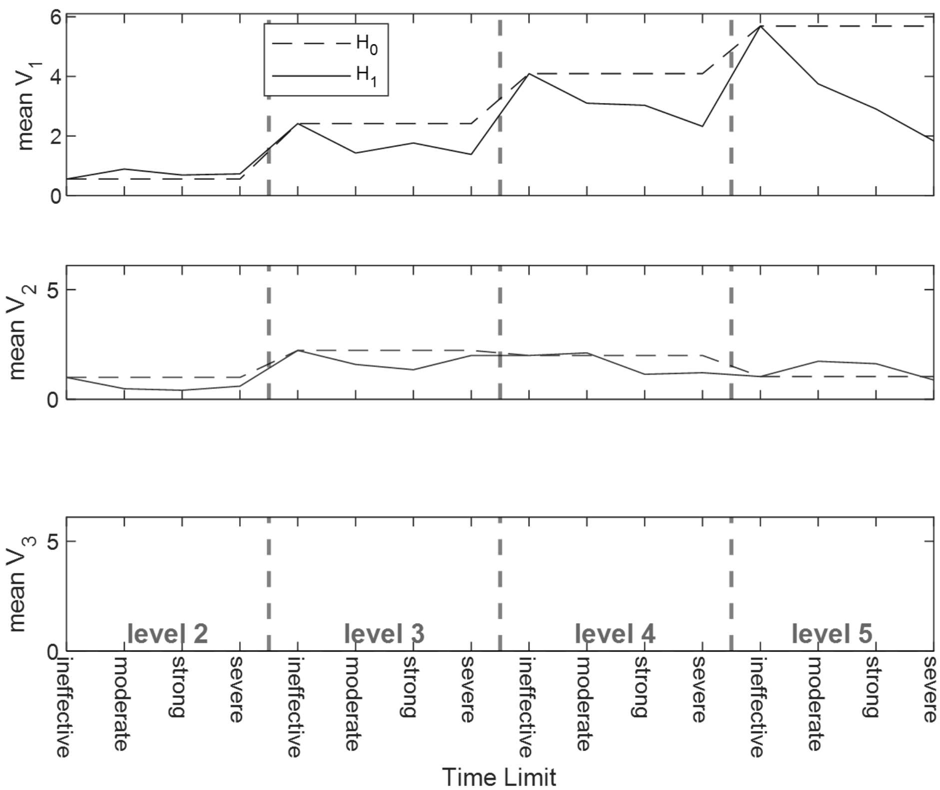
10.2.3 Evaluation of Hypothesis III

Hypothesis III claims an adverse influence of time limitation and a positive influence of complexity on the reduction of alternatives. The reduction of alternatives can only be observed for experiment versions V_1 and V_2. Due to the technical implementation of V_3 – cells are opened and closed by hovering over and out a cell – 'mouse clicks' are not recorded. However, this is necessary to identify a reduction. V_1 and V_2 exclusively implement such a function. The corresponding means of versions V_1 and V_2 are plotted in Figure 32 and Figure 33, whereas Table 64 and Table 65 (Appendix F) depict the p-values. The number of datasets of V_1 and V_2 considered for evaluation is 569 and 351, respectively.

Influence of Time Limitation

The impact of time limitation in case of V_1 is negative as Figure 32 illustrates. Still, it also notably depends on the level of complexity: for a small level, H_0 retains. Here, the mean of alternative reduction is almost constant under decreasing time limitation. In sections of larger levels of complexity, H_0 is very frequently rejected (seven out of nine cases). In case of V_2, p-values for small complexities start on a rather high level, partly inducing that the reduction of alternatives decreases – as expected beforehand. For growing levels of complexity p moderately decreases with H_0 far from being rejected. Still, the values generally remain larger than the ones of V_1.

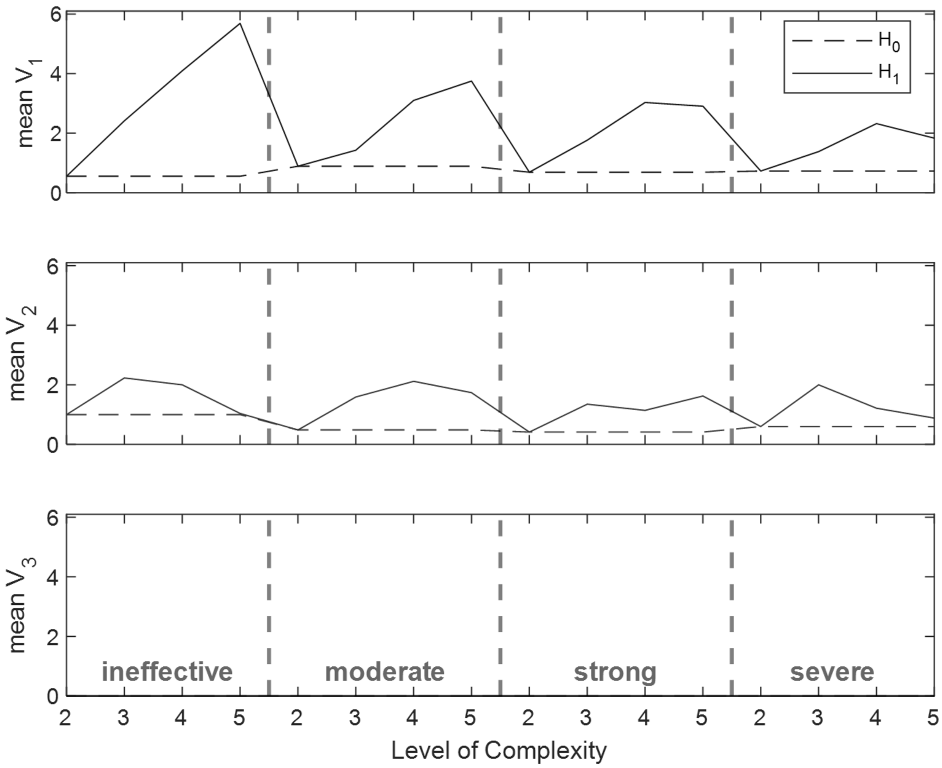
FIGURE 32 – HYPOTHESIS III: INFLUENCE OF TIME LIMITATION ON REDUCTION OF ALTERNATIVES



Influence of complexity

Testing the hypothesis for complexity in V_1, p-values are consistently very small with the consequence of rejecting H_0 in all cases. This rejection supports the alternative hypothesis which claims a positive influence of growing complexity – and thus growing number of occasions – on the reduction of alternatives. Experiment version V_2, in turn, provides p-values consistently smaller 0.5 but reaches the critical value just in five out of twelve cases.

FIGURE 33 – HYPOTHESIS III: INFLUENCE OF COMPLEXITY ON REDUCTION OF ALTERNATIVES



Altogether, one can notice that both experiment versions show tendencies of supporting the alternative hypothesis, yet only V_1 shows significant sensitivity for both task variables. The experiment version V_2 instead does not reach the necessary rejection rate for any task variable, even though in case of complexity, the values close by.

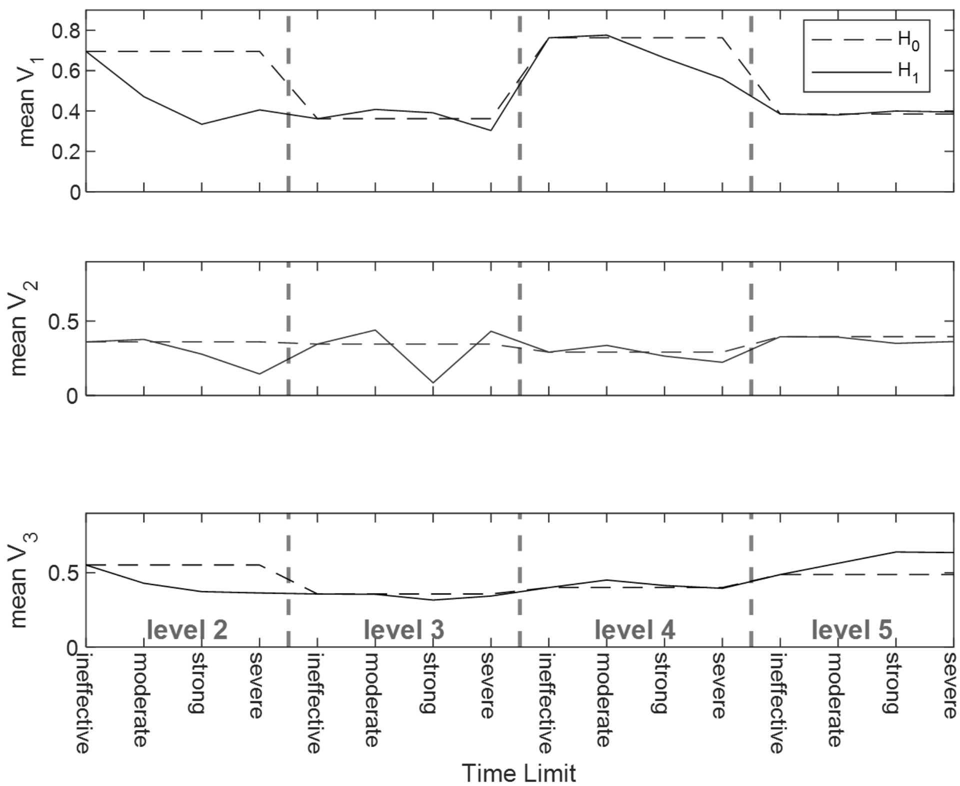
10.2.4 Evaluation of Hypothesis IV

This hypothesis questions the ability to generate payoff under the use of potentially strategic heuristics or elements of it. If time pressure increases, the payoff value should decline. Figure 34 and Figure 35 depict the means for all experiment versions. The corresponding p-values are listed in Table 66 and Table 67 in Appendix F. Several datasets are excluded from consideration to meet the requirement 'contain at least one strategic EIP COMPARE I'. Based on the dataset of adequate mouse users, V_1 has 355 of 569 datasets remaining for evaluation (V_2: 251 of 351 and V_3: 559 of 646).

Influence of time limitation

Means of V_1, as well as V_3 and to a minor extent of V_2, deviate significantly for all levels of time limitation when complexity is small. At higher complexities', p-values almost never pass the critical value (see Figure 34). In case of level-5 complexity, the p-values of V_3 are almost small enough to support an alternative hypothesis. However, its orientation is opposite to the one expected when the payoff is rising under decreasing time limit.

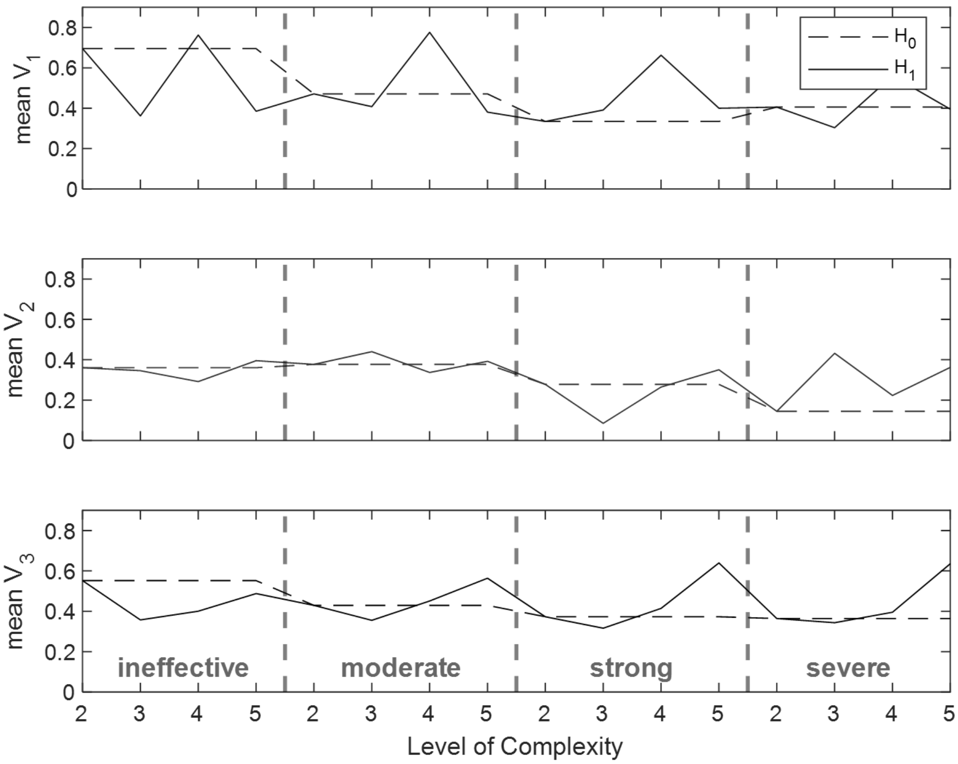
FIGURE 34 – HYPOTHESIS IV: INFLUENCE OF TIME LIMITATION ON RELATIVE PAYOFF OF COMPARE I-USERS



Influence of complexity

The picture is even more ambivalent when focussing on complexity. Throughout all experiment versions, H_0 is rejected for specific conditions. Figure 35 shows peak values of the means for a complexity of level four for V_1 and level five for V_3. Especially under severe time limit conditions, V_2 reaches notable deviations. Most of those mentioned deviations are significant and contrary to the alternative hypothesis. Support for the alternative is rare and can be found for ineffective and moderate time restrictions at V_1 and V_3 and in a single case also at V_2 when time limitation is strong, and the complexity level is three. In general, these results are not suitable to support the alternative hypothesis.

FIGURE 35 – HYPOTHESIS IV: INFLUENCE OF COMPLEXITY ON RELATIVE PAYOFF OF COMPARE I-USERS



Altogether, the stated hypothesis is in large parts not supported by the experiment data. The results of the three experiment versions are highly contradictory for the studied set of parameter. The hypothesis finds support at the smallest level of complexity when the time limit is decreasing. Under severe time limit conditions, an opposite alternative is partially backed. For other conditions, the sensitivity is not significant. One needs to examine those issues in connection with the results of the following Hypothesis V, as the complementary dataset, and Hypothesis II, which examines payoff generation for the combined dataset. Also, a dependency of the underlying game type can be expected: the means of V_1 and V_3 show for each section similar extrema (see Figure 35) for same levels of complexity. The plots in Figure 34 consequently show an almost constant trend for complexity levels of three and five (V_1) as well as three and four (V_3). This result indicates that time limits are almost ineffective at this point. In these cases, the same game type is applied in V_1 and V_3: at level three, players of both versions face the Chicken game. Subjects play the Hawk-Dove game at level five (V_1) and level four (V_3), respectively. In experiment version V_2, subjects play the Hawk-Dove game at level three, the Chicken game at level four and Stackelberg's Leadership game at level five. Constant values are found in the latter two games, indicating that choice remains almost the same among different time limit conditions. In case of the Hawk-Dove game, the situation differs.

Analyzing differences in behavior among social categories, similar results as reported for Hypothesis II occur: people educated in game theory realize significantly different payoff than those who are not. The effect is limited to V_3 at the smallest level of complexity. For the first game – Costa 2A – the experienced players gain fewer payoffs in comparison to new ones and generate more in the other three games.

10.2.5 Evaluation of Hypothesis V

This hypothesis evaluates the link between payoff generation and use of EIP patterns other than the strategic COMPARE I. According to simulation experiences, growing time pressure should have a strong impact on the ability to generate payoff – especially in the stage of severest time limitation. In opposition to Hypothesis IV, the datasets which at least provide one EIP of COMPARE I are excluded from consideration. Thus, V_1 has 214 of 569 datasets remaining for evaluation (V_2: 100 of 351, V_3: 87 of 646). Table 68 and Table 69 (Appendix F) list the corresponding p-values.

Influence of time limitation

The graphs of means show an entirely different behavior throughout all experiment versions (Figure 36). Whereas in case of V_1 values are close to the reference value of H_0, especially at V_2 and to a minor extent at V_3 the means are frequently below the reference value. This outcome is unexpected, yet significant as the corresponding p-values reveal (see Table 68).

Influence of complexity

Varying complexity leads to changing means as can be seen in Figure 37. V_1 and V_2 show for stronger time restrictions expectable deviations from the reference value. These are also significant (see Table 69, Appendix F). Values of V_3 surpass the reference mean for moderate time limits, but fall behind when the time limit decreases. Similar to Hypothesis IV, the influence of the complexity on the comparative values can be observed.

FIGURE 36 – HYPOTHESIS V: INFLUENCE OF TIME LIMITATION ON RELATIVE PAYOFF OF NON-COMPARE I-USERS

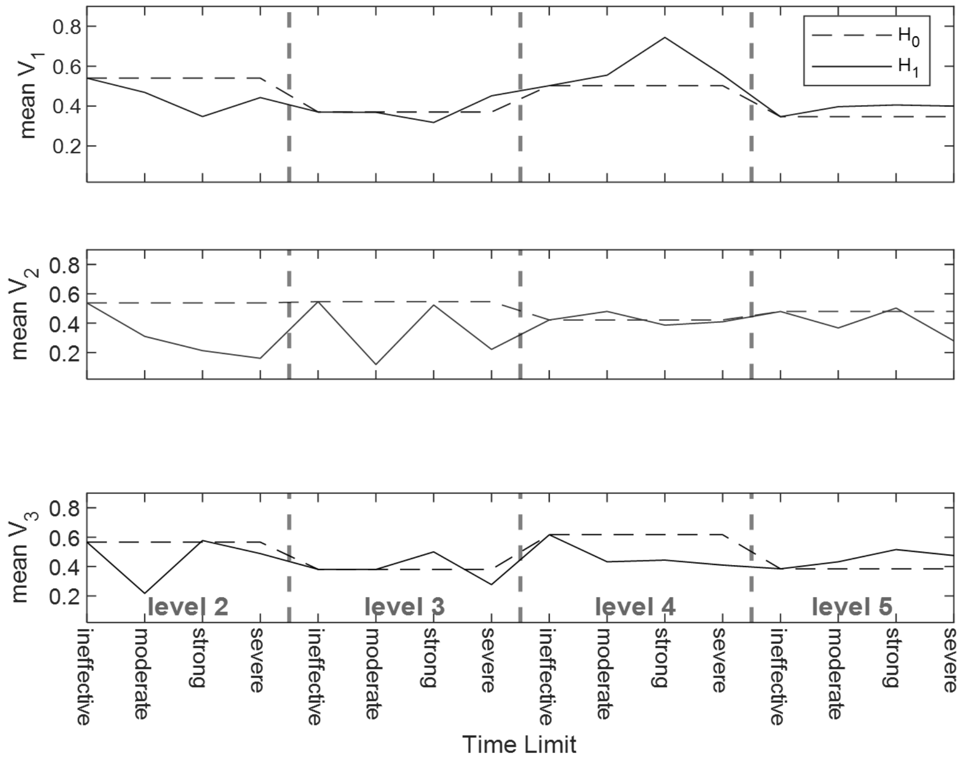
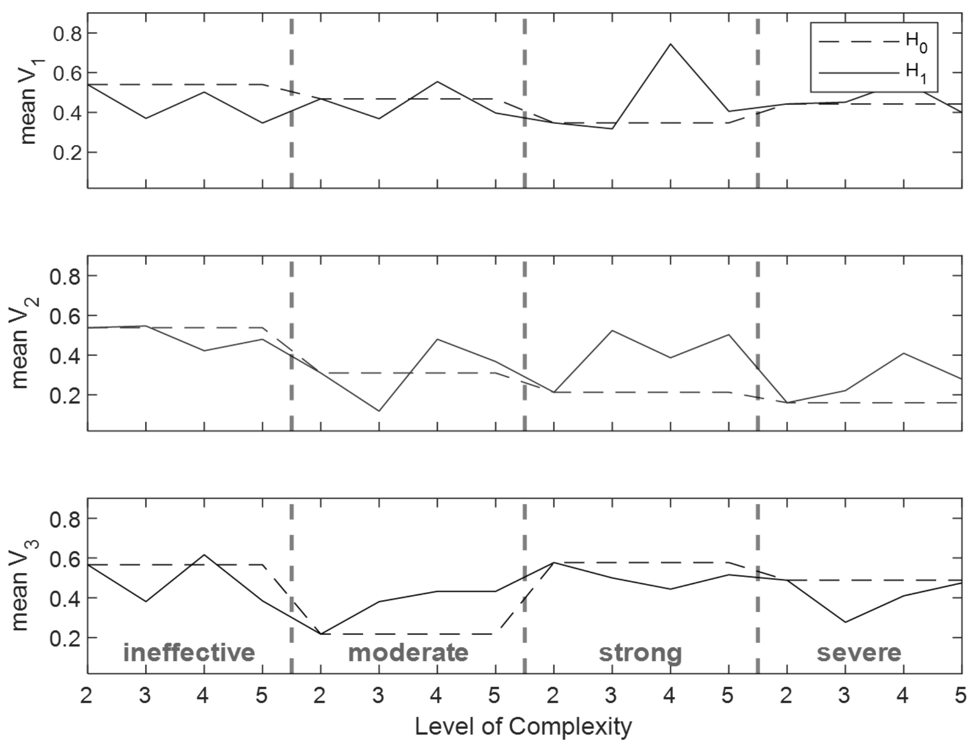


FIGURE 37 – HYPOTHESIS V: INFLUENCE OF COMPLEXITY ON RELATIVE PAYOFF OF NON-COMPARE I-USERS



Analogous to the previous hypothesis, the data partly confirms affectedness of generating payoff by time pressure. Still, Hypothesis V is backed up by the experimental data just in 5 out of 72 cases, all versions and tests for both task variables taken together. The alternative hypothesis is supported in 8 out of 72 cases. In the overwhelming majority of 59 cases, H_0 retains. However, especially the mean values need to be treated very carefully here since the database is extremely small due to the filter restrictions of this hypothesis. Appendix F provides the corresponding numbers of evaluated tasks.

Regarding social characteristics, similar results as reported for Hypothesis II and IV occur on V_3 at the smallest level of complexity: game theory-trained participants generate significantly different payoff than those who are not. For the first game type – Costa 2A – the experienced players gain lower payoffs in comparison to new ones and realize higher payoffs in the other three.

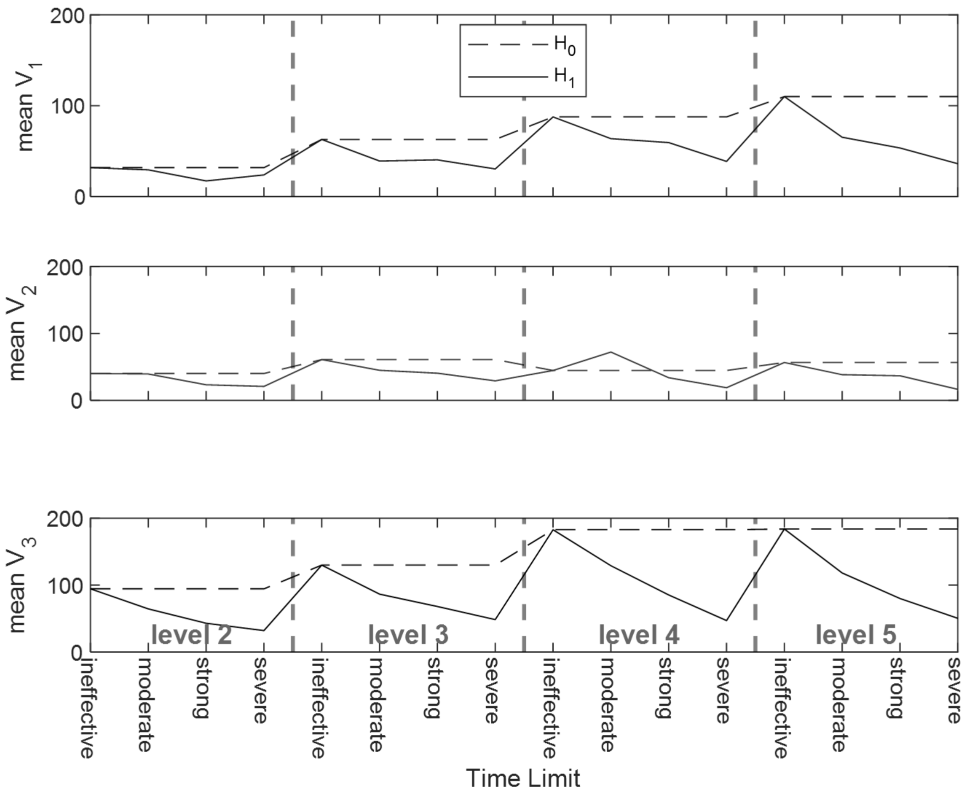
10.2.6 Evaluation of Hypothesis VI

The idea of this hypothesis is based on the assumption that time pressure negatively influences the processing capacity per time unit which in turn potentially reduces the number of EIPs during the process. The null hypothesis, representing the number of EIPs in case of no time pressure, serves as reference value for the values obtained under time pressure conditions. The analysis of this hypothesis is limited to subjects who drew a decision and used their mouse appropriately. This limitation leads the size of the dataset as 569 out of 664 for V_1, 351 out of 578 for V_2 and 646 out of 724 for V_3 in numbers of fulfilled tasks. In Figure 38 and Figure 39 the means are plotted. Table 70 and Table 71 (Appendix F) list the corresponding p-values.

Influence of time limitation

Throughout all experiment versions, H_0 is rejected quite frequently, giving support for the alternative hypothesis. Means are openly deviating from the reference value throughout all complexities with minima predominantly occurring at severe time restrictions (see Figure 38). In the eight cases where H_0 cannot be rejected, time restrictions are either moderate or strong. Subjects show no significant decrease in the number of EIPs here. This outcome could either result from an ineffective time limit or from an acceleration of processing speed here. The latter aspect is discussed in more detail in Subsection 10.2.11 where a classification is presented. Seven out of eight tasks where H_0 is valid belong to V_2. This circumstance highlights the impact of the experimental design once again.

FIGURE 38 – HYPOTHESIS VI: INFLUENCE OF TIME LIMITATION ON MEANS OF EIPS



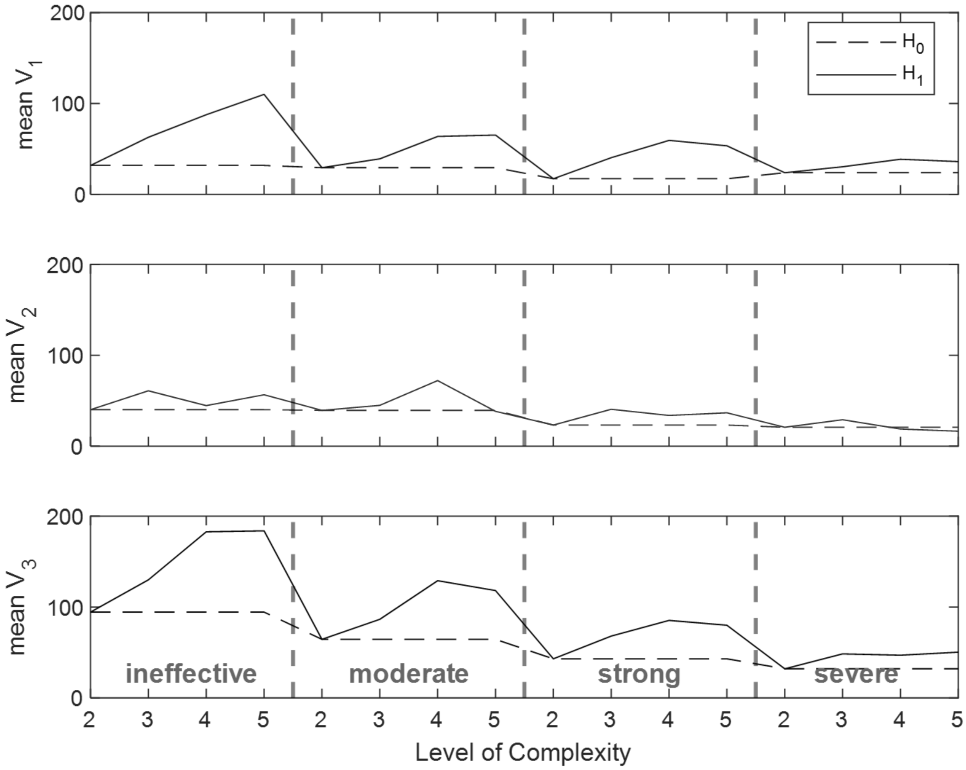
Influence of complexity

The means show a clear positive deviation from the reference value when focussing on complexity. Complexity levels of four and five present peak values. The p-values are expectedly small for V_1 and V_3, whereas the data of V_2 frequently produces values that support H_0 (8 out of 12 cases). At this point, it is to be noticed that V_1 and V_3 perform quite similar.

The p-values and the plotted means almost reflect the significant influence of time limitation and complexity on the number of EIPs entirely. However, the effect of limiting time on the numbers of applied EIPs is in contrast to that of complexity.

The subjects' occupational situations seems impact the number of applied EIPs: in 14 out of 16 cases employed participants invest more EIPs in the decision-making process than students. However, this fact is significant in just four cases.

FIGURE 39 – HYPOTHESIS VI: INFLUENCE OF COMPLEXITY ON MEANS OF EIPS



10.2.7 Evaluation of Hypothesis VII

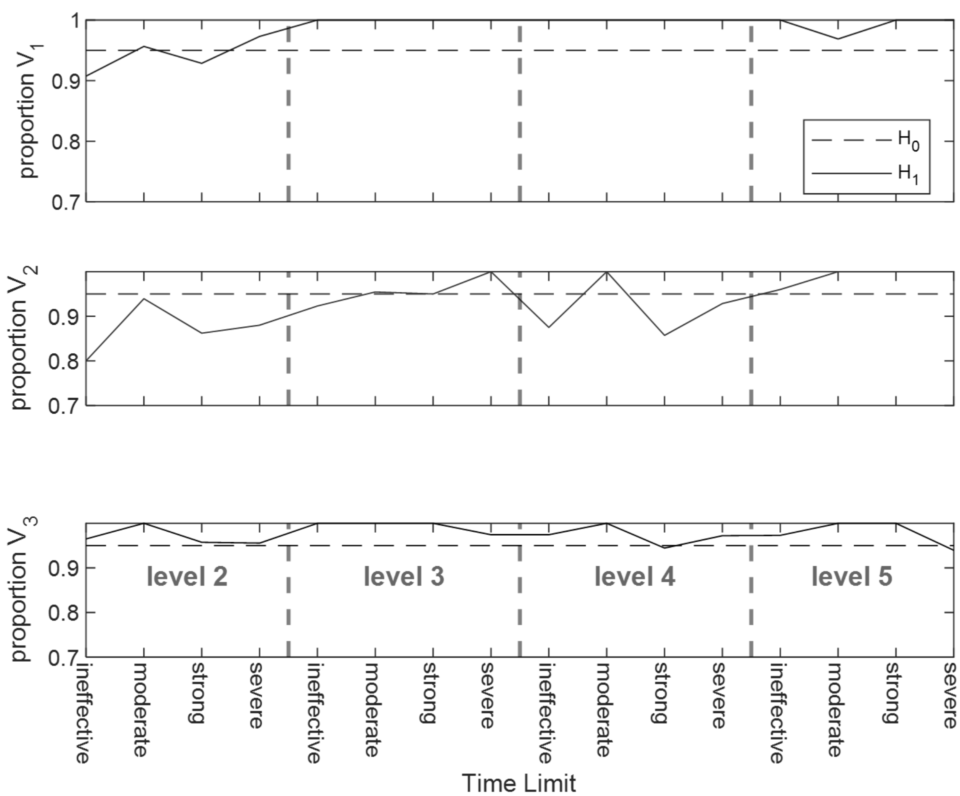
This hypothesis claims the application of heuristics or at least parts of it for decision-making under any time-pressure conditions. The proportion of usage can be assumed to be over zero percent as some decision makers have learned common heuristics within their training. Additionally, even those who did not receive such education might intuitively apply decision-making tools of their original repertoire, as the studies of Payne et al. (1988), Costa-Gomes et al. (2001), and Johnson et al. (2008) imply. The particular number is thus expected to be high – yet not necessarily equal to 100 %. It is quite possible that subjects reduce mouse usage during their mental decision-making process to an individual absolute minimum. The opposite is conceivable, too, when participants produce so much noise in their movement that the context can no longer be traced adequately by the present tracing framework. Furthermore, the selected set of heuristics could be incomplete, lacking certain EIPs in the model.

The configuration of the test statistic analogous to Eq. (15) requires a reference proportion $p_0(V_e) < 1$ to be valid. Due to that, this hypothesis cannot be tested in the way proposed in Subsection 9.1.2.6. The null hypothesis H_0 is thus adjusted, claiming a proportion of heuristics application among the experiment population of 95 %. This rate is assumingly lower than the rate the alternative hypothesis claims. In the test procedure, this value will be approximated by its boundary value, keeping the α-value claim valid. For V_1, V_2 and V_3 569 datasets, 351, 646, respectively remained for analysis after excluding inadequate mouse users. A separation of results regarding time pressure variables cannot be given since all p-values are based on a reference proportion of $p_0 = 0{,}95$. Corresponding proportions are plotted in Figure 40, whereas Table 72 (Appendix F) lists the p-values.

Results

Throughout all three experiment versions the shares are higher than or equal to 0.8. In 23 out of 48 cases it reaches 1.0 with most cases found at V_1 (11 cases), as can be seen in Figure 40. However, H_0 is rejected in three cases – all of them belong to V_2 – when the shares are significantly less than the reference value. It is not surprising that V_2 is affected since this version is supposed to motivate mouse usage the least effective with its cells being uncovered right from the beginning.

FIGURE 40 – HYPOTHESIS VII: INFLUENCE OF TIME LIMITATION ON PORTIONS OF HEURISTIC PATTERN USERS



Confidence intervals for proportions of subjects who apply heuristics or at least parts of it are presented in Table 32. All records that fit the regulations of Hypothesis VII are included – regardless of tasks and time pressure conditions. Features in social aspects cannot be reported.

TABLE 32 – CONFIDENCE INTERVALS OF HYPOTHESIS VII

$\boldsymbol{CI(V_1) = 0.98 \pm 0.01}$
$CI(V_2) = 0.92 \pm 0.03$
$CI(V_3) = 0.98 \pm 0.01$

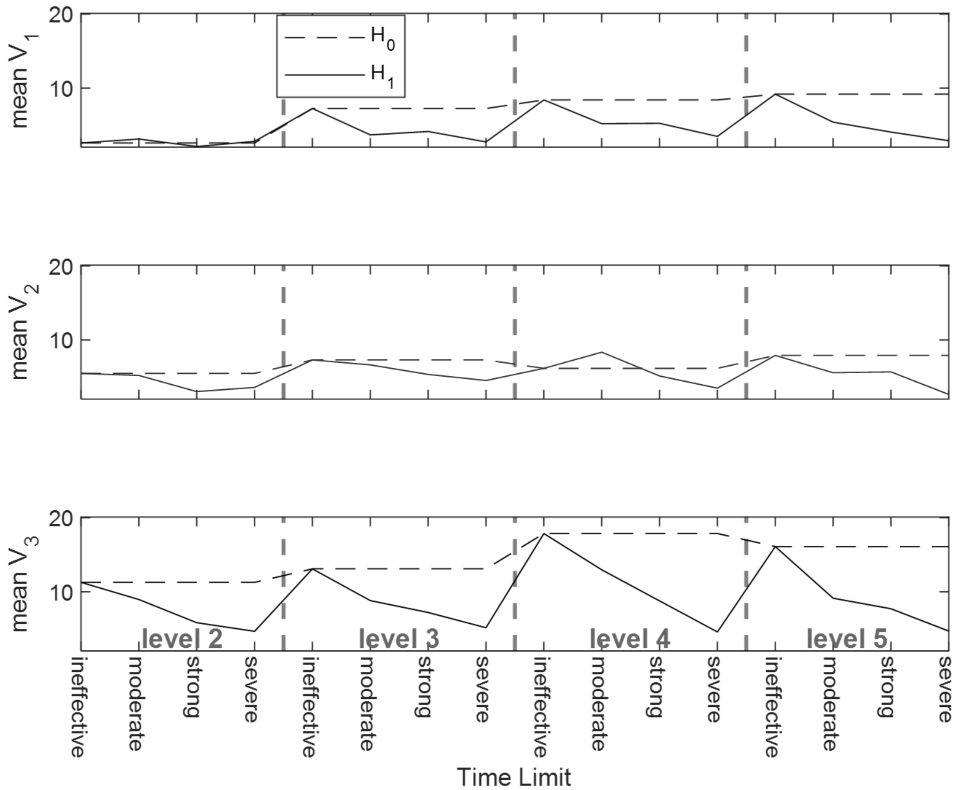
10.2.8 Evaluation of Hypothesis VIII

This hypothesis questions the use of heuristic patterns and its development under changing time pressure conditions. It postulates a negative impact, with growing time pressure evoking a decreasing amount of patterns. All records are included that show adequate mouse usage and no premature termination of tasks (V_1: 569, V_2: 351, V_3: 646). The corresponding means are presented in Figure 41 and Figure 42. Table 73 and Table 74 (Appendix F) list the associated p-values.

Influence of time limitation

The plotted means widely expose the assumed behavior. With two exceptions, comparative values are less than the reference value. The hypothesis H_0 is rejected in 29 out of 36 cases among all experiment versions and time pressure conditions. The p-values are consistently high in such cases. At the lowest level of complexity, the strongest support for H_0 occurs with four out of nine cases. Data of V_1 is at this stage almost similar to the reference value (see Figure 41). Among all experiment versions, V_2 shows the weakest support for an alternative hypothesis, retaining H_0 in eight out of twelve cases. However, mean values decline in most cases. One exception occurs for the complexity of level four at moderate time restrictions. All other time pressure conditions offer an almost similar picture, supporting the assumed alternative hypothesis.

FIGURE 41 – HYPOTHESIS VIII: INFLUENCE OF TIME LIMITATION ON MEANS OF HEURISTIC PATTERNS



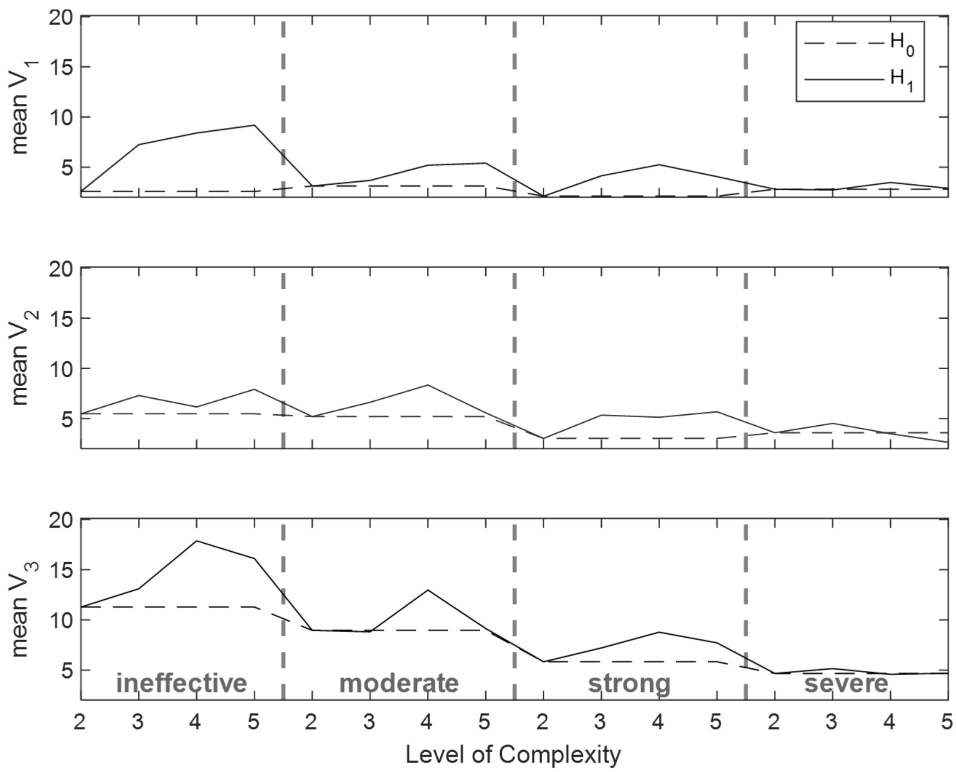
Influence of complexity

When focussing on complexity, the dropping rate of H_0 markedly decreases in comparison to the time limit effect (18 out of 36 occasions). In cases of rejecting H_0, the p-value is small, implying a positive effect on the number of employed heuristic patterns with growing complexity. When comparing the different time pressure sections, it becomes evident that with increasing time restrictions the scale of the mentioned effect declines. For severe time limit conditions deviations are hardly significant. However, the data of V_2 again show the least frequent significance among all experiment versions.

The sensitivity to time pressure is apparent here. Time limitation's influence is overwhelmingly negative, with some exceptions at moderate levels where the influence is not significant. Complexity, in turn, shows a less obvious impact, especially for the data of V_2 and for extremely narrow time limits. However, it affects the number of observable heuristic patterns positively and is thus in opposition to the time limiting effect.

Differences regarding social aspects are observed for age at V_3. In 13 out of 16 cases, the cohort aged 26 to 33 years used more patterns for their decisions than the younger group of 17 to 25 years old participants. This deviation is significant in four cases.

FIGURE 42 – HYPOTHESIS VIII: INFLUENCE OF COMPLEXITY ON MEANS OF HEURISTIC PATTERNS



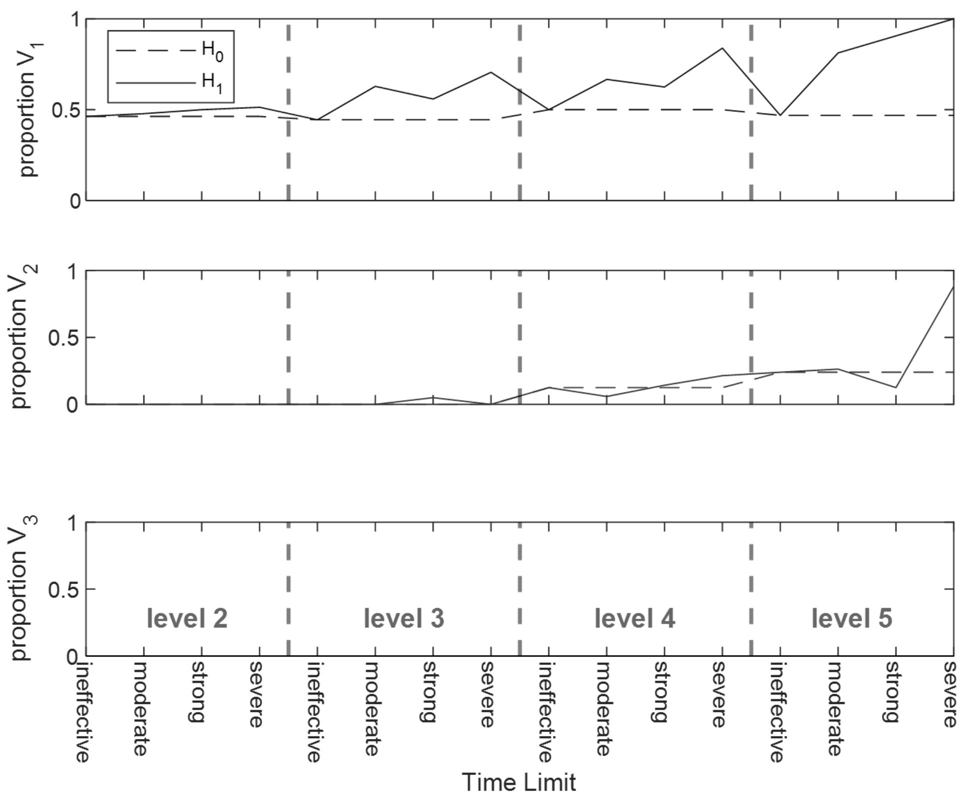
10.2.9 Evaluation of Hypothesis IX

This hypothesis inspects the share of random choices among the players' population under changing time pressure conditions. Positive influence is assumed. For evaluation, records are used that show adequate mouse usage and no premature termination of tasks (V_1: 569, V_2: 351, V_3: 646). Corresponding shares can be found in Figure 43 and Figure 44. The p-values are listed in Table 75 and Table 76 (Appendix F). For V_3 random choices cannot be detected with the methods applied. The share is hence stated as constant zero and test data for this case are not available.

Influence of time limitation

Time limitations show no significant impact on the smallest level of complexity. With growing level, shares deviate positively, and H_0 is rejected more frequently – starting with cases of severe time limitation. The p-value is at such points as small as expected from the alternative hypothesis. The complexity's impact on the scale of deviation is evident for V_1. For V_2 comparative values and reference values are quite similar, except for the highest level of complexity and severe time restrictions where share differs dramatically.

FIGURE 43 – HYPOTHESIS IX: INFLUENCE OF TIME LIMITATION ON PROPORTIONS OF *RANDOM* CHOICE

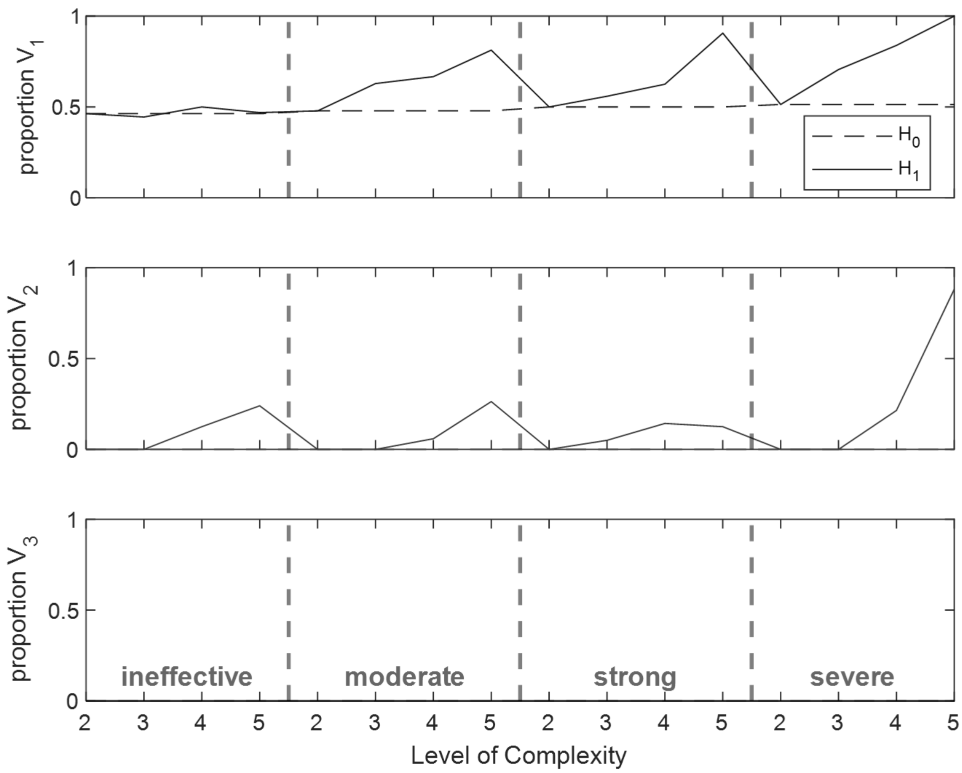


Influence of complexity

The hypothesis H_0 is rejected comparatively often. With time limits becoming more effective, the frequency of such cases increases – beginning with higher levels of complexity. Similar to the time limiting-evaluation, the p-values are entirely small when H_0 is rejected, supporting the stated alternative hypothesis. Peaks occur for the highest level of complexity.

The positive impact of time pressure on the ratio of random choices is quite clearly observable. Time limit and complexity show the same tendency here, even though significance is usually not recognized until time pressure conditions are extreme. Social influences cannot be observed for this hypothesis.

FIGURE 44 – HYPOTHESIS IX: INFLUENCE OF COMPLEXITY ON PROPORTIONS OF *RANDOM* CHOICE



10.2.10 Evaluation of Hypothesis X

This hypothesis examines the development of the equilibrium-choice proportion under time pressure. The null hypothesis postulates no changes of equilibrium-choice proportions under growing time pressure, whereas the alternative hypothesis assumes decreasing proportions. For the evaluation of H_0, all available records are in use. Herewith, 664 records of V_1, 578 of V_2 and 724 of V_3 are analyzed.

Before evaluating Hypothesis X, one needs to examine whether the proportion of equilibrium choices can be explained by chance.[163] For this, datasets are limited to those where records provide a choice. In case of V_1 608 are included in the analysis (V_2: 467, V_3: 596). Table 77 (Appendix F) presents the p-values of the corresponding hypothesis test.

For all three versions, there is usually one alternative designated as equilibrium choice. The expectable proportion of a randomly selected alternative is thus one by the number of alternatives

[163] This fact can be assumed when the proportion of equilibrium choices realized in the experiment is equal or close to the proportion of equilibrium choices one would expect when alternatives in a task are randomly selected with equal probability.

per player and task. When examining the influence of complexity, the same proportion is expected. One notable exception is found in versions V_1 and V_3 at a complexity level of three. Here, all alternatives describe equilibrium choices, making the proportion equal to 100 %. The expectable proportion further serves as reference proportion according to Eq. (15), with a null hypothesis postulating a proportion of equilibrium decisions equal to a random choice. The alternative hypothesis assumes a significant deviation from the portion of random choices, referencing intentional equilibrium decisions.

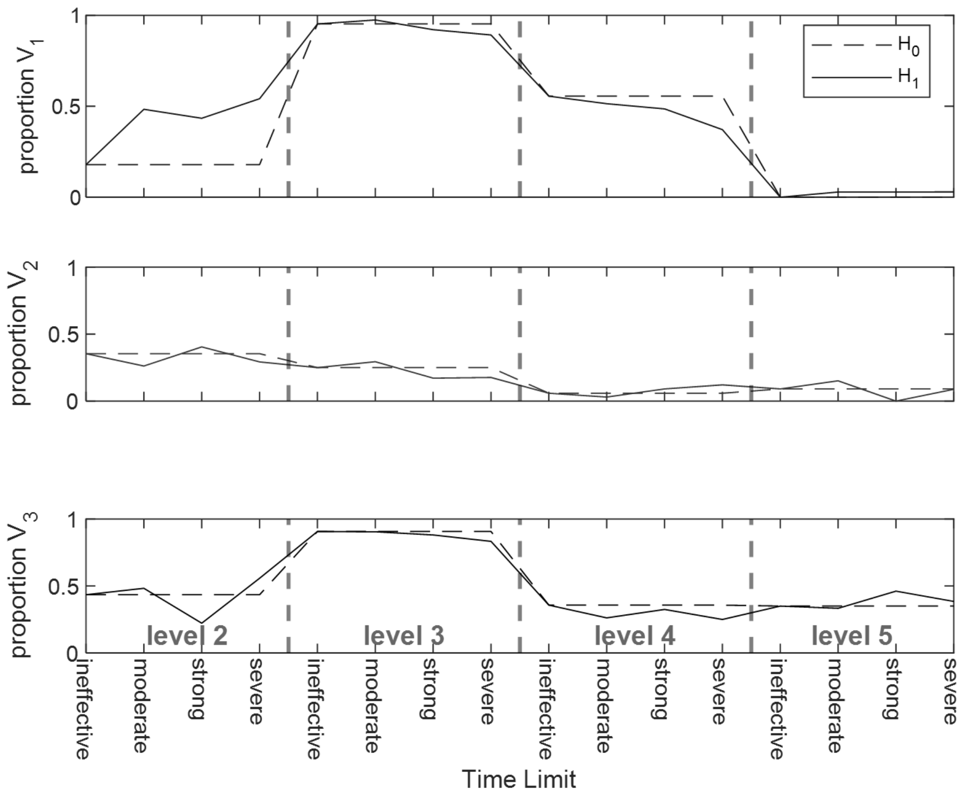
Looking at the p-values, it is remarkable that data from V_1 and V_3 mostly reject H_0, with five exceptions in total out of 24 occasions. It thus can be deduced that versions V_1 and V_3 support the alternative hypothesis to a large extent. Herewith, equilibrium choices are declared as intended rather than randomly made. However, H_0 mostly cannot be rejected under restricting time limits. This result indicates that choice, in general, is somewhat random under such conditions. The experiment version V_2 shows p-values consistently equal to and larger 0.5, yet the score is mostly lower than the critical value. At this point, H_0 is rejected just in 4 out of 16 cases. Hereby, equilibrium choice rate at V_2 is predominantly indistinguishable from random choice, except at complexity of level four. When examining the sensitivity of equilibrium choice to time pressure in the next step, V_2 is not expected to be affected here.

In the following, Figure 45 and Figure 46 illustrate the development of equilibrium-choice proportions for changing time pressure conditions. Table 78 and Table 79 (Appendix F) list the p-values corresponding to the proportion test.

Influence of time limitation

The comparative values are mostly in line with the reference values, considering the shares presented in Figure 45. The p-values support this outcome. Only five cases for all experiment versions are recorded where H_0 is rejected. Four of them belong to data from the lowest level of complexity. Despite that, the datasets show very little support for the suggested alternative hypothesis, mainly supporting the null hypothesis: time limitation does not affect equilibrium choice.

FIGURE 45 – HYPOTHESIS X: INFLUENCE OF TIME LIMITATION ON PROPORTIONS OF EQUILIBRIUM CHOICE



Influence of complexity

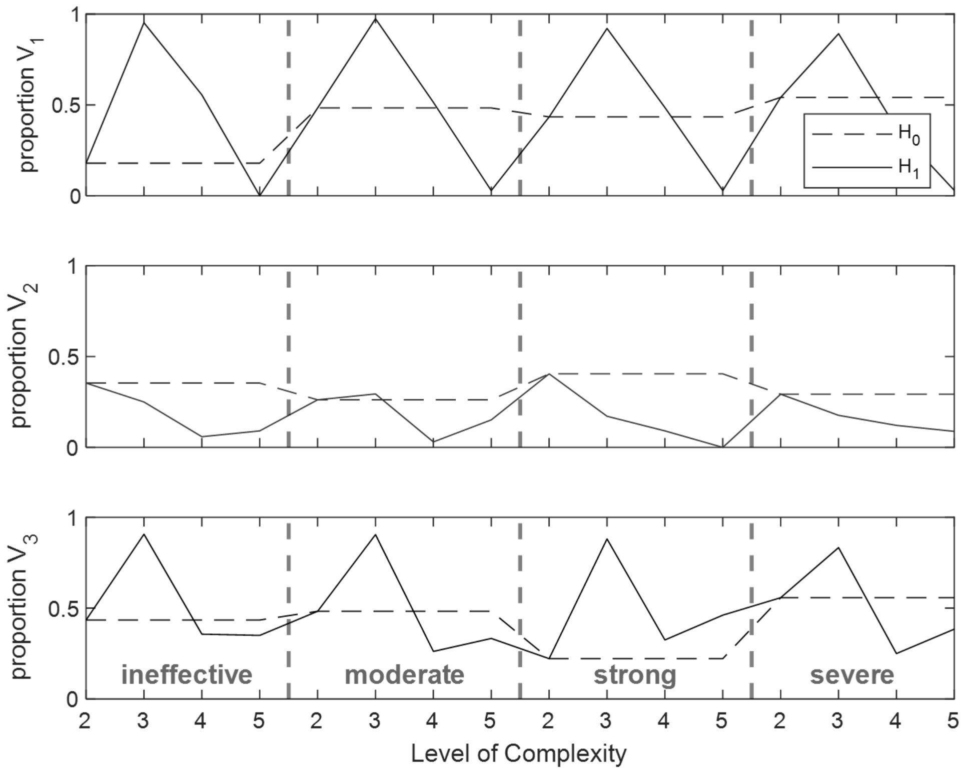
Concerning p-values, the results are changing throughout all experiment versions: in 24 out of 36 cases H_0 is rejected. However, the p-values show no stringent tendency in their change with growing complexity – neither across tasks nor versions. This issue hints to the influence of the experiment versions' design in combination with the underlying game type's payoff structure that superimposes the effect of complexity.

The sensitivity of equilibrium choice to time pressure is observed for complexity issues, even though its influence is ambivalent. For smaller complexity levels, the idea of increasing proportions of equilibrium choices gains support whereas the highest level of complexity suggests a decreasing proportion. The influence of the underlying game type's payoff structure assumingly overlays this effect. Thus, one needs to consider it in the evaluation. However, the proportion of equilibrium choices within the population is not significantly changing in tasks of even complexity levels and varying time limits. It is just varying within tasks of different levels of complexity.

Time limitation thus has no examinable impact. In general, equilibrium choice tends to be more frequently random when time limitations are rather narrow.

Regarding social differences, the occupational situation seems to have an impact: the soldiers' frequency of equilibrium choice is in 10 out of 16 cases higher than the one of students. Still, this is significant in just four cases.

FIGURE 46 – HYPOTHESIS X: INFLUENCE OF COMPLEXITY ON PROPORTIONS OF EQUILIBRIUM CHOICE



10.2.11 Other Aspects of Time Pressure induced Behavior Adaptation

During the experiment, several aspects of behavior occur that can be related to cognitive patterns. They are presented in Table 33 together with their contribution according to the underlying evaluation concept (compare Section 5.1). The systematic and significance of those aspects are studied in the following sections.

TABLE 33 – ADDITIONAL HYPOTHESES AND CONTRIBUTION

Hypotheses	**Goal 1**	**Goal 2**	**Goal 3**	**Goal 4**	**Heuristics/ patterns**
XI Acceleration of processing				x	x
XII Early decision	x				x
XIII Strategic patterns and strategic choice			x		x
XIV Focus on own payoff information				x	x
XV Initially skim all information once		x			x

10.2.11.1 Hypothesis XI: Accelerating Processing

The evaluation of Hypothesis VI revealed a strong but unanticipated influence of the time pressure task variables. As expected, time limitation has a negative and complexity has a positive impact on the numbers of EIPs in the decision-making process. The net effect of both influences is of certain interest here, since relevant studies describe an acceleration of processing under time pressure. The objective of this hypothesis is thus to verify such behavior in the context of this study. One can measure the acceleration in several ways. Two obvious approaches take into account the relationship between time and EIPs: the number of EIPs per unit of time and its quotient, the time needed per EIP. In the first approach, acceleration would be represented by an increased number of EIPs per unit of time. In the second approach, acceleration would be represented by a decreasing time per EIP. Eq. (8), represents the latter approach:

$$d_i = \frac{t_{finish,i}}{\sum EIP_i}, \tag{8}$$

with $t_{finish,i}$ as net processing time of task i until it is finished and $\sum EIP_i$ as the sum of detected EIPs in this task. Equation (8) is further applied for evaluation. With this, one can formulate Hypothesis XI as follows:

XI. Time pressure accelerates information processing and thus has a negative impact on the time used to execute an EIP, expressed by the ratio d_i.

Given a statistical distribution of the behavior, represented by the players' data stream in a certain experimental design V_e and for a fixed complexity c. Then the time-related sensitivity is assumed to impact the mean μ of the corresponding distribution of ratio d_i subject to the time limit t_i according to the following Hypothesis:

H_1: "$\mu(V_e, t_1, c) > \mu(V_e, t_2, c)$", testing the corresponding null hypothesis

(XI.a) H_0: "$\mu(V_e, t_1, c) \leq \mu(V_e, t_2, c)$" with $t_1 > t_2$ and $c = const.$

The case of evaluating the influence of complexity c is similarly formulated, with the value of c varying:

H_1: "$\mu(V_e, t_1, c_1) > \mu(V_e, t_2, c_2)$" and

(XI.b) H_0: "$\mu(V_e, t_1, c_1) \leq \mu(V_e, t_2, c_2)$" with $c_1 < c_2$ and $\frac{t_i}{c_i^2} \approx const.$

For evaluating this hypothesis, all records are considered where mouse usage is appropriate, and no premature terminations occur (V_1: 569, V_2: 351, V_3: 646). The graphs of the ratio means are presented in Figure 47 and Figure 48. Corresponding p-values are listed in Table 80 and Table 81 (Appendix F).

Influence of time limit

The sensitivity to time limitation is less distinct. More substantial deviations only occur for V_1 and V_2 for conditions of lower complexity. Here, p-values are high enough to reject H_0. The data of V_3 never produce p-values that reach the critical value. However, it is interesting that the means peak at severe time restrictions. Those peaks imply that actions are either taken more carefully instead of an expected acceleration. Alternatively, subjects more or less resign in that case. This result truly is surprising and needs further detailed analysis to determine potential reasons. A discussion is subject to the subsequent Chapter 11.

Influence of complexity

When looking at Figure 48, V_1 and V_3 show declining means with growing complexity-levels. The mainly high p-values indicate an impact as proposed by the alternative hypothesis. The null hypothesis H_0 is rejected more frequently in V_1 and V_3 (twelve and ten out of twelve cases, respectively), compared to V_2 which only shows this once. However, the acceleration assumption is widely confirmed for increasing levels of complexity.

FIGURE 47 – HYPOTHESIS XI: INFLUENCE OF TIME LIMITATION ON TIME PER EIP

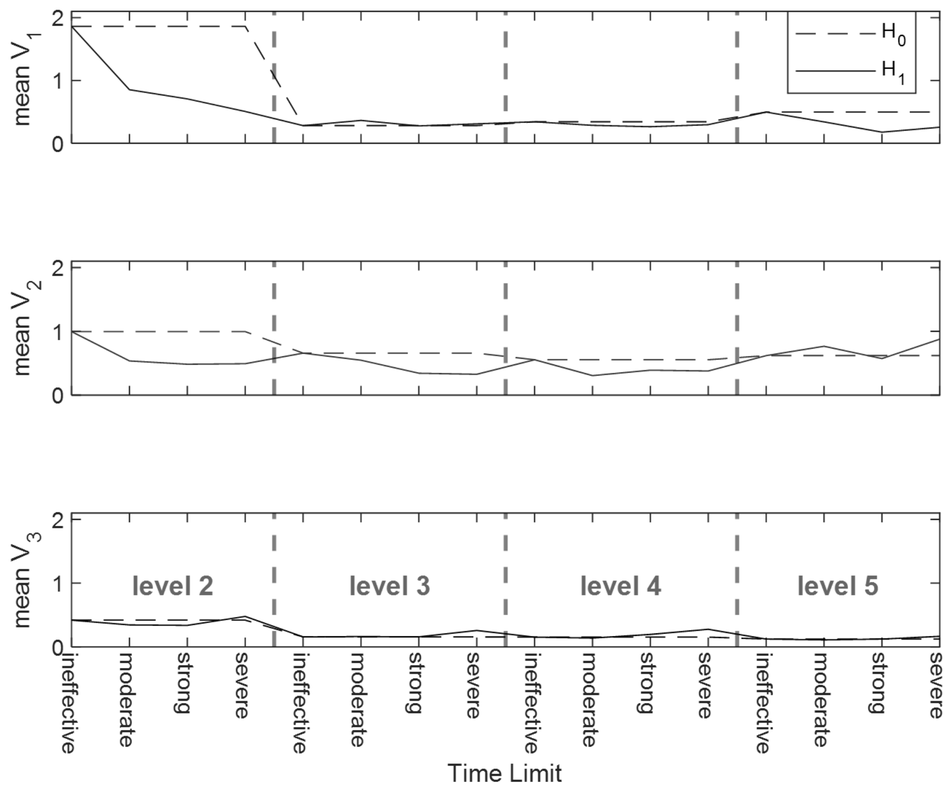
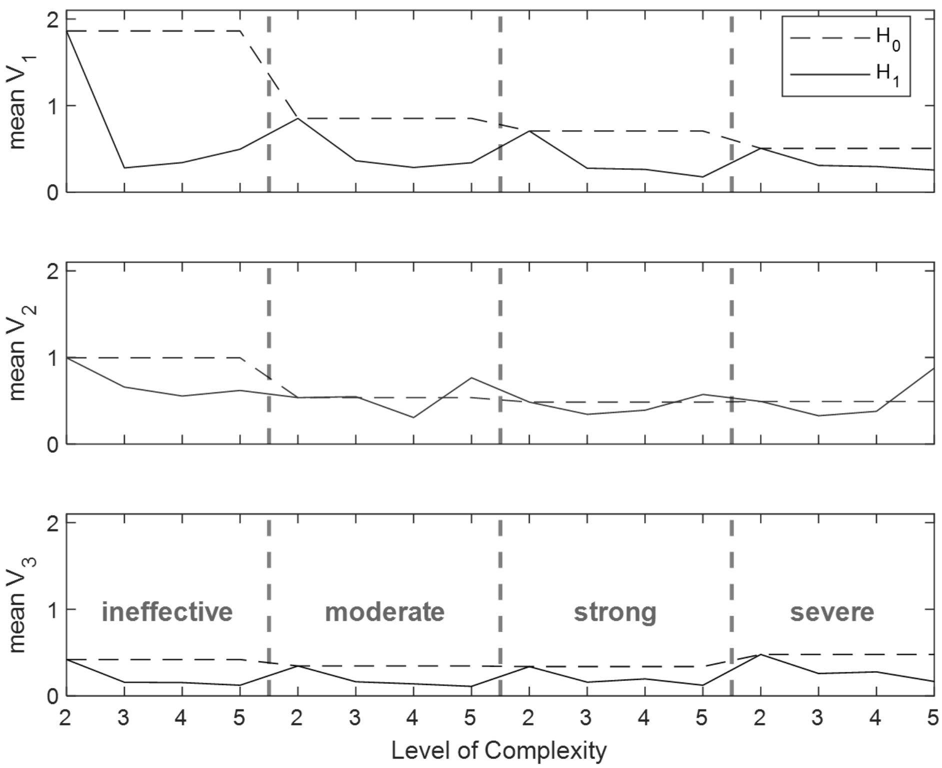


FIGURE 48 – HYPOTHESIS XI: INFLUENCE OF COMPLEXITY ON TIME PER EIP



The final result is that acceleration is observable for increasing time pressure. However, this entirely holds only for the influence of complexity, whereas time limitation shows – at best – such impact to an insufficient extent. It is thus necessary to study this effect in more detail from an individual perspective. Classification with respect to types of the observed dataset is supposed to reveal more insights into the acceleration phenomenon. No social features are observed for this hypothesis.

10.2.11.2 Hypothesis XII: Relative Decision Time

This section deals with the frequently observed feature in behavior that player make a decision in an early stage of the process and revise it with continuous time. A potential reason for this feature lies in the experimental design. The termination rule sanctions every non-decision with the lowest achievable outcome – equal to zero in relational payoffs. It is investigated here whether time pressure has any effect on this behavior. Although this strategy could be used for every task, it is possibly more important to do so in cases of severer time limitation. This way, a player can deal with the problem of sanctioning in advance and can focus on information gathering and processing for the rest of the short remaining time. In case of greater time restrictions, it is assumed that the decision-making is possibly interrupted at one time to concentrate on the time left. Herewith, one can formulate the resulting hypothesis as follows:

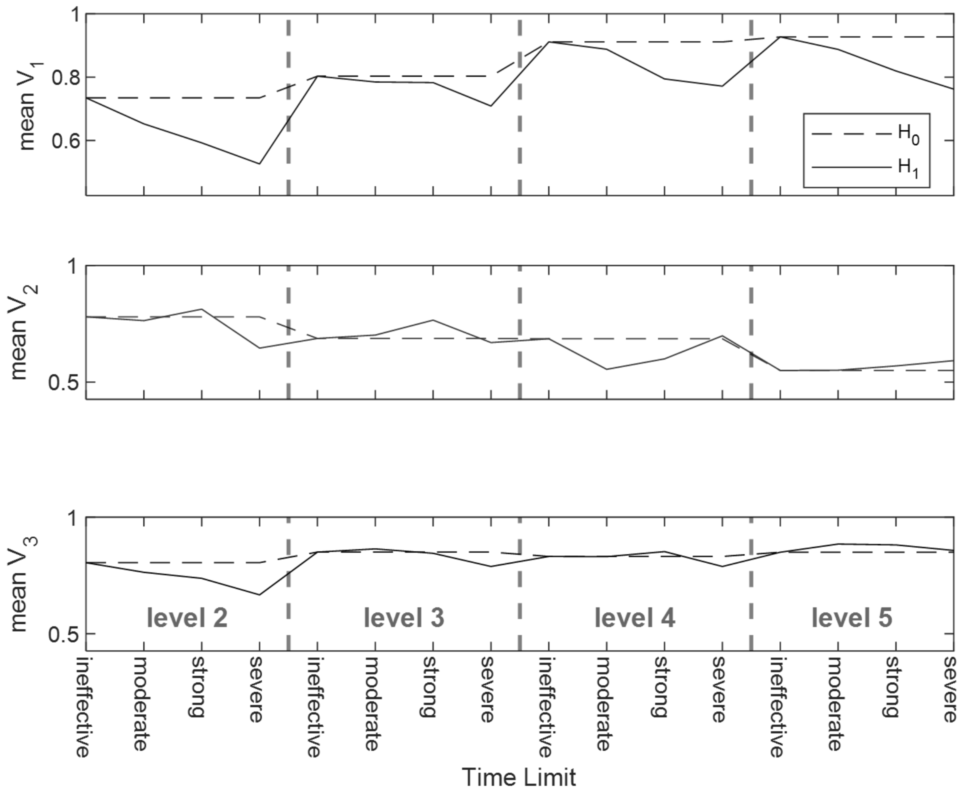
XII. Early decisions – concerning the whole processing time of a task – are likely to occur in earlier stages of the decision-making process when time pressure is rising.

For evaluation, the time duration is measured as the time difference between starting the task and first choice (i.e., click on a strategy button). The decision time is related to the finishing time of the task. Subjects can finish the task by clicking the finish button. Alternatively, *MouselabWeb* automatically ends the task when the time limit is exceeded. Hypothesis XII is evaluated for all records that provide appropriate mouse movement and a decision (V_1: 527, V_2: 254 and V_3: 529). Means of the decision-time ratio are depicted in Figure 49 and Figure 50. Corresponding p-values are listed in Table 83 and Table 84 (Appendix F).

Influence of time limit

Time limitation has an impact in most of the cases for all experiment versions. Following the means plotted in Figure 49, one can see that especially for severe time restrictions the decision-time-ratio is decreasing. For those instances V_1 is supporting the alternative hypothesis, rejecting H_0 in eight out of twelve cases. The other two versions do not provide such a high rejection rate.

FIGURE 49 – HYPOTHESIS XII: INFLUENCE OF TIME LIMITATION ON DECISION-TIME RATIOS

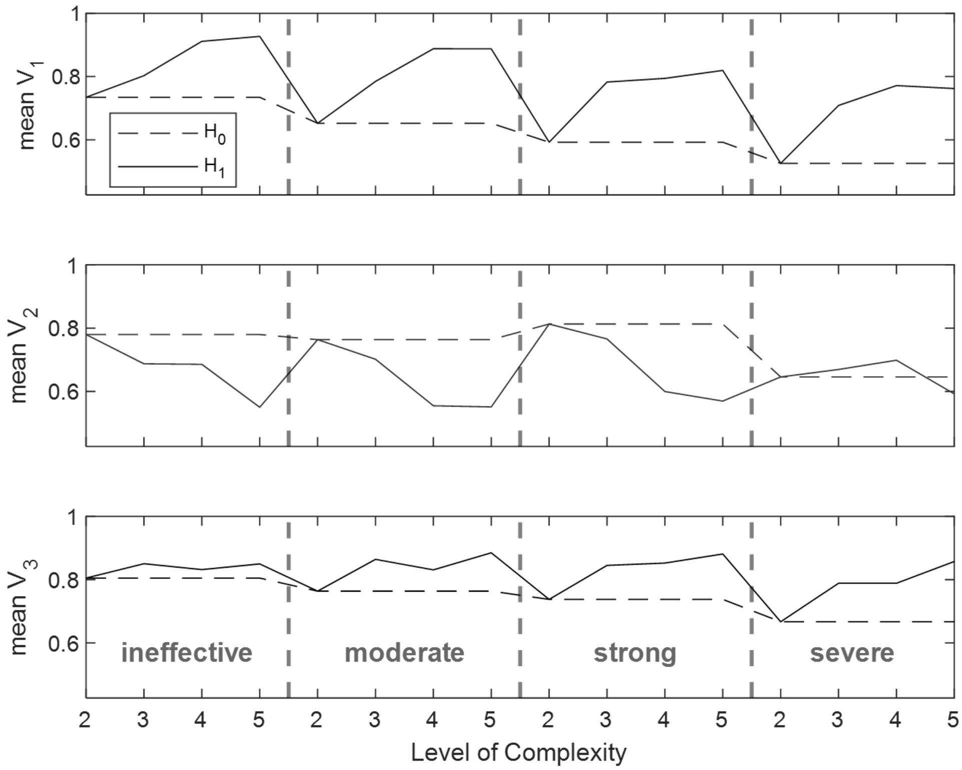


Influence of complexity

The picture fundamentally changes when examining the impact of complexity, with considerable differences among the experiment versions. As expected, V_2 supports the alternative hypothesis in five out of twelve cases. In opposition to that, the versions V_1 and V_3 show a trend of making decisions under increasing levels of complexity later. Hereby, an explanation could go either way – dependent on the experiment version in use. However, the impact of complexity, in general, is more evident than the one of time limitation.

The results of the experiment versions draw a complicated picture. However, in 38 out of 72 cases H_0 is rejected, implying sensitivity to time pressure. For V_1, the revealed impact of time limitation on decision times supports the assumptions: the stronger the time limit, the earlier the choice. This hypothesis moreover provides another example for the impact of the experiment version on the data.

FIGURE 50 – HYPOTHESIS XII: INFLUENCE OF COMPLEXITY ON DECISION-TIME RATIOS



10.2.11.3 Hypothesis XIII: Application of Strategic EIP COMPARE I

This hypothesis deals with the connection between the strategic EIP COMPARE I and strategic choice. It follows the basic idea of decision-making by heuristics. The classification of this EIP is similar to the one presented in Subsection 8.3.2. All possible decisions of the heuristics *D1*, *D2* and *Equilibrium* are characterized as strategic choices. The heuristic *L2* is excluded from consideration since this strategic heuristic works without using COMPARE I.[164] One can question whether the application of COMPARE I itself is a good indicator of strategic choice and thus linking behavior and choice. As reference serves the number of EIP COMPARE I in unsophisticated choices (i.e., decisions linked to heuristics *Altruist*, *Optimist*, *Pessimist* and *Naïve*). Thus, the hypothesis can be formulated as follows:

XIII. The use of the strategic EIP COMPARE I in strategic decisions is more frequent than in nonstrategic ones.

[164] See the production system of *L2* in Section 4.2.

To test this hypothesis, the means of the EIP COMPARE I regarding sophisticated and unsophisticated choices are compared:

H_1: "$\mu_{unsophisticated}(V_e, t_i, c_i) < \mu_{sophisticated}(V_e, t_i, c_i)$", testing the corresponding null hypothesis

(XIII.a) H_0: "$\mu_{unsophisticated}(V_e, t_i, c_i) = \mu_{sophisticated}(V_e, t_i, c_i)$",

with t_i and c_i as corresponding time and complexity variables of task i.

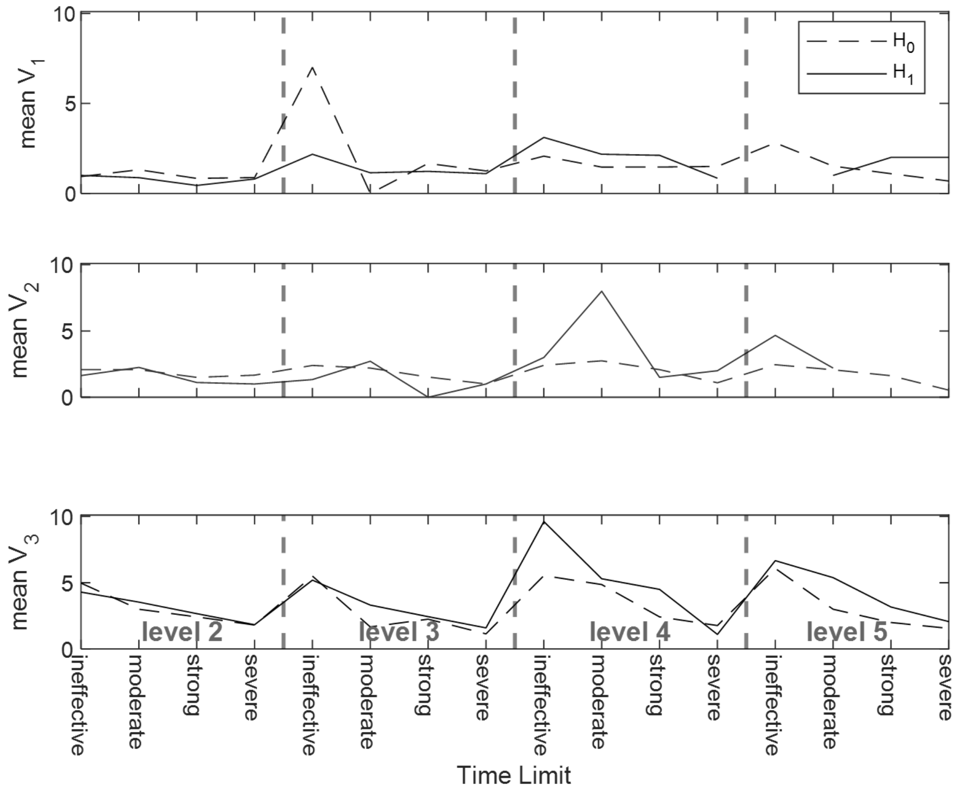
For evaluating this hypothesis, records are employed where a decision is made, mouse usage is appropriate, and no premature terminations occur. With this, the size of the dataset is 527, 254 and 529 for V_1, V_2, and V_3, respectively in numbers of fulfilled tasks. In the next step, the records are categorized regarding sophisticated and unsophisticated choices. Figure 51 presents the means with dashed lines depicting the dataset of unsophisticated choices. Table 85 (Appendix F) lists the corresponding p-values.

Analysis

Sensitivity to time limitation, in general, can be noted best at V_3. Peaks within sections occur, especially under ineffective time restrictions. The experiment versions V_1 and V_2 show less distinctive and fewer peaks. Looking at Figure 51, deviations between reference values and comparative values are seldomly observed. Four cases approach significance, with three of them supporting the expected alternative hypothesis. All of them belong to V_3. With this number of cases, one must infer that applying the EIP COMPARE I is insensitive to choice and thus no proper indicator predicting choices. However, time-bound sensitivity is indicated by the mere existence of maxima and minima and their distribution. Since the peaks' amounts are increasing with complexity increasing, applying COMPARE I is also sensitive to complexity as can be seen in Figure 51.

The age category provides a notable feature regarding social aspects. In 13 out of 16 cases, the group of 26 to 33 years olds employed COMPARE I to a more considerable extent in their decision-making process compared to the group of 17 to 25 years old. Six cases approach a significant level, mostly for moderate and strong time restrictions.

FIGURE 51 – HYPOTHESIS XIII: INFLUENCE OF TIME LIMITATION ON USE OF EIP COMPARE I



10.2.11.4 Hypothesis XIV: Time Spent Reading Own Payoff

The idea to analyze this aspect results from the findings of several reference studies which note a specific way of selecting information under growing time pressure (among others, Payne et al. 1988; Zakay 1993; Hwang 1994; Maule et al. 2000). For a strategic task environment, Devetag et al. (2016) report about players spending more time fixating their payoff than that of the opponent (pp. 190 ff.). It is thus examined here, whether one can detect such a selection concerning focussing on individual payoff or opponent's payoff with increasing time pressure. The corresponding hypothesis directly follows from the results of reference-papers:

XIV. The proportion of time a subject spends focussing on own payoff is increasing with increasing time pressure.

As null hypothesis serves the status quo – no sensitivity can be examined. The corresponding proportion of interest r_i is formed by following equation (Eq. (9)):

$$r_i = \frac{t_{own\,focus,i}}{t_{total\,focus,i}}, \tag{9}$$

with $t_{own\,focus,i}$ as time spent with the focus on own payoff in task i and $t_{total\,focus,i}$ as total time spent watching the payoff cells in task i.

All records are included that provide adequate mouse usage and no premature termination of tasks (V_1: 569, V_2: 351, V_3: 646). The ratio's means are depicted in Figure 52 and Figure 53. The flat gray dashed line at a ratio of 0.5 represents the reference level, where greater values can be interpreted as favoring the focus on own payoff over the opponent's payoff and vice versa. Corresponding p-values are listed in Table 86 and Table 87 (Appendix F).

Influence of time limitation

The plotted means of the proportion of interest in Figure 52 oscillate around the reference level section by section, indicating no stringent dependency. The expected rise of numbers with decreasing time restrictions is observed in five out of twelve sections among all experiment versions. Those deviations are yet rarely significant as the hypothesis tests reveals. The null hypothesis is rejected in 5 out of 36 cases. The p-values also show no clear tendency – neither among experiment versions nor within task variables. It is thus evident that the expected alternative hypothesis concerning the sensitivity to time limitation is generally not supported, but restricted to certain time pressure conditions and experimental designs.

Influence of complexity

In case of increasing complexity, the ratio is usually above the reference level, even though it is overpassed less frequently (see Figure 53). The hypothesis test results assert this trend with tendencies of generally smaller p-values. That supports the expected alternative hypothesis, even though only 9 out of 36 cases reach significance.

In this case, the influence of complexity mostly dominates the one of time limitation. The expected alternative hypothesis – focus on own payoff is rising under time pressure – finds support more frequently here. Increasing time limitation has not shown evident tendencies of sensitivity.

FIGURE 52 – HYPOTHESIS XIV: INFLUENCE OF TIME LIMITATION ON MEAN RATIOS OF FOCUSSING ON OWN PAYOFF

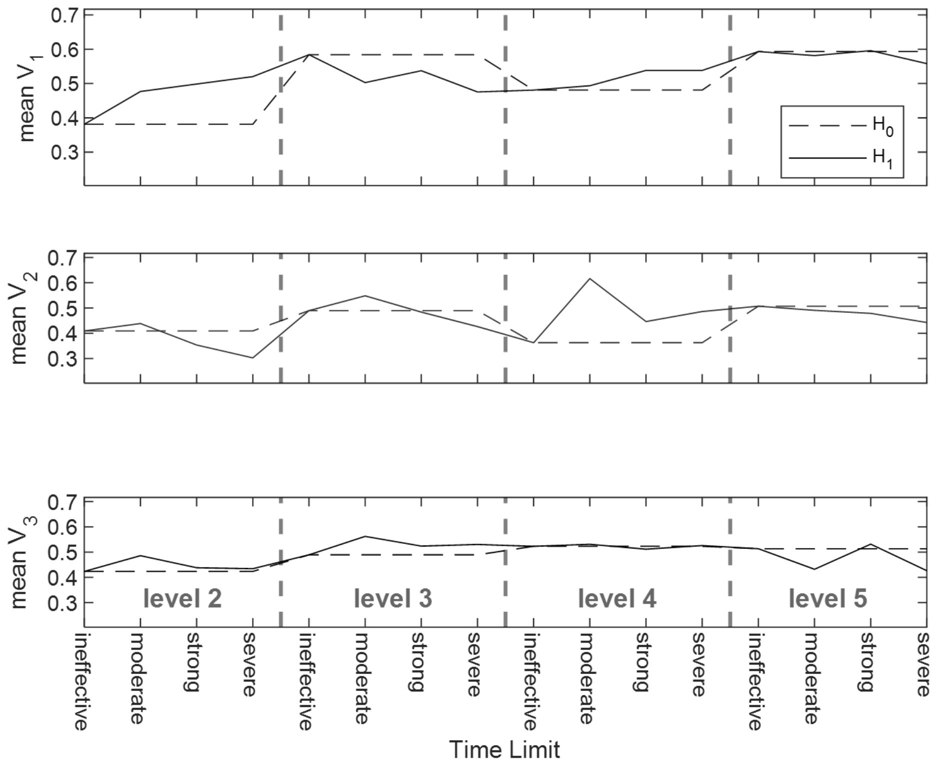
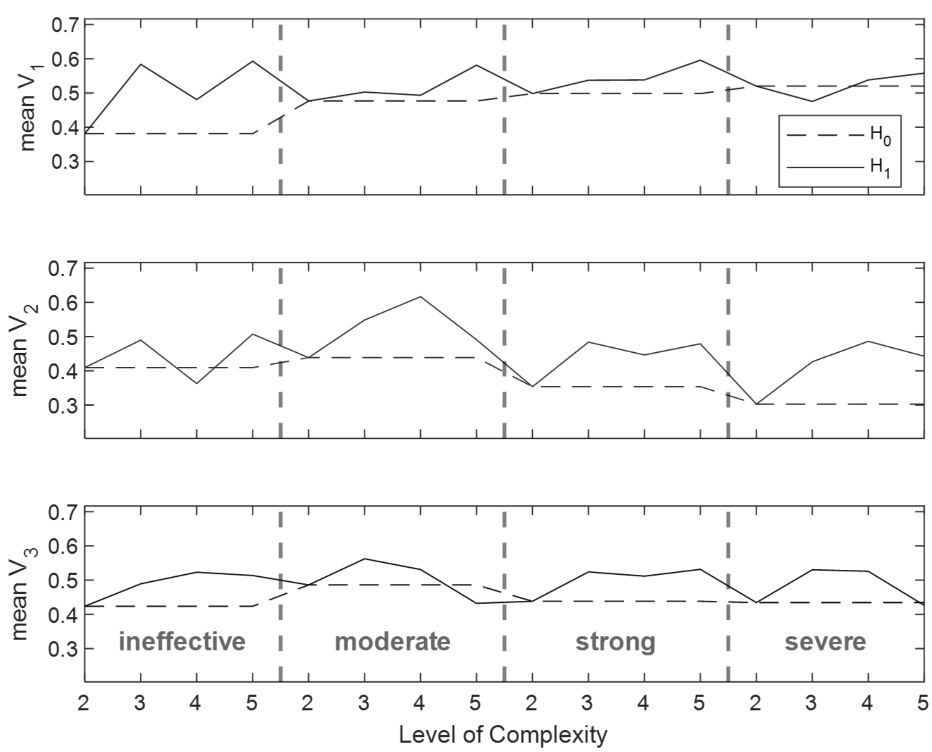


FIGURE 53 – HYPOTHESIS XIV: INFLUENCE OF COMPLEXITY ON MEAN RATIOS OF FOCUSING ON OWN PAYOFF



10.2.11.5 Hypothesis XV: Initially Skim All Information Once

This subsection deals with frequently observed behavior that is V_1-immanent.[165] Players tend to uncover all cells first before starting to process the revealed information. It is investigated at this point whether this behavior is sensitive to time pressure. The share of the population who uncovers all cells first is thus determined and compared for each task. One could expect a continuous reduction of the share under increasing time pressure since uncovering cells is a time-intensive action. Also, for higher levels of complexity, this action might lead to a confusion of the players, since this action is not directly linked to information processing, which in that case could start with 25 open cells. Herewith, the hypothesis is formulated as follows:

XV. Time pressure negatively influences the proportion of subjects who open up all cells first.

The null hypothesis depicts the case where time pressure does not significantly influence the proportion. Again, the proportion of subjects who open up all cells first determined under no time pressure serves as reference value. The examination leads to a dataset of 569 tasks from V_1. All records show adequate mouse usage and no premature termination. Proportions are depicted in Figure 54 and Figure 55. Table 88 and Table 89 (Appendix F) list the p-values of the corresponding hypothesis tests.

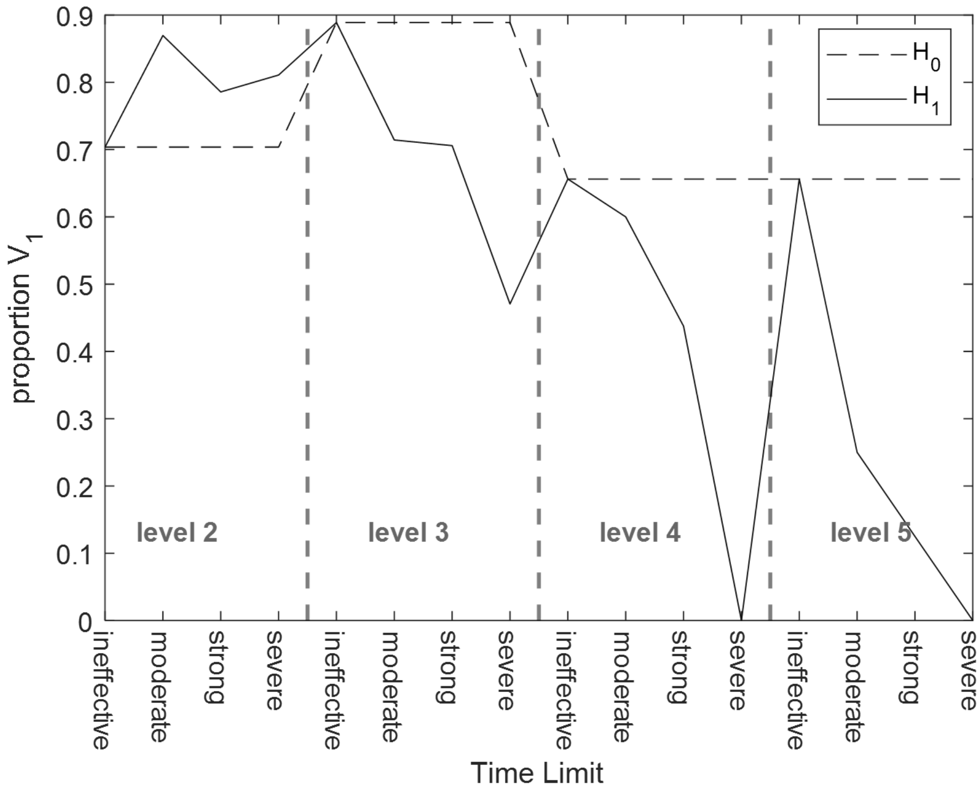
Influence of time limit

The proportion graphs in Figure 54 indicate an evident sensitivity here. For complexity level of 2, the comparative values surpass the reference value. Since the number of cells within the payoff matrix is minimal, subjects can uncover all of them in no time. Also, a priming effect needs to be considered that follows the initial task in the experiment. The process of opening cells is realized as necessary and hence automatically processed at the beginning. Hereby, the observed behavior is a lesson learned at the very beginning of the experiment. This behavior is frequently observable until the time pressure is extreme. The latter case can be already observed for the following sections of higher complexity levels where the proportion of subjects who open up all cells first massively declines with increasing time restrictions and lies henceforth below the reference val-

[165] Due to different technical designs, only V_1 starts with covered cells that can be opened and closed independently. This combination is necessary to facilitate and hence observe such behavior. Recall, that V_2 starts with open cells and at V_3 subjects cannot open cells independently.

ue. The results from hypothesis testing widely confirm that impression. In eight of twelve cases, H_0 is rejected in favor of the expected alternative hypothesis.

FIGURE 54 – HYPOTHESIS XV: INFLUENCE OF TIME LIMITATION ON PROPORTIONS OF OPENING UP CELLS FIRST



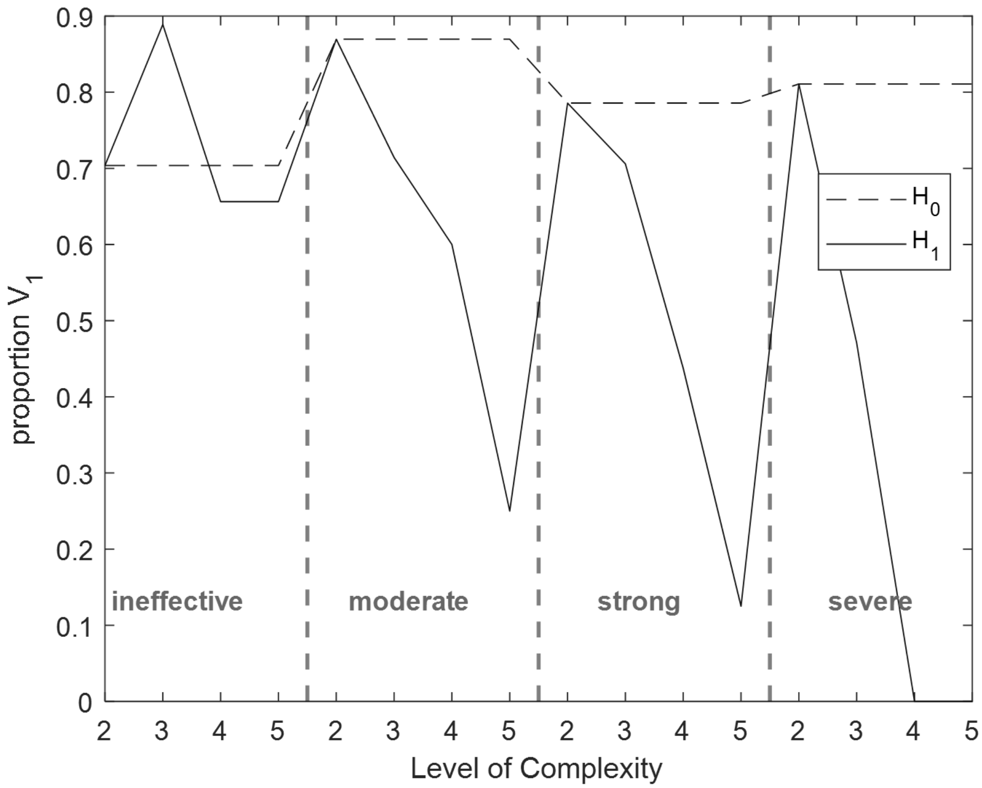
Influence of complexity

The picture does not change when examining the impact of complexity as can be seen in Figure 55. Beginning already in the section of ineffective time limitation for complexity levels of 4 and 5, the proportion values fall beneath the reference value and never exceed them subsequently. Significant support for the alternative hypothesis is observed in eight cases, rejecting H_0 nine times.

The sensitivity to time pressure is evident. Both time limitation and complexity show a clear negative influence. Differences in social aspects cannot be observed.

Results from hypothesis testing indicate a remarkable influence of time pressure on diverse aspects of decision-making. However, differences in the behavior of participants occurred. It is thus investigated whether those differences can be attributed to the existence of distinctive types of decision-making. If so, it is of further interest to define such types. The following section is dedicated to these questions.

FIGURE 55 – HYPOTHESIS XV: INFLUENCE OF COMPLEXITY ON PROPORTIONS OF OPENING UP CELLS FIRST



10.3 Classification of Types

With the results of hypothesis testing on hand, it is now possible to search for cognitive patterns and derive types of decision-making. The classification is based on the fifteen characteristics that describe the decision-making process and that already served for hypothesis testing. The characteristics are categorized according to the underlying model of the cognitive process (see Section 3.2). Categories are 'Information acquisition', 'Information processing' and 'Decision', complemented by a 'Total' category that considers the whole process. In Table 34 the matching of characteristics to categories is presented.

Preceding roman numbers relate the characteristics to the corresponding hypothesis. For the first two categories, a distinction between time pressure conditions ('tpc') and no time pressure conditions ('no tpc') is necessary to reasonably match the characteristic 'Execution accelerated' (see Hypothesis XI). This category together with 'Own payoff focussed' (see Hypothesis XIV) is regarded as relevant for both categories, ‘Information acquisition’ and ‘Information processing', and hence multiply included. Additionally, the datatype is given, with ‘b’ marking binary classification and ‘r’ rational values. Binary classification allows for determining type ratios, and thus averages can be interpreted (Coombs et al. 1975, pp. 33 ff.). This information is essential for selecting a proper clustering method.

TABLE 34 – CATEGORIZATION OF CHARACTERISTICS

	Categories						
Characteristics of behavior	**Information acquisition (no tpc)**	**Information acquisition (tpc)**	**Information processing (no tpc)**	**Information processing (tpc)**	**Decision**	**Total**	**Data type**
XII Decision made early and revised	x	x				x	r
XIV Own payoff focussed	x	x	x	x		x	r
XV All cells opened first	x	x				x	b
XI Execution accelerated		x		x		x	r
III Set of alternatives reduced			x	x		x	r
VI Total number of EIPs employed			x	x		x	r
VII Heuristic patterns used			x	x		x	b
VIII Total number of heuristic patterns used			x	x		x	r
IX *Random* pattern used			x	x		x	b
X *Equilibrium* pattern used					x	x	b
I Decision made					x	x	b
II Relative payoff generated					x	x	r
IV Relative payoff generated when having employed strategic EIPs					x	x	r
V Relative payoff generated when not having employed strategic EIPs					x	x	r
XIII Strategic EIPs used and strategic (equilibrium-) decision made					x	x	r

Remarkable differences between types of behavior within each category are expected. This requires grouping of behavioral data in the first place. The approach to derive types from experimental data is subject of the following section.

10.3.1 Classification Proceeding

For classification, Costa-Gomes et al. (2001) compared subjects' behavior with the decision-making processes of heuristics. Their supervised-learning approach[166] is well in line with Johnson and Payne (1985), Payne et al. (1988) and Bettman et al. (1990). A distinct disadvantage is that all types necessarily belong to the basic set of heuristics. Types that are not part of this set remain unidentified. Also, Costa-Gomes et al. (2001) examined subject by subject. Polonio et al. (2014) and Devetag et al. (2016) follow a different approach in their eye-tracking studies. The authors performed a cluster analysis to group their subjects according to similar eye movement patterns. This way types can be identified independently of a specific set of heuristics. Also, cluster analysis classifies all subjects together. The method is thus adopted in the following.

Cluster analysis describes a set of methods to group large datasets of objects according to some dissimilarity definitions.[167] It is sometimes referred to as ‘unsupervised classification’, since it derives classification only from data, without pre-labeled reference objects (Tan et al. 2013, p. 490 f). The goal is to find groups that are useful, meaningful or both (Tan et al. 2013, p. 487). Cluster analysis can provide a prototype for each group that represents its members and their characteristics (Tan et al. 2013, p. 488). Results of clustering usually depend on the selected clustering algorithm, and the similarity/dissimilarity measurement applied (Hastie et al. 2013, p. 459). One can categorize clustering techniques in partitional clustering (all data is grouped in disjunctive clusters), hierarchical clustering (allows nested clusters organized in a tree) and density-based clustering (clusters are built around areas of high data density in Euclidean space). The limits of the techniques' application are frequently discussed in the literature (e.g., Hastie et al. 2013; Knaak 2013; Tan et al. 2013). With this information on hand, one can select a proper algorithm according to the application's requirements.

[166] In this classification approach, data objects are usually classified according to an underlying model of known, class-representing data objects (Han et al. 2011, p. 24). Comparing characteristics of unknown data objects with those of the representatives enables classification. Besides, there are a number of other established concepts for cluster analysis and classification.

[167] Tan et al. (2013) provide a comprehensive introduction to basic concepts of cluster analysis and common algorithms in the context of data mining.

Density-based methods are considered to be difficult to be applied to high-dimensional datasets. The reasons are that density is difficult to define under such conditions, and the corresponding algorithm is computationally intensive (Tan et al. 2013, p. 532). Its application in this context is thus not further considered. Instead, a combination of hierarchical clustering and the 'k-means'-algorithm is selected from the established pool of cluster methods for treating the fifteen-characteristic dataset.[168]

Hierarchical clustering also is computational expensive in case of high-dimensional datasets and reveals problems when data is noisy (Tan et al. 2013, p. 526). However, when applied to a dataset sample it can help to determine the real cluster number of the dataset and proper centroids. An effective clustering approach is to use these values to initiate the 'k-means'-algorithm (Tan et al. 2013, p. 503).

The popular 'k-means' partitioning algorithm provides an exclusive assignment of data objects to clusters (Tan et al. 2013, p. 496 ff.). Its cluster definition is prototype-based, in which the prototype is interpreted as the center of a cluster. For real-valued data types, the center is the average of the cluster members' data points. The algorithm is simple and effective in finding these centroids (Tan et al. 2013, p. 510). Tan et al. (2013, pp. 499 ff.) provide a description of the underlying algorithm.

Frequently addressed drawbacks of this method are its dependency on the two initialization parameters 'cluster number' and 'cluster centroids'. One can meet the initialization problem by employing hierarchical clustering or initialization heuristics such as the 'k-means++'-algorithm proposed by Arthur and Vassilvitskii (2007) to determine parameters.[169] For improving the results, multiple iterations with different values for the initialization parameters are usually suggested (Tan et al. 2013, p. 502). The goodness of a clustering can be evaluated using the 'silhouette coefficient' (or 'overall average silhouette width') introduced by Rousseeuw (1987, p. 59) and developed by Kaufman and Rousseeuw (2008, p. 96), which is independent of the cluster number. It considers both, homogeneity within and heterogeneity between clusters, with values bound to the interval [-1,1]. Low values represent poor clustering, where average distances between cluster members are higher than average distances between non-cluster members. Useful clustering typically have a silhouette coefficient larger 0.5 (Kaufman and Rousseeuw 2008, p. 88).

[168] This approach follows a similar proposal of Chen et al. (2005).
[169] *Matlab* (in version R2014b) includes this algorithm as default setup for the k-means-algorithm.

Another significant drawback of the k-means-algorithm is that the search is limited to convex clusters. It cannot identify other clusters that are based on data density or hierarchical structures. Convex clusters are expected to appear when minimizing the distance of a given data object to the centroid of the cluster. They thus require the notion of a center and a reasonable interpretation of this center concerning the dataset (Tan et al. 2013, p. 513). This fact corresponds with the intention of classifying the present experimental dataset. Identified types are supposed to show differences in values of the fifteen characteristics, with the types' values representing centroids of their clusters. The algorithm's limitation in search is thus expected to cause no false clustering here. However, this issue needs to be addressed in the discussion and potentially leads to further research that compares the results of different clustering methods.

Following the argumentation above, the experiment dataset is clustered according to the subsequent procedure:

For all experiment versions and games, do the following steps:

1. Identify records of all players and adjust for application of the algorithm.[170]
2. Chose all integer values from 2 to 10 as array of initial cluster numbers.

Do while the silhouette coefficient is < 0.5 (and thus not representing a useful clustering), and the algorithm has not exceeded the iteration maximum (size of the array from step 2):

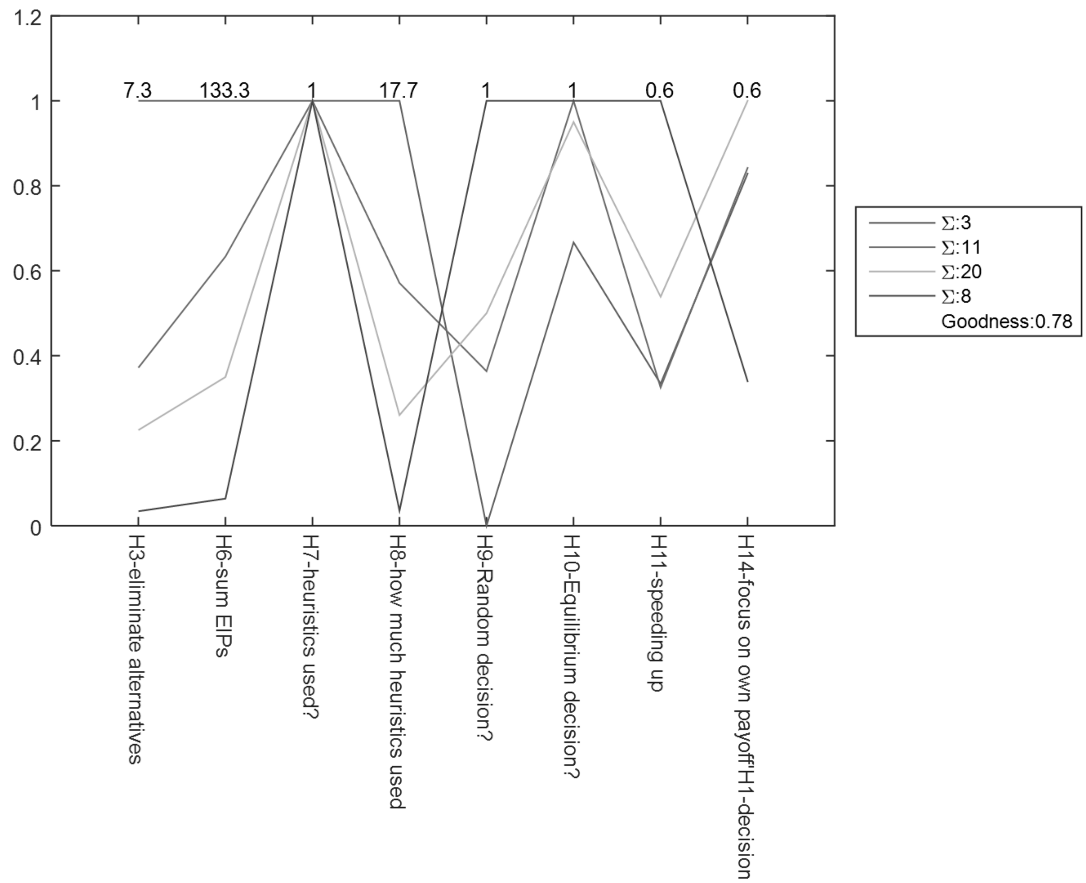
3. Find values for initial cluster centroids by k-means++-algorithm.
4. Use values from steps 2 and 3 as initial parameters of k-means-algorithm.
5. Run k-means-algorithm five times[171] and return values according to best fitting (measure: the smallest sum of the squared Euclidean distances between data points and centroid).
6. Determine silhouette coefficient of values of step 5.
7. Compare the value of step 6 to the current maximum of silhouette coefficient: if greater, store clustering results. If not greater, discard clustering results.

[170] The data scale has a significant influence on clustering based on Euclidean distances. It is thus normalized to the interval [0,1] as proposed in the literature (e.g., Han et al. (2011, p. 113) and Tan et al. (2013, p. 553)).
[171] The number of repetitions is the same as *Matlab*'s default value for the algorithm. A variation of the repetitions in the value range [1,10] has confirmed this setting: Increasing the number of repetitions beyond the value of five did not improve the fitness values.

Matlab (in version R2014b) implements this method where hierarchical clustering, the k-means- and k-means++-algorithms, as well as the silhouette coefficient, are available as predefined algorithms.

Clustering is followed by classification. Comparing the visualization of the values of the cluster prototypes, it is evident that some of them show similar tendencies in the values. Figure 56 depicts one example with the dataset partitioned into four clusters.

FIGURE 56 – EXAMPLE COMPARISON OF CLUSTER CENTROIDS



Characteristics' categories are on the horizontal axis and corresponding values on the vertical axis. Values are normalized to the interval [0,1], with the maximum value mapped to the interval's upper limit. The maximum value of each category is given in absolute terms above the lines. Each line in the figure represents the values of one cluster prototype.[172] The legend contains some additional information: the cluster size as the number of members with a leading up-

[172] A line connects the values for better illustration.

per-case sigma and the silhouette coefficient with the leading term ‘Goodness', representing the clustering goodness-of-fit.

While the lines colored in darker shades of gray differ, the lines in lighter shades have similar shapes. They seem to differ only by an absolute term or in very few characteristics. One can suppose that those clusters are connected. The results of the k-means algorithm (cluster centroids) are applied to hierarchical clustering to study their connection as proposed by Chen et al. (2005, p. 406). The resulting cluster tree is cut off at the distances between clusters that are larger 1.0. This value represents differences in the orientation of at least two characteristics. It results in a reasonable number of distinctive clusters, which the author further refers to as decision-making types.

10.3.2 Results of Type Classification

The method proposed in the prior subsection results in a specific number of types with different properties in decision-making described in the following. Types are determined for each experiment version separately and every part of the anticipated behavioral model (‘Information acquisition’, ‘Information processing’, and ‘Decision’). The presentation of the results follows this order. One can generally observe those types in each of the experiment versions. They are quite similar. However, differences exist and hence are presented along with the types' descriptions.

Characterization

In the experiment versions, five types are distinguished. While most of the types are not continuously present in all characteristics and time pressure conditions, one type is persistent. This type is further addressed as ‘Type A’. The other types are named ‘Type B’ to ‘Type E’.

The characteristics of Type A are presented in Table 35. The first two columns of Table 35 contain the parts of the behavioral model and the hypotheses that correspond with the studied characteristics. Each row refers to a single characteristic. The third column contains the characteristics under altering time limitation and complexity.

TABLE 35 – CHARACTERISTICS OF TYPE A

Part	Hypothesis	Characteristics
Information acquisition	XII: Early decision	High ratio values (0.7 – 0.9), increase with complexity and decrease with the time limitation, dominant type.
Information acquisition	XV: Open cells first	Ratio values of greater than 0.6 for small complexities, decrease with increasing complexity, dominant type.
Information acquisition/ processing	XI: Acceleration	Constantly high velocity.
Information acquisition/ processing	XIV: Focus own payoff	Slightly favoring own payoff information under increasing complexity; dominant type concerning favoring own payoff.
Information processing	III: Eliminate alternatives	Eliminations increase with complexity. They increase with time limitation up to a maximum that depends on the complexity level: the higher the level, the more moderate the time limit at which maximum is achieved.
Information processing	VI: Sum of EIPs	Number of EIPs increase with complexity and decrease with increasing time limitation, dominant type.
Information processing	VII: Use heuristics (qual.)	Constantly 100 %, dominant type.
Information processing	VIII: Use heuristics (quan.)	Numbers decrease with increasing time limit and reach their maximum at Level 3 with growing complexity, dominant type.
Information processing	IX: *Random*	Omitting information increases with time limit and complexity.
Decision	X: *Equilibrium* decision	100 % for almost all occasions, dominant type.
Decision	I: Choice	Always comes to a decision, dominant type.
Decision	II: Relative payoff	Almost constant with increasing time limitation, reaches maximum values at a complexity of level 4, dominant type.
Decision	IV: Payoff with COMPARE I	Slightly decreasing with growing complexities, reaches maximum values at a complexity of level 4, dominant type.
Decision	V: Payoff without COMPARE I	Constantly zero, since EIP COMPARE I is always used.
Decision	XIII: Strategic EIPs & choice	Generally, small numbers of strategic EIPs used, except for small complexities and moderate time limitations, dominant type.

One can describe the essentials of Type A's characteristics as follows: this type of decision maker is analyzing the given information set with a consistently high velocity (see the development of Hypothesis XI (H-XI) values in Table 35). However, probably for strategical reasons, a choice is often made long before the task processing is over and from time to time revised (H-XII). This type tends to regard all information at once at first but refuses to do so when time pressure is increasing (H-XV). Under such conditions, the own payoff information is becoming more critical (H-XIV). Eliminating alternatives seems to be an activity that is discarded fast when time limits are becoming severer (H-III). This type applies the highest number of EIPs (H-VI), even though it declines when time limits are getting severer. Among those usually are strategic EIPs (H-XIII). Both characteristics support the assumption of Type A frequently applying EIP-intensive, strategic heuristics or similar problem-solving strategies. Hereby, it selects the Equilibrium decision in most of the time (H-X). No other type achieves comparable frequencies here. Consequently, Type A develops the highest payoff, even though this value is slightly decreasing for severer time limits (H-II, H-IV). All in all, this type represents the strategic decision maker, dominating the other types in most of the observed characteristics linked to strategic reasoning.

Type B is second in frequency, even though it is not present in ‘Information acquisition’. Table 36 presents the characteristics. This type accelerates its processing speed when time pressure increases by complexity or time limitation (see the development of H-XI values in Table 36). Under such conditions Type B tends to focus more on its payoff (H-XIV) and increases to filter information (H-IX). There is evidence that this type changes problem-solving strategies under increasing complexity (H-VIII). Such behavior is similar to the conjectures of Ben Zur and Breznitz (1981, p. 43), as introduced in Section 3.3. However, a decrease in choice ratio with growing time pressure could not be observed (H-I). Type B provides little intention to eliminate alternatives (H-III) or to use strategic EIPs to develop Equilibrium decisions (H-XIII). Other, less EIP-intensive heuristics are employed instead (H-VI, H-VII). The payoffs realized this way are smaller compared to Type A, even though they increase under growing time pressure (H-II). If the strategic EIP COMPARE I is not in use, payoffs drop under extreme time pressure (H-V). In its characteristics Type B reveals close resemblance to the adaptation process of human decision-making under time pressure described in literature (compare Miller (1960, p. 697), Ben Zur and Breznitz (1981, p. 102); Zakay and Wooler (1984, p. 279), Zakay (1993, p. 60) among others).

TABLE 36 – CHARACTERISTICS OF TYPE B

Part	Hypothesis	Characteristics
Information acquisition	XII: Early decision	Not observed.
Information acquisition	XV: Open cells first	Not observed.
Information processing	XI: Acceleration	Velocity increases with increasing complexity and severer time limits.
Information processing	XIV: Focus own payoff	Ratio increases with complexity and for higher levels of complexity with time limit from low (0.2) to moderate values (0.5).
Information processing	III: Eliminate alternatives	Small values, moderately increasing with complexity, constant with the time limit.
Information processing	VI: Sum of EIPs	Small values (max = 25 EIPs), increases with complexity, decreases with the time limit
Information processing	VII: Use heuristics (qual.)	High values (80 – 100 %), increase with complexity up to Level 4, constant with the time limit.
Information processing	VIII: Use heuristics (quan.)	Slightly increasing with complexity; constant with the time limit.
Information processing	IX: *Random*	High values (80 %) from the beginning which increase with complexity and time limit up to 100 %.
Decision	X: *Equilibrium* decision	High ratios for small complexities and moderate time limits (> 90 %), values decrease sharply with growing complexity and severer time limits (0 %).
Decision	I: Choice	High values (90 – 100 %).
Decision	II: Relative payoff	Ratio increases with complexity and time limit from zero up to 0.7 of the maximum payoff achievable.
Decision	IV: Payoff with COMPARE I	Ratio is zero in most cases, with sparse exceptions (0.2 – 0.4).
Decision	V: Payoff without COMPARE I	Ratio increases with complexity from zero to 0.8, sharp decline for high complexity combined with severe time limit.
Decision	XIII: Strategic EIPs & choice	Constantly small values (0.0 – 0.2).

Type C shows a single occurrence (in ‘Decision’) with one value larger zero for each of the corresponding heuristics. Characteristics regarding behavior under time pressure conditions cannot be inferred from that. A detailed characterization is thus omitted here.

Type D is mainly present in ‘Information acquisition’. In ‘Information processing’ and ‘Decision’, this type is only present under conditions of little complexity and modest time limitation. Its characteristics are shown in Table 37. In opposition to Type A, players of this type show little intention to make early decisions under moderate time pressure. However, this attitude changes under growing complexity (see the development of H-XII values in Table 37). Similar to Type B, information acquisition and processing speed accelerate with growing time pressure (H-XI). Certain adaptation in ‘Information processing’ under time pressure cannot be stated because of Type D's limited occurrence. However, this type tends to favor its payoff information (H-XIV). Alternatives are rarely eliminated (H-III). The small sums of EIPs in use point to the application of less effortful heuristics and problem-solving strategies (H-VI – H-VIII). Omitting information seems to increase with complexity (H-IX). In ‘Decision’, Type D only occurs on two occasions, both for the smallest level of complexity under moderate and strong time limits. Characteristics can thus not be inferred.

Similar to Type C, Type E shows very few occurrences (three in ‘Information acquisition’ and two in ‘Decision’ within the game set) with values larger zero for each of the corresponding heuristics. Characteristics regarding behavior under time pressure cannot be inferred from that. A detailed characterization is thus omitted at this point.

TABLE 37 – CHARACTERISTICS OF TYPE D

Part	Hypothesis	Characteristics
Information acquisition	XII: Early decision	Start with low to moderate ratios (0.1 – 0.4), which increase with complexity (up to maximum values of 1) and decrease with time limitation.
Information acquisition	XV: Open cells first	Values > 0 only for small levels of complexity or ineffective to moderate time limits; in that case increases with the time limit and level. Values vary markedly with experiment version (V_1: 0.6 – 1.0, V_2: 0.0 – 0.1, V_3: 0.2 – 0.4).
Information acquisition /processing	XI: Acceleration	Velocity increases with increasing complexity and severer time limits.
Information processing	XIV: Focus own payoff	Values > 0 only for small levels of complexity with moderate to strong time limits, favoring own payoff information.
Information processing	III: Eliminate alternatives	Small values > 0 only for small levels of complexity or ineffective to moderate time limit.
Information processing	VI: Sum of EIPs	Values > 0 only for small complexity and / or ineffective to moderate time limit (~ 50 EIPs).
Information processing	VII: Use heuristics (qual.)	Ratios > 0 % only for small levels of complexity or ineffective to moderate time limit (in that case ratios are equal to 100 %).
Information processing	VIII: Use heuristics (quan.)	Values > 0 only for small levels of complexity or ineffective to moderate time limit (in that case absolute values are constantly 4).
Information processing	IX: *Random*	Ratios > 0 % only for small levels of complexity or ineffective to moderate time limit. In that case, ratios increase with complexity from 0 % up to 100 %.
Decision	X: *Equilibrium* decision	Constantly zero.
Decision	I: Choice	Ratios > 0 % only occur for smallest complexity. In that case, ratios are always 100 %.
Decision	II: Relative payoff	Ratios > 0 % only occur for smallest complexity. In that case, values reach a maximum of 0.6 at strong time limits.
Decision	IV: Payoff with COMPARE I	Values > 0 only occur for smallest complexity. In that case, values reach a maximum of 0.2 at strong time limits.
Decision	V: Payoff without COMPARE I	Values > 0 only occur for smallest complexity. In that case, values reach a maximum of 0.4 at strong time limits.
Decision	XIII: Strategic EIPs & choice	Constantly zero.

Changes of Type Frequency

After characterizing the five types, it is of interest how the frequency of the types changes with respect to the game set. The rates of remaining at one type and switching to another can point to changes in the general problem-solving strategy with varying time pressure conditions. The following figures (Figure 57 to Figure 65) illustrate that aspect for each experiment version and process part. Each figure contains three stacked subplots.

The upper plot illustrates the frequency of the types in numbers for increasing time pressure conditions – similar to the order of tasks in the experiment. The size (in absolute numbers) of the most frequent type group per game is added in the upper region of this plot.

The middle graph shows the ratio of remaining players to all players. That is the share of players who are assigned to the same type in game i and the following game $i + 1$ compared to the type's group size in game i. The size of the largest group of remaining players is additionally presented in absolute values in the upper part of this plot. Note that there is no valid value of the remaining players in the last game because of this ratios' definition. A value of zero is thus stated.

The lower plot contains information about the ratio of players who switch types. Analogous to the ratio of remaining players, the switch ratio is given by the share of players assigned to different types in game i and the following game $i + 1$ compared to the type's group size in game i. Some data points are depicted as diamonds. They represent high switching rates of the corresponding type. Those peaks are marked when they meet the following condition: ratios that are at least one standard deviation above the arithmetical mean of the type's switching rates of all game tasks. They are essential for analyzing the volatility of type memberships. Again, the size of the largest switching group in absolute numbers is added in the upper part of this plot. The legend in the lower region of the figure informs on depicted types. Similar to the size of the group of remaining players, there is no valid value of switching subjects in the last game because of the underlying definition. A value of zero is thus stated.

The following example taken from Figure 57 shall illustrate the diagrams' information: in the first game of the experiment – complexity level 2 and ineffective time limitation – 54 players are assigned Type A. Of those, 38 remain Type A in the next game, which equals a ratio of about 0.70. Further 10 players switch to either Type D or Type E which equals a ratio of 0.19. Some players decide to quit the experiment at this stage. This group consists of 6 players

$\left(= 54 - (38 + 10)\right)$ or a ratio of about 0.11. The number of players that switch to Type A group from game i to game $i + 1$ can be found by calculating the difference between the size of Type A in game $i + 1$ and the type's size of the remainder in game i. In the example studied here, 13 players $(= 51 - 38)$ of Type D or E switch after the first game to Type A.

FIGURE 57 – CHANGES OF TYPE FREQUENCY IN INFORMATION AQCUISITION IN V_1

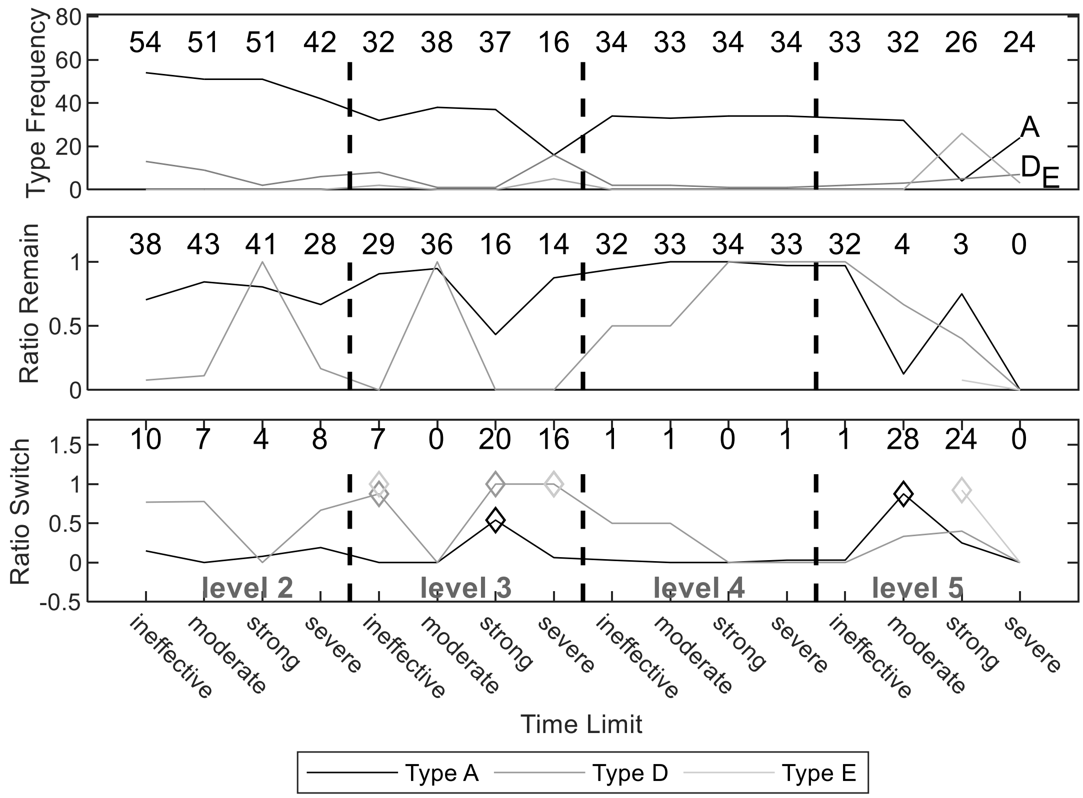


FIGURE 58 – CHANGES OF TYPE FREQUENCY IN INFORMATION PROCESSING IN V_1

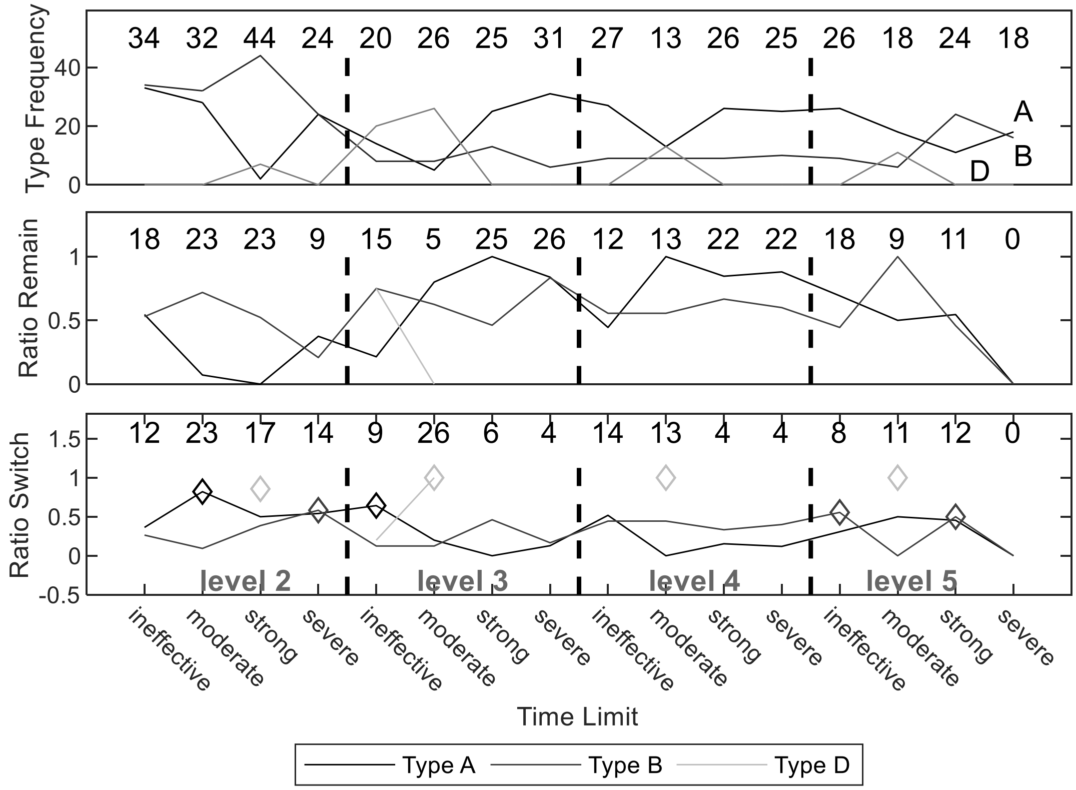


FIGURE 59 – CHANGES OF TYPE FREQUENCY IN DECISION V_1

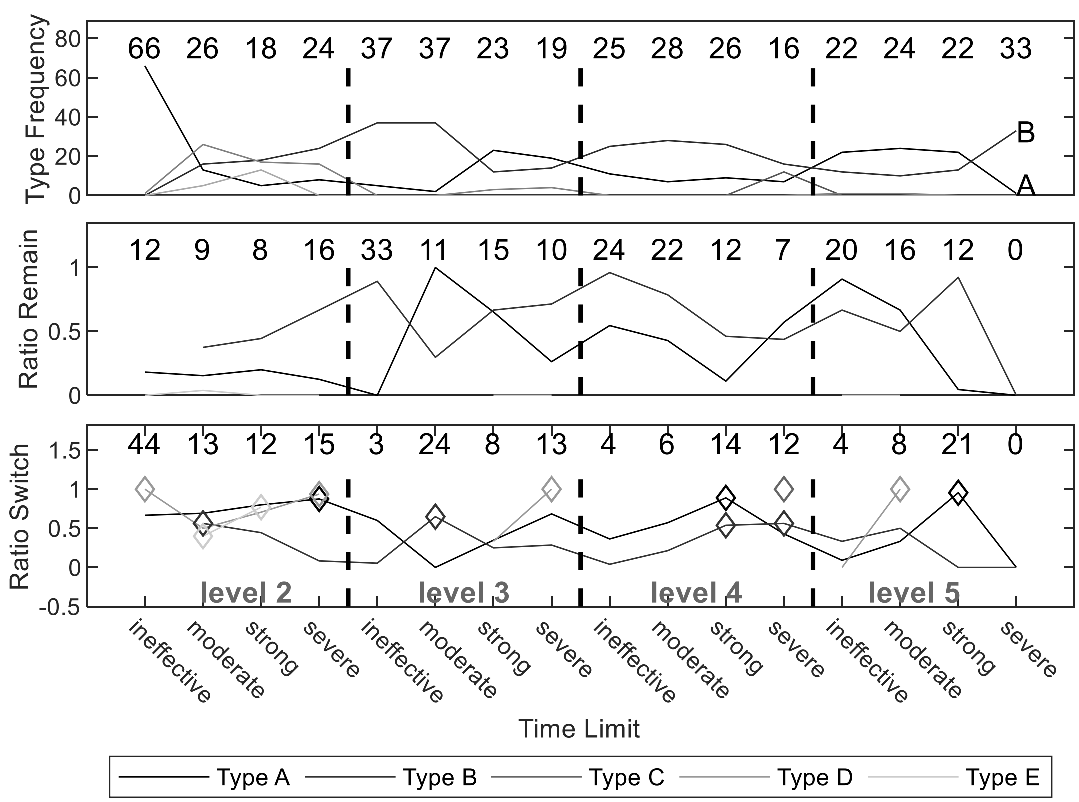


FIGURE 60 – CHANGES OF TYPE FREQUENCY IN INFORMATION AQCUISITION IN V_2

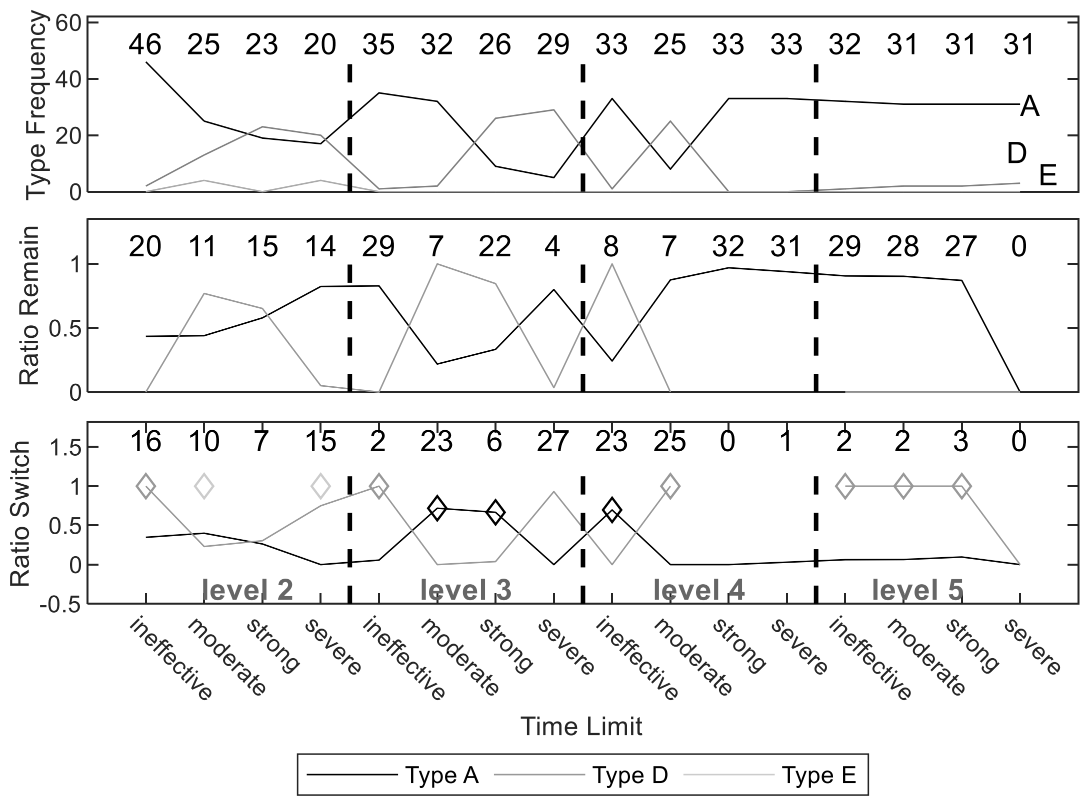


FIGURE 61 – CHANGES OF TYPE FREQUENCY IN INFORMATION PROCESSING V_2

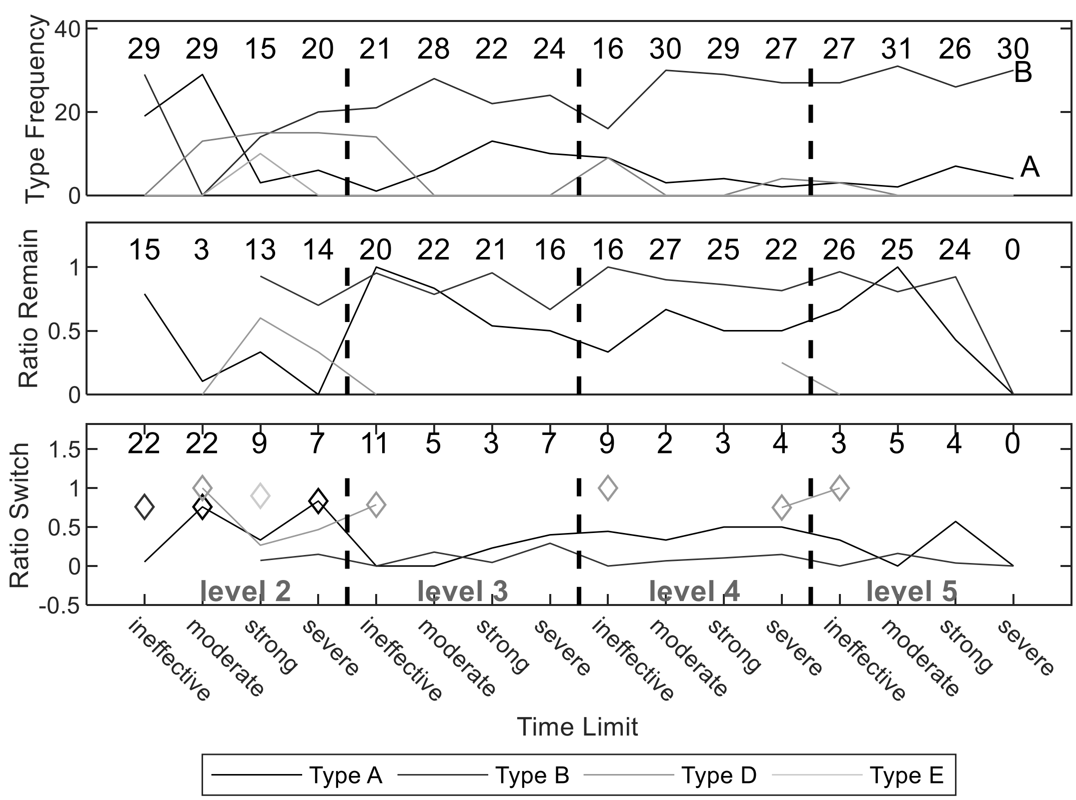


FIGURE 62 – CHANGES OF TYPE FREQUENCY IN DECISION V_2

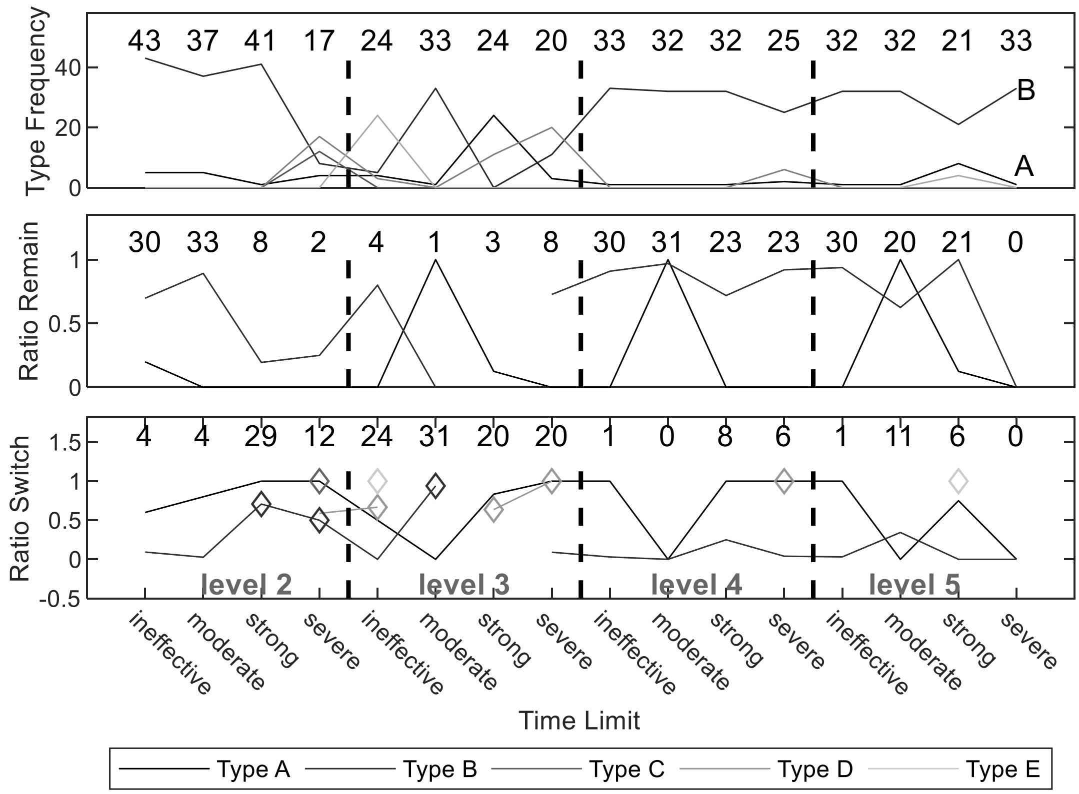


FIGURE 63 – CHANGES OF TYPE FREQUENCY IN INFORMATION ACQUISISTION V_3

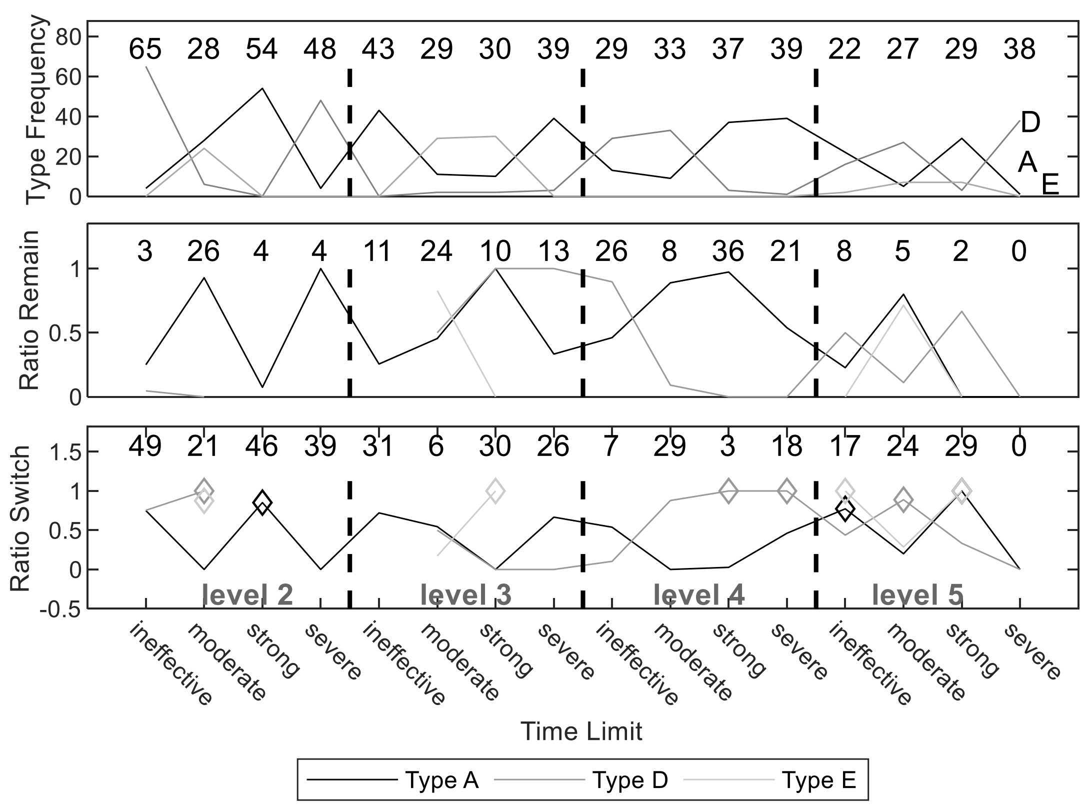


FIGURE 64 – CHANGES OF TYPE FREQUENCY IN INFORMATION PROCESSING V_3

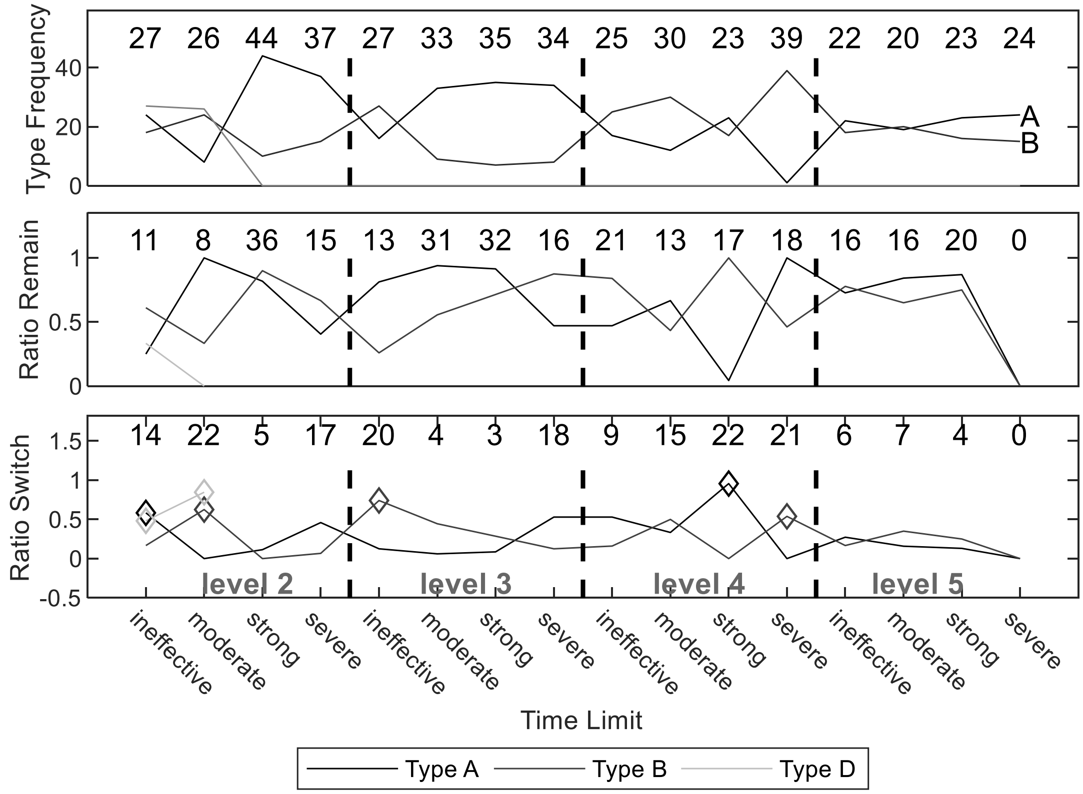


FIGURE 65 – CHANGES OF TYPE FREQUENCY IN DECISION V_3

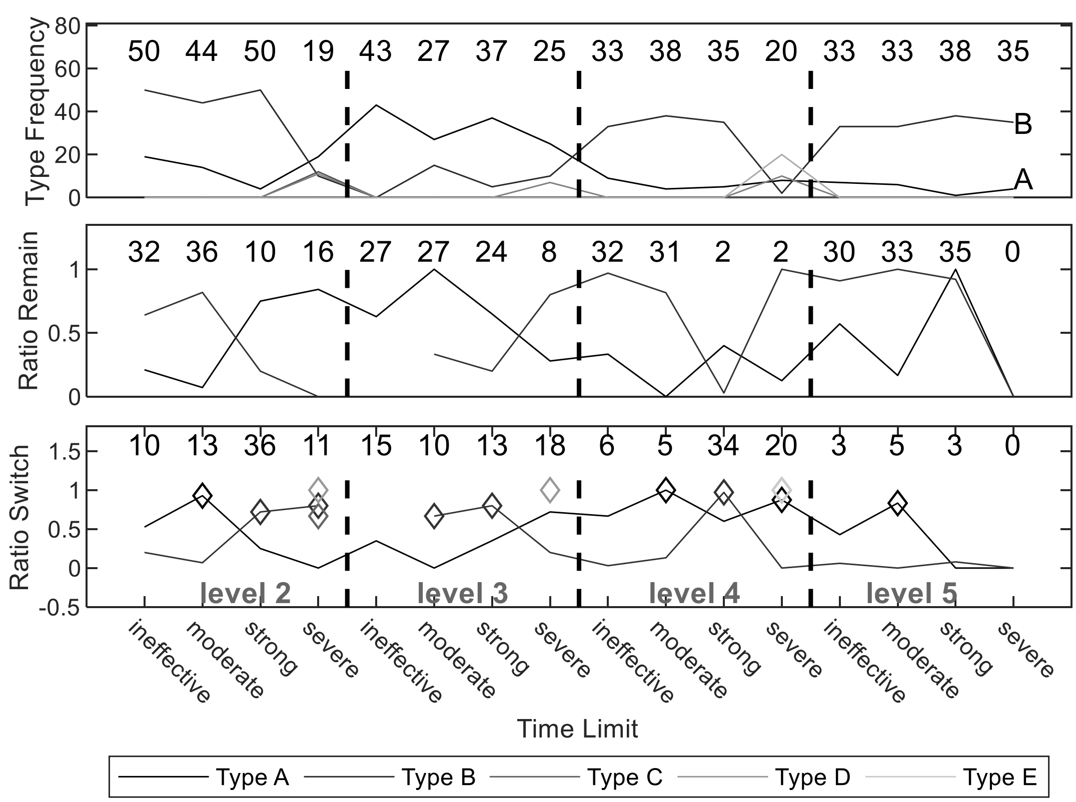


Table 38 provides a ranking of the group sizes under different time pressure conditions. Ranking information is presented for each experiment version and part of the process separately. gray cells mark cases where the first ranked type dominates all other types for all time pressure conditions.[173]

TABLE 38 – SUMMARY CHANGES OF TYPE FREQUENCY

Version	**Process part**	**Complexity** 2	3	4	5
V_1	Acquisition	A,D,E	A,D,E	A,D,E	A,D,E
V_1	Processing	B,A,D	A/D,B	A,B,D	A,B,D
V_1	Decision	B/D,A,E,C	B,A,D,C/E	B,A,C,D/E	A,B,C/D/E
V_2	Acquisition	A/D,E	A/D,E	A,D,E	A,D,E
V_2	Processing	B,A,D,E	B,A,D,E	B,A,D,E	B,A,D,E
V_2	Decision	B,A,D,C,E	A/B/D/E,C	B,D,A,E/C	B,A,E,C/D
V_3	Acquisition	D,A,E	A/E,D	A/D,E	A/D,E
V_3	Processing	A/D,B	A,B,D	B,A,D	A,B,D
V_3	Decision	B,A,D,C/E	A,B,D,C/E	B,E,A,D,C	B,A,C/D/E

Version	**Process part**	**Time limit** **ineffective**	**moderate**	**strong**	**severe**
V_1	Acquisition	A,D,E	A,D,E	A,D,E	A,D,E
V_1	Processing	A,B,D	A/D,B	A/B,D	A,B,D
V_1	Decision	A/B,C/D/E	B,A,D,E,C	A/B,D,E,C	B,A,D,C,E
V_2	Acquisition	A,D,E	A,D,E	A/D,E	A/D,E
V_2	Processing	B,D,A,E	B,A,D,E	B,A,D,E	B,A,D,E
V_2	Decision	B,A,E,D,C	B,A,C–E	B,A,D,E,C	B/D,A,C,E
V_3	Acquisition	D,A,E	D,A/E	A,E,D	A/D,E
V_3	Processing	B,A,D	B,A,D	A,B,D	A,B,D
V_3	Decision	B,A,C–E	B,A,C–E	B,A,C/D/E	A,B,E,D,C

<u>**Legend**</u>:

, separate ranks

/ share ranks

rank #1 dominates

[173] Consider the ranking for version V_1 at process part 'acquisition' and a complexity level of 2 for example. Type A is ranked first, followed by Type D and finally Type E. The cell is marked gray, showing that Type A dominates the other types for all game tasks of complexity level 2. That comprises four game tasks, each with a different time limit (ineffective, moderate, strong and severe).

It can be observed that Type A is most frequent in ‘Information acquisition’, followed by Type D and at some distance by Type E. For the ‘Information processing’ part, the predominant type varies with the experiment version. In versions V_1 and V_3, Type A again is the most frequent, followed by Type B. In case of version V_2 the types’ positions are interchanged. For all three versions, Type D follows third in group size at some distance. In the decision part, Type B dominates, closely followed by Type A and again at a distance by Types D, E and finally C.

Remaining and Switching

Costa-Gomes et al. (2001, p. 1194) assume in their paper that types are relatively constant for all game tasks they study (i.e., no time pressure, moderately varying complexity conditions). This assumption is supported by Polonio et al. (2015, p. 95), who find that patterns of a subject’s visual information acquisition does not change within a single game task.[174] That is interesting since the game tasks studied vary concerning their strategic intensity.[175] Their finding implies that subjects do not adjust their behavior to the specific strategic level of a task.

In tasks with time pressure conditions an adaptation in behavior, as described by Ben Zur and Breznitz (1981, p. 102), is confirmed for all identified types within the present experiment. Since types represent different concepts of problem-solving, high switching rates between the types imply fundamental changes in problem-solving strategies. It is thus essential to identify conditions under which they occur to determine the influence of time pressure.

The volatility of type memberships, as switching rate of players between two game tasks, is presented in following. Figure 57 to Figure 65 provide the rates of remaining players and switching players along the game set. Those ratios are directly linked together by the frequency of types. It is thus sufficient to focus on one of the two ratios to analyze the volatility of types. In this treatise, the focus lies on the switching ratio which offers direct access to identify changes in prob-

[174] They focused on one shot two by two games without time limitation, where either, neither or only one of the players had a dominant strategy (= four types). ‘One shot’-games last a single round. After a player makes one choice, this specific task is finished and either the task changes or the matching of players changes. This way, learning effects can be largely avoided. The characteristic ‘two by two’ refers to the number of alternatives involved in a game: here, each of the two players can choose between two alternatives. This setup is similar to a complexity level of 2.

[175] Polonio et al. (2015) conclude that the ability of sophisticated decision-making varies among humans – a problem-centered strategy adoption is usually not made. In other words, the subjects tend to apply the same strategy of information processing for the entire experiment. The term ‘sophistication' is also often referred to as the depth of reasoning (e.g., in the cognitive hierarchy model introduced by Camerer et al. (2004)) or the ability of strategic thinking (e.g., in Costa-Gomes et al. (2001)). Which of the terms is in use depends on the underlying model of examining the process of cognition.

lem-solving strategies. Diamonds depict high switching rates in the corresponding plots. Results of the analysis are given for each type separately.

For ‘Information acquisition’, Type A provides peaks in switching rates for the highest level of complexity (i.e., 5 alternatives) or demanding time limits (i.e., strong and severe). Values range from 0.5 to almost 1. However, phases of constant ratios at a value of zero can be found for complexities of levels 2 and 4 (experiment version V_1) or levels 4 and 5 (V_2). For ‘Information processing’, peaks occur for low levels of complexity (i.e., 2 and 3 alternatives) or moderate time limits. For ‘Decision’, peaks occur at high levels of complexity or demanding time limits. Herewith, peaks of ‘Information acquisition’ and ‘Decision’ correlate, suggesting a close link between the information acquisition strategy and choice. Switching from the information acquisition strategy of Type A is benefitted by demanding time pressure. The information processing strategy already shifts under moderate time pressure.

Type B does not occur for ‘Information acquisition’. For ‘Information processing’, few peaks are found. They occur under various time pressure conditions. However, continuous phases of moderate to low switching ratios for higher levels of complexity are present in all experiment versions. In case of V_2, a consistently low ratio is present for all levels. Peaks in making decisions are observed only for experiment versions V_1 and V_3: for strong time limits when complexity is of level 2 to 4. Higher levels of complexity at V_2 and V_3 go along with continuous phases of low ratios (i.e., about zero). The information processing strategy of Type B shows lowest switching ratios among all types, implying a particular attraction to corresponding participants. However, decision behavior shifts when time limitation is strong.

Type D provides evenly distributed peaks over time pressure conditions in ‘Information acquisition’. Continually high ratios (i.e., about 1) occur for higher levels of complexity at V_2 and V_3. In ‘Information processing’, one can observe maximum switching values under moderate time limits for all levels of complexity. Type D is not present in decision-making. Except for experiment version V_1, subjects assigned to Type D change the ‘Information acquisition’ strategy when complexity increases. Similar to Type A, information processing behavior shifts under moderate time limits.

In summary, five decision-making types are identified, characterized and described regarding their frequency and volatility under time pressure conditions. Especially Types A, B, and D are frequently present in the analyzed data set. They show different characteristics in the decision-making process. Decision makers of Type A can be assumed to be the most sophisticated type of decision-making, followed by D and B. All frequently observed types adapt behavior under time pressure, with Type B showing most similarities to the process described by Ben Zur and Breznitz (1981, p. 102). While Type A is the most frequent type of 'Information acquisition' and 'Information processing', it is Type B for 'Decision'. Participants further adapt to time pressure by changing their fundamental problem-solving strategy. The most frequently observed switches occur from EIP-intensive (like Type A) to less intensive (Type B) strategies. This result supports the first three results of the preparation time model (Section 3.3).

This chapter was dedicated to present the results from the experiment. Fifteen behavioral aspects, related to the process described in the preparation time model, were studied with respect to the influence of varying time pressure conditions. Some general characteristics could be determined this way. The findings generally support the results of the preparation time model. However, differences between the predictions of the model and the experimental dataset occurred in individual game tasks, requiring a more nuanced consideration of decision-making behavior. Classification revealed the existence of five types which differently adapt to time pressure. Their frequencies within the group of participants are determined as well as switching and remaining ratios. The latter two showed support for the preparation time model. In the following chapter, the results from the analysis of experimental data are discussed regarding the initial research questions of this treatise.

11 Discussion of Experimental Results

The present chapter discusses selected findings of the experiments in more detail. This discussion seems especially necessary in case of results differing from assumptions of the underlying preparation time model. Also, findings from the present study are compared to those reported in the relevant literature. This way, consistency, as well as improvements in theory, can be outlined. The chapter begins with a discussion of the data quality produced in the experiment before concentrating on the results of the hypothesis tests which rely on this data. It concludes by discussing results from classifying behavior.

Laboratory-based experimentation generally provides a high level of control and thus a sufficient data quality without influences by external bias. Those advantages in control are not automatically transferable to online experiments. Their experimental design needs to consider this circumstance carefully. As described in Chapter 9, various precautions are taken to deal with uncertainties in data entry. A minimum standard regarding the properties of the datasets is defined for each hypothesis. Those actions cannot entirely prevent mal-intended participation that otherwise matches all requirements. Nevertheless, the probability of such behavior is assumed to be negligibly small, since no aspects of the experimental design offer corresponding incentives or evoke such motivation.

A serious problem with data quality is unintended behavior that is the result of misunderstanding the experimental tasks. The introductory part comprises of explanations, tips, and a short test, in which participants need to correctly answer two questions regarding their understanding of design and content.[176] During the experimental sessions, no feedback occurred that explicitly mentioned problems in understanding the tasks. However, anticipating the interactive character of the tasks as well as spontaneously developing proper problem-solving methods without previously knowing any concepts of the game theory are demanding challenges.

Whether or not subjects could have used the mouse more intensively during decision-making when familiar to those solution concepts cannot be evaluated appropriately. This evaluation demands knowledge of a subject's individual mouse movement behavior which, in turn, forms the basis for an individualized interpretation metric. That *de facto* identification of types of mouse users, as Huang et al. (2012) suggest, is not undertaken in the context of this study. However, this

[176] Recall that the principle is studying initial behavior rather than training effects. Thus pre-experimental exercise is not favored.

is not necessary to generate reasonable inferences from individual mouse use behavior, as Huang et al. (2011) prove in their study.

The application of a general metric also had implications for the tested set of hypotheses. Quantitative differences between subjects regarding certain aspects needed to be relativized (e.g., Hypotheses IV and V) or even transformed into a qualitative statement (e.g., Hypotheses I and II) to achieve generality. The limited validity is regarded as acceptable, even though individualized analysis is assumed to bear high potential regarding decision-making insights.

The following sections of this chapter focus on the results of hypothesis testing and their implications in the light of changing complexity and time limitation. In this context, the author discusses observed dissimilarities between experiment versions as well as deviations of results of hypothesis tests from assumptions. A succeeding discussion of the experimental procedure also focusses on the examined derivation from the preparation time model's implications. Finally, the classified types of behavior are discussed and set in context to the underlying model of preparation time.

Table 39 provides an overview of the results of hypothesis testing. Similar to the notation in Chapter 10, light gray colored cells represent support for the stated hypotheses. gray color symbolizes that H_0 is rejected, yet a majority of cases point to a mostly unexpected alternative hypothesis. If H_0 is mostly accepted, the cell in the table is colored in dark gray. In case of white cells, the hypothesis is not tested for this particular experiment version.

As can be seen in Table 39, the three experiment versions perform quite differently with respect to the hypotheses. Whereas V_1 shows support for expected behavior on 14 out of 28 cases, V_2 and V_3, with 5 out of 26 cases and 9 out of 23 cases respectively, show a noticeably lower rate. However, the hypothesis test results from V_3 regarding time pressure sensitivity show a quite similar quality to the one of V_1 (21 out of 23 cases). The only two differences occur for Hypothesis II and XII. In the former case, both experiment versions share the missing sensitivity of general payoff generation regarding time limitation. This outcome characterizes the first of two parts that belong to the evaluation of Hypothesis II. The second one provides Hypothesis III where only V_1 and V_2 produced results. Since p is supportive in the case of V_1, the combined result supports also Hypothesis II whereas for V_3 the evaluation is only based on the negative vote of the test mentioned above.

TABLE 39 – OVERVIEW OF HYPOTHESIS TESTS

Hypothesis concerning influence of growing time pressure:	**Time limit** V_1	**Time limit** V_2	**Time limit** V_3	**Complexity** V_1	**Complexity** V_2	**Complexity** V_3
I General Sensitivity	support for H_1	support for H_1	support for H_1	support for H_1	support for H_1	support for H_1
II Qualitative Sensitivity: negative influence on effectiveness	support for H_1	support for H_1	H_0 cannot be rejected	support for H_1	support for H_1	support for H_1
III Reduction of Alternatives: negative influence	support for H_1	H_0 cannot be rejected	no data	support for H_1	H_0 cannot be rejected	no data
IV Payoff Strategic PSM: negative influence	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected
V Payoff Nonstrategic PSM: positive influence	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected
VI Qualitative Sensitivity: positive (time limit) /negative (complexity) influence on efficiency	support for H_1	H_0 cannot be rejected	support for H_1	support for H_1	H_0 cannot be rejected	support for H_1
VII Application of Heuristics: no influence	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	no data	no data	no data
VIII Heuristics' Completeness: negative (time limit) /positive (complexity) influence	support for H_1	H_0 cannot be rejected	support for H_1	support for H_1	H_0 cannot be rejected	support for H_1
IX Application of *Random*: positive influence	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	support for H_1	no data
X *Equilibrium* Choice: negative influence	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	support for H_1	support for tendency opposite of H_1	support for H_1
XI Acceleration: positive influence	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	support for H_1	H_0 cannot be rejected	support for H_1
XII Relative Decision-Time: negative influence	support for H_1	H_0 cannot be rejected	H_0 cannot be rejected	support for tendency opposite of H_1	H_0 cannot be rejected	support for tendency opposite of H_1
XIII Application of strategic EIP COMPARE I: negative influence	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	no data	no data	no data
XIV Focus Own Payoff: positive influence	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected
XV Initially Skim all Information once: negative influence	support for H_1	no data	no data	support for H_1	no data	no data

Legend:

no data	support for H_1	support for tendency opposite of H_1	H_0 cannot be rejected

In the other case, where the results of V_1 and V_3 are different (i.e. Hypothesis XII), both versions show similar tendencies in three out of four levels of complexity, yet V_1 surpasses the significance level more frequently. The discrepancy at a level-5 complexity where decision-time ratios unexpectedly increase for V_3 remains the only question mark when comparing the two experiment versions – which in all other cases produce similar test results.

The experiment V_2 instead shows eleven cases of different test outcomes when compared to V_1. Of the 15 cases where V_1 and V_2 show similarity, only four support the respective hypothesis: Hypotheses I and II – both sharing a rather general view on the subject matter – for the influence of time limitation and complexity. The experiment version V_2 suffers from at least three issues, most likely causing the observed discrepancy in comparison to the other two versions:

- mouse usage is not encouraged sufficiently,
- a small dataset, and
- different time limits.

The first point is due to the payoff matrix design the participants face when starting a task. With all cells initially uncovered, there is no need to use the mouse for gathering information – solely viewing is enough. In good compliance with the findings of Huang et al. (2012), the mouse pointer is inactive for the majority of the time. If at all, it is used as reading or concentration aid, especially when the information is presented complexly. However, this is not expressed in the proposed interpretation framework as it offers too little reasonable and particularly deterministic information to conclude on the underlying cognitive process from observed behavior.

As mentioned earlier, the interpretation metric requires the adequate employment of the mouse. With the little encouragement to intensively use the mouse that the design of V_2 provides, the dataset for hypothesis testing is often notably reduced. The circumstance that V_2 provides the smallest raw database amplifies this effect. It is hence difficult to reach significance in the hypothesis tests.

The last drawback is the difference in time limitation. As argued in Subsection 9.3.2.1, the restrictions are generally based on pre-experiments where the set of participants was smaller compared to this experiment. It is, therefore, possible that version-specific time limits do not represent the intended time pressure levels, although the same rules were used to determine the levels for all experiment versions.

The results from V_2 thus have to be treated with caution. An evaluation has to consider the three issues mentioned above. However, for several practical reasons this is not manageable within the scope of this treatise.

In the following, the test results of the remaining experiment versions V_1 and V_3 are reviewed. At this point, it seems practical to focus on V_1 only since the results of both experiment versions show considerable similarities. Beyond this, V_1 with 29 out of 30 hypothesis tests provides the richer database. The results are assumed to be valid for V_3, too, which is thus only implicitly reviewed at this point.

11.1 Supported Hypotheses

The impact of time pressure on decision-making is the first research objective and thus constitutes the primary focus of ongoing hypothesis evaluation. For a coherent picture, it is worth mentioning that the test results support the time limit aspect of seven hypotheses and the complexity aspect of eight hypotheses. Six of those are supported for both aspects at the same time: Hypotheses I, II, III, VI, VIII, and XV (Table 39). It can thus be summarized that (increasing) time pressure tends to

- have a significant influence on the decision-making process,
- have a significant negative influence on the effectiveness of choice,
- have a significant impact on the reduction of alternatives (the tendency is negative in case of time limitation and positive in case of complexity),
- have a significant positive influence on the efficiency of choice,
- have a significant impact on the number of heuristic-like problem-solving patterns (the tendency is negative in case of time limitation and positive in case of complexity), and
- have a significant negative influence on the frequently observed behavior of initially scanning the content of all cells once before starting structured information acquisition and processing.

The hypotheses that gain support for only one of the two aspects are Hypotheses X, XI and XII. Hypothesis XII, supported regarding its time limit aspect, provides the following characteristics:

- With a constant level of complexity and a decreasing time limit, decisions tend to be made earlier concerning the time available for task processing.
- With similar time limits, decision time relative to the time available is increasing with increasing complexity.

The latter characteristic was not anticipated by the time the hypothesis was formulated and is surprising at first glance. The behavioral pattern of making an early decision before starting information gathering and processing is less frequently applied or at least shifted to a later stage of problem-solving when complexity is growing. However, its position in the decision-making process is not traded in for the also observed behavioral pattern of visualizing the content of covered cells first, as Hypothesis XV reveals.

Questions arise why this advantageous pattern cannot retain its position within the process and which alternative behavior substitutes it. The answer to the latter question is: no different particular pattern substitutes it. Routine information acquisition and processing is used instead. Subjects make decisions when they have processed (not necessarily all available) information. The early decision-making pattern based on (nearly) no information is less often used.

Regarding the first question, the answer is more difficult. Since the effort needed to perform this action is relatively small, missing efficiency seems no reasonable argument. It can be assumed that either the motivation to apply this pattern or its availability in the current toolset of problem-solving methods drops. The intelligent cognitive system does not automatically trigger motivation and availability. However, both aspects exceed the scope of the proposed model of preparation time. Finding reasonable explanations for this phenomenon should be subject of further research projects.

In two cases – Hypothesis X and XI – support for the presumed alternative hypothesis is exclusively detected for their complexity aspects. For Hypothesis X, time limitation shows a significant impact on the share of equilibrium choice only for the smallest level of complexity when it unexpectedly grows. This outcome confirms findings of Lindner and Sutter (2013, p. 544). However, for this level, chances for picking an equilibrium choice do not significantly differ from a random choice. This makes chance a reasonable explanation. This conjecture is confirmed by the fact that for stages of higher complexity this effect is no longer observable. From a process-based view, one would expect fewer equilibrium decisions with growing time pressure, since determining these requires the repeated application of the complex EIP COMPARE I. However, this number is constant for the conditions examined here, as Hypothesis XIII shows.

In case of complexity, the sensitivity is apparent. However, this sensitivity is based in no small extent on the underlying game type of the task. Even though the share of equilibrium choice significantly differs from a random choice, this value is still depending on the number of alternatives labeled as equilibrium per total number of alternatives. This circumstance produces the observable, seemingly alternating values.[177]

[177] See Figure 46, p. 215.

Hypothesis XI deals with the processing speed of the subjects.[178] One can observe that speeding up processing of information under time pressure is possible. This behavior is applied up to a certain point which in this treatise is referred to as saturation point. For the influence of time limitation, only at level-2 complexity H_0 is rejected. Here, the time spent per EIP is monotonously decreasing, starting with a value that is about six times higher than the saturation point. At stages of level-3 complexity and higher, a significant acceleration cannot be observed. The saturation point is already reached for ineffective time bounds at these levels. Hence, H_0 is not further rejected. In case of observing the influence of complexity on accelerating processing speed, the time-per-EIP values at level 2 function as reference values. They are compared to the values of other levels of complexity with similar time limits. Since the values of level-2 complexity are comparably high, significant acceleration is observable. The affectedness by time bounds is therefore evident – until the subjects exceed the saturation point.

When combining this finding with the one from Hypothesis XII, one can state that participants facing growing complexity accelerate the information gathering and processing. However, they prolong their decision time within the permissible time limit. Similar effects are described by Reutskaja et al. (2011, p. 922) in their eye-tracking-based time pressure study.

11.2 Unsupported Hypotheses

In the following section, all unsupported hypotheses are discussed. To this end, they are categorized into two groups. The first group contains hypotheses that show tendencies which favor the corresponding alternative hypothesis for most tasks, even though they approach the level of significance only in less than half of all tasks. Hypotheses, where at least one of the time pressure stages (level 2 to level 5 complexity / ineffective to severe time limit) shows full support for the corresponding alternative hypothesis, complete this group. The second group follows the same rules as the first group, but the orientation of the detected tendencies is in contrast to the assumed correlation. All unsupported hypotheses fit in either of the two groups.

Members of group one are Hypotheses IV and V (both only for part-time limitation), IX, XIII, and XIV. Looking at the time component of Hypothesis IV, one can ascertain full support of the corresponding alternative hypothesis at a complexity of level 2. This level is the only stage where

[178] Recall that processing speed is measured in time per EIP in this study.

the generated relative payoff of COMPARE I users[179] significantly deceases when time restrictions increase. Since it is the only stage where payoff matrices change with tasks, varying time limits are not necessarily the only reason for the observed behavior. Still, payoff matrices of the game types played at that stage (Costa2A – Costa3B) are not varying as much in their values as would be necessary to explain that outcome fully. It is hence a remarkable effect triggered by decreasing time restrictions.

On the other side, there might be more general support for the proposed correlation between using the EIP COMPARE I and generating payoff. However, there are more influencing aspects to the resulting payoff values for the players: payoff matrices of the game types and choices of the randomly matched opponents. Furthermore, this hypothesis suggests a qualitative link between those two values: the mere existence of a single EIP COMPARE I is enough to consider a dataset for evaluation. The actual number of EIPs is not related to the generated payoff. However, even this might not be enough to evaluate the connection entirely: in the case of high numbers of detected EIPs COMPARE I a subject may apply a strategic heuristic. In the simulation, time limitation lead to an early termination of the heuristics' problem solving procedure (i.e. production systems). This early termination has a negative impact on effectiveness concerning generated payoff.

That said, the question arises why this cannot be observed for stages of higher complexity in the online experiment. The decisive difference between simulated and observed behavior is that in the simulation, heuristics are consequently applied – an intelligent adaptation to growing time pressure is not included. When formulating Hypothesis IV, application of the complete production systems of the heuristics without changes is implicitly assumed. This assumption does not comply with the results observed. With this, one can understand the observed effect as evidence for rather flexible decision-making: the EIP COMPARE I is applied, but not necessarily a complete strategic heuristic. To this extent, the finding supports the fundamental assumption of Bettman (1979, p. 33) that heuristics are not necessarily deployed with their complete production systems. At the same time, the assumption of Costa-Gomes et al. (2001, p. 1194), regarding continuity of decision types during task processing, must be rejected under time pressure conditions.

Argumentation for Hypothesis V and XIII is straightforward – except the fact that support for alternatives are rare and spread over all tasks rather than being concentrated on a single stage. For

[179] Recall that this EIP is the cognitive pattern for a pairwise comparison of the payoff of two alternatives in a game.

Hypothesis XIII it is essential to emphasize that time pressure has no significant impact on the number of the EIP COMPARE I applied by participants who make strategic choices. This finding is another piece of evidence supporting Bettman's assumption.

Test data of Hypothesis IX show clear tendencies for supporting the proposed alternative, claiming a rising share of random choices under growing time pressure. One can thus generally associate the effect with the influence of time pressure. However, the significance level is not reached before the later stages, making its influence evident at conditions of either high complexity or severe time pressure.

The last member of group one is Hypothesis XIV. After investigating the time spent focussing one's payoff during tasks, especially ‘complexity’ showed a clear, partly significant impact. In all but one task, the ratio of favoring the own payoff information over the one of the opponent's increased with growing complexity level. By the way, the results of V_2 and V_3 show even more support for the corresponding alternative hypothesis here. However, the tendency seems to be visible, albeit not substantial. With increasing time limitation just seven out of twelve cases are displayed where the tendency favors the alternative hypothesis. In those cases, p-values are close to the reference values of H_0.

Group 2 contains Hypotheses IV and V (both only for part ‘complexity’), and VII. The first two can be examined together, since both heuristics deal with payoff generation, yet based on disjunctive datasets. As already mentioned when discussing the influence of time limitation, ‘complexity’ shows a distinct impact on the relative payoff. However, it is once again unclear whether it is more affected by the number of alternatives or the concrete game type of the task. When comparing stages of equivalent time bounds, peaks always occur at complexity level 4. From those peaks at complexity level 4, values at level 3 and 5 decline more or less sharply. The payoff test of Hypothesis II shows similar behavior for a combined dataset. Thus, it can be concluded that neither the users of the EIP COMPARE I (as assumed in Hypothesis IV) nor their complementary counterparts (as assumed in Hypothesis V) show a specific sensitivity to time pressure in their payoff generation compared to a combined population.

Hypothesis VII concludes the discussion of members of group no. 2. Data confirms an apparent, yet unexpected tendency. As the confidence interval already suggests (cf. Table 32), the real ratio of users of heuristic patterns is higher than 95 % in the examined population. Of course, the distinction between 0.95 and higher values up to 1 regarding significance requires an appropriate size of the dataset. In case of V_1, the analyzed datasets comprise at least 30 subjects. Still, this size proves to be too small to determine significance of the heuristic's assumed tendency.

11.3 Social Aspects

Focussing on the impact of social aspects on decision-making, it is interesting that examined differences within the hypotheses make up such a small number in total. Sex and education are among those characteristics which have no observable impact on decision-making under time pressure. Since the overwhelming majority of the participants have a university education, the size of the group of participants with no such education is quite small. Obtaining significance regarding the differences between the groups is thus difficult.

The only examined aspects that are influential are age, job, and experience in game theory. The experiment V_3 provides the most effects – potentially because its dataset is the largest of the three versions which makes it easier to obtain significance. Therefore, further presented details rely on findings of V_3.

Age matters in categories of applying EIPs and mouse usage. The group of 26- to 33-year olds shows tendencies of using the mouse pointer more intensively than the younger reference group of 17- to 25-year olds. To a minor extent, the younger group also shows smaller numbers of applied EIPs than the group of participants older than 33 years of age. This fact allows for at least two conclusions: subjects of the younger group seem to act more efficiently since age has not proven to impact the payoff results. Also, subjects older than 25 seem to be more skilled concerning applying patterns of problem-solving methods.

The characteristic 'job' appears to be decisive only once as soldiers made more equilibrium decisions than students. However, this does not lead to significantly larger payoffs. Process-based reasons cannot be observed. Neither the number of sophisticated EIPs in use nor the number of problem-solving patterns in general is significantly higher. The experimental dataset needs to be reviewed again in more detail to discover explanations for this. This aspect is nonetheless not further considered at this point since it exceeds the goals of this research.

Game-theoretic preposition seems to of advantage for payoff generation as is shown in Hypotheses II and IV. However, having experiences in this field does not necessarily mean subjects permanently make use of them under growing time pressure. As a result, game-theoretic knowledge proved especially advantageous under stricter time limits. However, one can only determine statistical evidence for stages of low complexity. The influence of the game types' payoff structure on the observed effect also seems relevant. A significant application of strategic PSM patterns, as could have been assumed beforehand, is not observed for this group.

11.4 Types of Behavior

Classification identified five types, with three of them continuously present during the experiment (Types A, B, and D) and thus worth to focus on in the following. Their performance differs markedly for the process of decision-making. Decision-making of Type A is EIP-intensive and frequently includes strategic EIPs, like COMPARE I, where payoffs of two alternatives are compared to each other. The subjects acquire and process information with high speed, such that acceleration under growing time pressure seems impossible. They have potentially already reached their physical limits in mouse-moving speed and mental information processing. Despite that, Type A generates the highest payoffs and identifies the equilibrium strategy more frequently as compared to the other types. It thus can be labeled as the most sophisticated type. In 'Information acquisition' and 'Information processing', this type is also the most frequent. However, under severe time pressure, switching to the less time-consuming decision-making Types B and D increases sharply. Type A's dominating frequency in the population compared to less sophisticated Types B and D is well in line with findings of Arad and Rubinstein (2012, pp. 3570 f.). This frequency is also expectable, since the experiment's instructions explicitly pointed to the strategic decision-making interest.

While not present in 'Information acquisition', Type B is the most frequent in the 'Decision' part of the cognitive process. This type makes decisions very similar to the typical system described by Ben Zur and Breznitz (1981, p. 102): first accelerating information processing speed, second filtering information and third changing problem-solving strategy. The effort observed in the decision-making process is markedly smaller compared to Type A. However, the payoff Type B generates under moderate time pressure conditions is also smaller than the one of Type A. Under increasing time pressure, this payoff gap is almost closed.

Type D is mainly present in 'Information acquisition'. Its behavior under time pressure is very similar to the one described in the literature (Ben Zur and Breznitz 1981, p. 102; Zakay 1993, p. 60). This type thus seems to substitute Type B in the 'Information processing' part, where it is not present. Two other types are observed very rarely and are difficult to describe. Even though Types C and E form individual clusters and thus individual types, their incidence is low in most cases. In their sporadic occurrence, their behavior is similar to that of participants who refuse to proceed with task processing after beginning. In this case, some mouse move data is recorded, but it is entirely different from the one produced by participants who completed their tasks. One can thus reason that the occurrence of Types C and E mark game tasks where participants quit the experiment.

Concluding Part III, one can now summarize the results. The primary objective was to identify behavioral patterns in decision-making processes when executing normal-form games under various time pressure conditions. It was investigated how time pressure as a function of time limitation and complexity of tasks influences the effectiveness and efficiency of cognitive processes.

Mouse tracking was used to trace process data. The author presented a metric that allows for interpreting behavioral data, observed by mouse tracking, as cognitive operations (EIPs). This metric was used to interpret the behavioral data of the decision-making experiment. Findings were provided along with results of hypothesis testing and classifying behavior. Fifteen hypotheses were defined, according to the process model of preparation time, for the three process parts 'Information acquisition', 'Information processing', and 'Decision'.

Each hypothesis concerned a particular aspect of the decision process. Testing the hypotheses revealed a marked effect of time pressure on the decision-making process in all fifteen aspects observed. People adapt their behavior. Some patterns in 'Information acquisition' and 'Information processing' were identified that decision-maker use frequently. Those patterns are sensitive to time pressure, too.

Participants' results regarding the fifteen hypotheses' aspects were clustered to identify distinct types of behavior that can be interpreted as problem-solving strategies. This way, five types were defined and their performance under time pressure described. Each type represented an individual class of problem-solving strategies. One of those types proved to be a sophisticated decision-maker, while another behaved similarly to the adaptation scheme described by Miller (1960, p. 697). These two types were the most frequently observed among the group of participants. However, severer time pressure conditions generally increased the willingness to switch problem-solving strategies.

12 Concluding Remarks

In the final chapter, the introductory research questions are answered using the results of simulation and experiment. A short recapitulation of the applied research approach and results will be given in advance. In the next step, the research contributions are presented along with answers to the initial questions. The chapter ends with a summary, including open questions and an outlook on further research regarding human decision-making in strategic tasks under time pressure.

12.1 Research Approach

The purpose of this study was to identify strategic decision-making patterns and examine as well as characterize the influence of time pressure. Noncooperative game theory served as the problem-related environment offering tools for interpretation and a set of established and well-defined strategic decision situations. In literature, very few studies contributed to the field of strategic choice behavior from a process perspective or consider time pressure aspects. However, none of the studies extensively discussed the goals of the current research project.

Before answering one central research question, a proper process model of cognitive decision-making that supports studying the impact of time pressure was established (Chapter 3). An appropriate framework was found with the concept of Newell and Simon (1972). In this framework, the problem-centered model of preparation time was embedded. It models the cognitive process of decision-making on the basis of elementary information processes (EIPs) and implies aspects which relate to the behavior-adaptation rationale under time pressure, described by Ben Zur and Breznitz (1981).

To investigate the cognitive process, available patterns of decision-making behavior, which all relate to heuristics, were identified in Chapter 3, following suggestions of Costa-Gomes et al. (2001). Those heuristics were transformed into production systems on the base of EIPs in a further step. This approach was recommended by Newell and Simon (1972), enabling the derivation of a minimum set of EIPs.

A simulation was conducted to study the implications of the preparation time model (Chapter 5). Therefore, a set of heuristics, represented by their production systems, was applied to strategic games in normal-form under different time pressure conditions. Modeling time pressure was limited to its operational components ‘time limitation’ and ‘complexity’, well knowing that there are several more critical aspects. However, both components were parameterized in a reasonable range to produce a rich dataset of performance information. In this context, an evaluation concept was developed that comprises aspects of effectiveness and efficiency to measure the heuristics' performance.

The simulation results presented in Chapter 6 verify a general sensitivity to time pressure. One can distinguish two time phases, where the heuristics perform utterly different. In a transition phase, execution of the production system continues. A constant phase follows where the processing of the heuristics' production systems is already finished. The length of these two phases varies from heuristic to heuristic. When reaching the time limit in a decision task, the state of the heuristics' production systems can be described by one of these phases. Determining performance under such conditions allows grouping of the examined heuristics. Nonstrategic heuristics which do not imply mutual dependencies in choice show different values in their performance compared to the corresponding group of strategic heuristics. The latter proves to be more resource-intensive regarding EIPs and hence reaches its full capabilities after a more extended period compared to nonstrategic heuristics. This circumstance generally has negative implications for effectiveness under time pressure.

From those results, one can identify successful behavioral patterns as presented in Chapter 7. Up to a certain point in time which depends on the complexity, the group of nonstrategic heuristics generally shows better performance indicators. The heuristic *L2* shows an unexpectedly high performance quality. It originally belongs to the group of strategic heuristics but combines the advantages of nonstrategic and strategic heuristics. Herewith, *L2* frequently achieves top ranks within the determined categories under various time pressure conditions and is thus generally recommendable. The results from simulation helped to improve the assumptions of the preparation time model. They served as a basis for formulating hypotheses regarding behavior in the experiment.

An interpretation concept was developed that enables studying behavioral patterns in experiments (Chapter 8). The concept applies the process tracing method *mouse tracking*. This technique is a further development of the well-established *Information Board* technique initially introduced by Payne (1976). With this concept, one can interpret mouse movement information as cognitive process operations, according to the minimum set of EIPs. The interpretation metric uses a broad variety of behavioral as well as contextual information to infer on the particular cognitive operation. It uses an argumentation similar to production systems to reason behavior.

An experiment was designed and conducted to study subjects' decision-making behavior under time pressure conditions for specific strategic tasks (Chapter 9). With this, one can identify, describe as well as classify patterns of real cognitive processes. It offered the possibility to study the implications of the preparation time concept and to evaluate the results of the simulation. Drawing from these sources, a set of hypotheses was formulated that concerns manifold aspects of expected behavior.

The experiment was conducted online, applying the open-source *mouse tracking* software *MouselabWeb*, developed by Willemsen and Johnson (2011). An adaptation of functionalities was necessary to meet the requirements of the interpretation metric concerning information quality and quantity. For these objectives, the author implemented some new HTML functionalities in the software to handle a broader range of mouse events.

In a four-week time span, 174 participants took part, with 104 finishing the online experiment that lasted about 14 minutes and consisted of 16 tasks. Three experiment versions with differences in certain design aspects were offered, matching participants randomly. Tasks were formed by applying and adjusting a set of well-established game types of noncooperative game theory.

Chapter 10 presented the results from the experiment. As a general finding, the experimental data confirmed the sensitivity to time pressure of human decision-making behavior in strategic choice situations in principle. Concerning effectiveness, this impact was slightly negative, whereas the impact on efficiency was positive. The influence of time on the activity regarding applied EIPs, and thus problem-solving patterns was validated on various occasions. Making a decision occurred less frequent under increasing time pressure. One can explain this effect by the increasing refusal of the participants to continue the task under such conditions. This effect supports the fourth implication of the preparation time model and is consistent with the behavior-adaptation rationale described by Ben Zur and Breznitz (1981, p. 102).

Different impacts on ‘orientation’ and ‘intensity’ occurred between ‘time limitation’ and ‘complexity’ for several aspects. These deviations could mostly be explained with the proposed ‘Preparation time model’. The prominent exception was an often occurring behavioral pattern that led to early decision-making before processing the whole information set given. While reasons for applying this pattern lay in the design of gaming-rules, no explanation from a process-based view can be given when this behavior is less frequently observable with growing complexity.

Unsupported hypotheses mostly showed consistent tendencies in their values, even though they lacked statistical significance. Among those was the ratio of subjects who use heuristic patterns within the observed population. With about 98 % $\pm$ 1 % it was unexpectedly high as the corresponding confidence interval proves. However, supporting evidence for Bettman (1979) was found. He doubted that people regularly employ complete heuristic patterns. Instead, Bettman favored the application of a preferably flexible as well as less complicated toolkit of problem-solving methods (Bettman 1979, p. 33).

Another example was the proportion of random choices within the population when time pressure is increasing. As expected, this share was increasing – even though significance could not be determined until the subjects met conditions of either higher levels of complexity or stronger time restrictions. Those conditions could serve as a reference value in case of quantifying the steps of the adaptation process reported by Ben Zur and Breznitz (1981, p. 102) and others.

Somewhat surprising was the general finding that no significant relationship between the use of the strategic EIP COMPARE I and the realized payoff could be determined.[180] Potential interactions between identified sources of influence that were not considered before were likely. To explain this result, further intensive investigations were required. Data clustering and classification behaviors were the preferred means to study individual differences in the decision making process.

Using classification, five decision-making types were identified, characterized and described by their frequency within the observed population and their volatility under time pressure conditions. Especially Types A, B, and D were frequently present in the analyzed data set. They showed different characteristics in the decision-making process. A decision maker of Type A was assumingly most sophisticated, followed by D and B. All frequently observed types adapted behavior under time pressure. Among those, Type B showed good compliance with the process described by Ben Zur and Breznitz (1981, p. 102). While Type A was the most frequent type of 'Information acquisition' and 'Information processing', Type B was for 'Decision'. Participants further adapted to time pressure by changing their fundamental problem-solving strategy. The most fre-

[180] To a limited extent, the hypothesis is valid for the influence of time limitation when complexity is low.

quently observed switches were from EIP-intensive (like Type A) to less intensive (Type B) strategies. This behavior supported the first three implications of the preparation time model (Section 3.3).

Aspects not explicitly discussed in this section are not regarded as neglectable. Instead, the present selection provides an introduction rather than a complete overview. Chapters 6, 10 and 11, as well as Appendix C and Appendix F, might be of aid here, as they provide many more results.

12.2 Research Contribution

The presentation of the research contribution achieved in this treatise is based on the issues raised in Section 1.2. To answer these questions, the central research findings in Chapters 3 to 10 are used.

The goals of this treatise were to identify and characterize patterns of the decision-making processes for strategic tasks and to evaluate the impact of time pressure on them. Four research questions were formulated to specify those goals. Concepts had to be developed to answer them.

Those concepts deal with understanding and describing the cognitive process, with the reasoning of its performance under time pressure conditions, and with identifying process components in real decision-making. Each of them represents a separate contribution to this area of research, which could be usefull in further studies.

The primary research question is how to describe, explain and determine the influence of time pressure on the decision-making process. With the concept of preparation time, a model of the cognitive process for two-person normal-form games is developed. This model enables studying this research question as well as explaining the cognitive process of decision-making under time pressure conditions.[181]

The second research question is how cognitive processes and behavioral patterns of such processes can be identified. According to the preparation time model, such a process is composed of a particular set of elementary information processes (EIPs). The configuration of the set depends on the tasks which need to be solved as well as the problem-solving strategy that is employed. For a set of selected heuristics, representing such normative problem-solving strategies, produc-

[181] This concept is in line with well-established approaches in behavioral game theory (such as Camerer's Cognitive Hierarchy Model (Camerer et al. 2004)).

tion systems are designed. From those, a minimum set of employed EIPs is derived. This set proved to be sufficient as it was applied by 97 – 99 % of participants during the experiment.[182]

To identify processes and patterns from behavior, a process tracing technique was to be selected and adjusted considering the problem tasks. Mouse tracking is selected which is capable of identifying a specific set of elementary manual motor acts (EMMAs) from observed behavior. On this basis, an interpretation metric is developed to convert EMMAs into EIPs and thus relate observable mouse movements to the cognitive process. This concept helps to answer the second research question.

The third research question concerns what patterns and types can be identified from the decision-making process. Behavioral patterns are interpreted as part of the cognitive process when decision-making. With the preparation time model, decision-making can be described by EIPs and observed via the interpretation metric. Per definition, each EIP represents a particular behavioral pattern. Some are composed only of EMMAs, and some require EMMAs and additional EIPs. They thus differ in levels of complexity (cf. Figure 23, p. 140). One can further assign those EIPs to strategic and nonstrategic problem-solving strategies because of behavioral similarities.

Consequently, the production systems of the particular heuristics make use of them. By composition, the EIPs form even more complex behavioral patterns here. That allows for interpreting a complete heuristic as a single pattern itself.

The heuristics' behavior is simulated to study characteristics of normative decision-making patterns and their performances under time pressure. For this purpose, a concept to measure performance on the basis of effectivity and efficiency is developed, contributing to four goals of decision-making, namely choice, reduction of alternatives, generation of payoff, and minimize effort. With this measurement concept, the performance of patterns can be evaluated and successful patterns can be identified.

Transferring those results to human decision-behavior leads to the definition of hypotheses on specific patterns. Those hypotheses are tested in an experiment, assessing the use of assumed patterns and also revealing other patterns of problem-solving strategies that are not expected. Based on that pool of data, fifteen characteristics of behavior are classified. The clustering of the data further leads to the identification of five types of behavior. Each type has a characteristic

[182] These numbers result from evaluating Hypothesis VII Use of heuristics (qualitatively), see Subsection 10.2.7.

way to acquire and process information and develop a decision. These characteristics imply employing different patterns of behavior. Those findings regarding patterns and types observed in decision-making answer the third research question.

The fourth question concerns what influence time pressure has on the decision-making process and its patterns. Starting with the model of behavior adaptation under time pressure from Miller (1960, p. 697), the preparation time model requires specific adaptions. Among those are the acceleration of information acquiring and processing, the information selection and change of strategies. Simulation of heuristics' behavior under time pressure shows remarkable differences in the ability to cope with such conditions. Again, the categorization of heuristics into strategic and nonstrategic groups proves reasonable, implying that patterns show sensitivity to time pressure. This sensitivity of behavior could be observed in the experiment, too. The fifteen characteristics vary in many cases significantly with changing time pressure conditions.

Consequently, the identified types differently cope with time pressure, too. Findings also comprise information on the change of type's group sizes and time pressure conditions that support switching to other types. The findings regarding changing behavior with changing time pressure answer the fourth research question.

Besides that, a software program called ‘Automatic game task designer’ is developed that allows full content control of the experimental tasks. It is easy to adapt and fast to apply with its *Excel* front-end, enabling rapid creation of experiments. The software program is supplemented by a partially *Excel*-based and partially *Matlab*-based analyzer, which supports transparent data analysis and quickly visualizes the results. Both tools are adjusted to games in normal-form, making an application especially interesting in this context.

12.3 Summary and Prospects

Altogether this treatise offers a comprehensive study of individual decision-making patterns under time pressure conditions in specific strategic tasks. The process-based approach proved to be reasonable to study behavior under time pressure. The findings are well in line with related studies to either strategic decision-making (such as Costa-Gomes et al. (2001), Arad and Rubinstein (2012) and Devetag et al. (2016)) or decision-making under time pressure conditions (e.g., Ben Zur and Breznitz (1981), Zakay (1993)) or both (e.g., Lindner and Sutter (2013)).

Furthermore, the results from the present study go beyond that scope and enhance insights on strategic decision-making under time pressure. The identification of a set of fifteen characteristics to describe behavioral patterns under such conditions helped to generate findings. On that basis, the subjects' behavior from experiment is classified and linked to five types, each with different properties to adapt to time pressure.

The developed concepts for modeling, simulating and observing strategic decision-making supported the analysis of behavior significantly in both the pre-experimental normative approach as well as the post-experimental descriptive one. The research questions, concerning what patterns of decision-making can be identified, how they can be observed and be described, and how they change under time pressure, could be answered successfully within this study.

The experiment provides a rich set of mouse-movement data that can be used for analysis in the field of information gathering, processing and decision-making by participants, especially when emphasizing the individual focus. The evaluation concerning this issue is far from being completed. Questions that arose in this context during data analysis remain unanswered. Those aspects are presented in the context of the following prospects.

The present study uses fifteen criteria to describe behavioral patterns. Some consider EIPs, which in turn represent patterns of the cognitive process to acquire or handle information. However, no attempt was made to compare sequences of EIPs observed in the experiment with those resulting from simulating heuristic-based behavior. This way, participants' behavior could have been even closer related to heuristics, enhancing the approach of Costa-Gomes et al. (2001). Such an approach could provide results on how often a specific participant applies a specific heuristic as problem-solving strategy or certain patterns of a heuristic. That would give further insight into the cognitive process of decision-making. However, this approach requires significant computer processing capacities for data mining to find similar patterns within this large data set. Within this project, the author was not able to pursue this approach but suggests considering it whenever possible.

Other notable implications are expected to occur from extending the set of examined time pressure components. Having shown that time limitation and complexity crucially influence the decision process, additional aspects – especially those favored by psychologists – bear potential. Of those, the most prominent are emotions, motivation, memory phenomena, and learning effects.

One can add external criteria to these internal, i.e., unobservable ones. Bodenhausen, for instance, suggests time of the day as a significant source of influence (Bodenhausen 1990, pp. 321 f.).

The influence of cluster algorithms is only considered to a minimal extent in this study. The discussion of advantages and disadvantages of methods is limited to selected expert literature. Other, more appealing algorithms indeed exist or emerge in the future. Adapting the method of clustering and type categorization thus potentially changes the resulting number of types and its characteristics. However, present findings fit with related studies. It might thus serve as starting point for further research.

A rather fundamental question that arises while developing the concept of preparation time without finding process-based explanations concerns the studies of Bettman (1979) and Kahneman (2012): how do individuals choose problem-solving methods? This question explicitly comprises aspects of consciousness of choice, the prospect of success of such a method as well as knowledge of characteristics and performance under certain conditions. Such a choice is likely to occur before starting to apply the method to a problem task. The proposed approach of studying decision-making processes by mouse tracking is thus not capable of answering this question sufficiently. In this sense, the question is out of the scope of the present study's objectives. However, an answer to it would undoubtedly have broader implications for this research goal and the interpretation of the findings.

All of the aspects named in these prospects positively contribute to the understanding of decision-making processes and patterns applied. However, for various reasons they did not fit in the project of this dissertation. The question arises as to how far the presented concept can be refined in order to be able to be applied to the issues of individual decision-making processes. In research projects that go beyond the considered factors influencing time pressure, the presented approach could possibly be transferred to a consideration of further influencing factors on the cognitive decision-making process. These include, for example, learning effects, the time of day during task fulfillment, as well as the cultural background and the education of the subjects.

Bibliography

Abed, F. (1991): Cultural influences on visual scanning patterns. In *Journal of Cross-Cultural Psychology* 22 (4), pp. 525–534. DOI: 10.1177/0022022191224006.

Amann, Erwin (1999): Evolutionäre Spieltheorie. Grundlagen und neue Ansätze. Heidelberg: Physica (Studies in contemporary economics).

Amosov, Nikolaj M. (1967): Modeling of Thinking and the Mind. New York: Spartan.

Arad, Ayala; Rubinstein, Ariel (2012): The 11–20 money request game: a level-*k* reasoning study. In *The American Economic Review* 102 (7), pp. 3561–3573.

Ariely, Dan; Zakay, Dan (2001): A timely account of the role of duration in decision making. In *Acta Psychologica* 108 (2), pp. 187–207. DOI: 10.1016/S0001-6918(01)00034-8.

Arthur, David; Vassilvitskii, Sergei (2007): *k-means++*: The advantages of careful seeding. In Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, pp. 1027–1035.

Beach, Lee Roy; Mitchell, Terence R. (1978): A contingency model for the selection of decision strategies. In *Academy of Management Review* 3 (3), pp. 439–449.

Ben Zur, Hasida; Breznitz, Shlomo J. (1981): The effect of time pressure on risky choice behavior. In *Acta Psychologica* 47 (2), pp. 89–104. DOI: 10.1016/0001-6918(81)90001-9.

Bettman, James R. (1979): An Information Processing Theory of Consumer Choice. Reading, Mass.: Addison-Wesley Pub. Co. (Advances in marketing series).

Bettman, James R.; Johnson, Eric J.; Payne, John W. (1990): A componential analysis of cognitive effort in choice. In *Organizational Behavior and Human Decision Processes* 45 (1), pp. 111–139.

Bodenhausen, Galen V. (1990): Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. In *Psychological Science* 1 (5), pp. 319–322.

Brams, Steven J. (2000): Theory of Moves. Reprinted. Cambridge: Cambridge Univ. Press.

Braun, W. John; Rousson, Valentin; Simpson, William A.; Prokop, Jennifer (2003): Parametric modeling of reaction time experiment data. In *Biometrics* 59 (3), pp. 661–669.

Brehmer, Berndt (1992): Dynamic decision making: Human control of complex systems. In *Acta Psychologica* 81 (3), pp. 211–241.

Brocas, Isabelle; Carrillo, Juan D.; Wang, Stephanie W.; Camerer, Colin F. (2014): Imperfect choice or imperfect attention? Understanding strategic thinking in private information games. In *The Review of Economic Studies* 81 (3), 944–970. DOI: 10.1093/restud/rdu001.

Bronstein, Ilja N.; Hromkovic, Juraj; Luderer, Bernd; Schwarz, Hans-Rudolf; Blath, Jochen; Schied, Alexander et al. (2012): Taschenbuch der Mathematik: Springer-Verlag.

Brown, George W. (1951): Iterative solution of games by fictitious play. In *Activity Analysis of Production and Allocation* 13 (1), pp. 374–376.

Bull, Rebecca; Johnston, Rhona S. (1997): Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. In *Journal of Experimental Child Psychology* 65 (1), pp. 1–24.

Busemeyer, Jerome R.; Townsend, James T. (1993): Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. In *Psychological Review* 100 (3), pp. 432–459.

Camerer, Colin F. (1995): Individual Decision Making. In John H. Kagel, Alvin E. Roth (Eds.): The Handbook of Experimental Economics. Princeton, N.J.: Princeton University Press, pp. 587–703.

Camerer, Colin F. (2003): Behavioral Game Theory. Experiments in Strategic Interaction. New York, NY: Russell Sage Foundation (The Roundtable series in behavioral economics). Available online at http://www.loc.gov/catdir/description/prin031/2002034642.html.

Camerer, Colin F.; Ho, Teck-Hua; Chong, Juin-Kuan (2004): A cognitive hierarchy model of games. In *The Quarterly Journal of Economics* 119 (3), pp. 861–898. DOI: 10.2307/25098704.

Caplin, Andrew; Dean, Mark; Martin, Daniel (2011): Search and satisficing. In *The American Economic Review* 101 (7), pp. 2899–2922. DOI: 10.1257/aer.101.7.2899.

Card, Stuart K.; Moran, Thomas P.; Newell, Allen (1980): Computer text-editing: An information-processing analysis of a routine cognitive skill. In *Cognitive Psychology* 12 (1), pp. 32–74.

Chase, William G. (1978): Elementary information processes. In William K. Estes (Ed.): Human Information Processing. Hillsdale, New York: L. Erlbaum Associates (Handbook of learning and cognitive processes, v. 5), pp. 19–90.

Chen, Shu-Heng; Du, Ye-Rong; Yang, Lee-Xieng (2014): Cognitive capacity and cognitive hierarchy: A study based on beauty contest experiments. In *Journal of Economic Interaction and Coordination* 9 (1), pp. 69–105. DOI: 10.1007/s11403-013-0113-1.

Chen, Tung-Shou; Tsai, Tzu-Hsin; Chen, Yi-Tzu; Lin, Chin-Chiang; Chen, Rong-Chang; Li, Shuan-Yow; Chen, Hsin-Yi (2005): A combined *k*-means and hierarchical clustering method for improving the clustering efficiency of microarray. In Proceedings of the International Symposium on Intelligent Signal Processing and Communication Systems (2005). Hong Kong, China: IEEE, pp. 405–408.

Church, Allan H. (1993): Estimating the effect of incentives on mail survey response rates: A meta-analysis. In *Public Opinion Quarterly* 57 (1), pp. 62–79.

Claypool, Mark; Le, Phong; Wased, Makoto; Brown, David (2001): Implicit interest indicators. In Proceedings of the 6th international conference on Intelligent user interfaces. ACM, pp. 33–40.

Coombs, Clyde Hamilton; Dawes, Robyn M.; Tversky, Amos; Wendt, Dirk (1975): Mathematische Psychologie. Eine Einführung. Weinheim: Beltz (Beltz-Monographien Psychologie).

Coricelli, Giorgio; Rusconi, Elena (2011): Probing the decisional brain with rTMS and tDCS. In Michael Schulte-Mecklenbeck, Anton Kühberger, Rob Ranyard (Eds.): A Handbook of

Process Tracing Methods for Decision Research. A Critical Review and User's Guide. New York: Psychology Press (Society for judgment and decision making series), pp. 205–222.

Costa-Gomes, Miguel; Crawford, Vincent P.; Broseta, Bruno (2001): Cognition and behavior in normal-form games. An experimental study. In *Econometrica* 69 (5), pp. 1193–1235.

Costa-Gomes, Miguel A.; Crawford, Vincent P. (2006): Cognition and behavior in two-person guessing games: An experimental study. In *The American Economic Review*, pp. 1737–1768.

Damarla, Saudamini Roy; Just, Marcel Adam (2013): Decoding the representation of numerical values from brain activation patterns. In *Human Brain Mapping* 34 (10), pp. 2624–2634.

Dansereau, Donald F. (1969): An Information Processing Model of Mental Multiplication. Dissertation (unpublished). Carnegie-Mellon University, Pittsburgh, Pennsylvania, USA.

Dansereau, Donald F.; Gregg, Lee W. (1966): An information processing analysis of mental multiplication. In *Psychonomic Science* 6 (2), pp. 71–72. DOI: 10.3758/BF03327962.

de Spinoza, Benedict (1883): Tractatus Politicus. London: G. Bell & Son. Available online at http://www.constitution.org/bs/poltreat.txt, checked on 2017-01-20.

Devetag, Giovanna; Di Guida, Sibilla; Polonio, Luca (2016): An eye-tracking study of feature-based choice in one-shot games. In *Experimental Economics* 19 (1), pp. 177–201. DOI: 10.1007/s10683-015-9432-5.

Dillman, Don A. (2011): Mail and Internet surveys: The tailored design method. Update with new Internet, visual, and mixed-mode guide. Hoboken, NJ, USA: John Wiley & Sons.

Domschke, Wolfgang; Drexl, Andreas (2005): Einführung in Operations Research. Berlin/Heidelberg: Springer-Verlag.

Edland, Anne; Svenson, Ola (1993): Judgement and Decision Making Under Time Pressure. Studies and Findings. In Ola Svenson, A. John Maule (Eds.): Time Pressure and Stress in Human Judgment and Decision Making. Boston, MA: Springer US, pp. 27–40.

Efron, Robert (1970): The minimum duration of a perception. In *Neuropsychologia* 8 (1), pp. 57–63. DOI: 10.1016/0028-3932(70)90025-4.

Ericsson, K. A.; Moxley, J. H. (2011): Thinking aloud protocols: Concurrent verbalizations of thinking during performance on tasks involving decision making. In Michael Schulte-Mecklenbeck, Anton Kühberger, Rob Ranyard (Eds.): A Handbook of Process Tracing Methods for Decision Research. A Critical Review and User's Guide. New York: Psychology Press (Society for judgment and decision making series), pp. 89–114.

Ericsson, K. Anders; Kintsch, Walter (1995): Long-term working memory. In *Psychological Review* 102 (2), pp. 211–245.

Ericsson, K. Anders; Simon, Herbert A. (1980): Verbal reports as data. In *Psychological Review* 87 (3), p. 215.

Fischbacher, Urs (2007): z-Tree: Zurich toolbox for ready-made economic experiments. In *Experimental Economics* 10 (2), pp. 171–178.

Fisz, M. (1989): Wahrscheinlichkeitsrechnung und mathematische Statistik. Berlin: Deutscher Verlag der Wissenschaften (Hochschulbücher für Mathematik).

Fitts, Paul M. (1954): The information capacity of the human motor system in controlling the amplitude of movement. In *Journal of Experimental Psychology* 47 (6), pp. 381–391.

Ford, J. Kevin; Schmitt, Neal; Schechtman, Susan L.; Hults, Brian M.; Doherty, Mary L. (1989): Process tracing methods: Contributions, problems, and neglected research questions. In *Organizational Behavior and Human Decision Processes* 43 (1), pp. 75–117.

Förster, Jens; Higgins, E. Tory; Bianco, Amy Taylor (2003): Speed/accuracy decisions in task performance: Built-in trade-off or separate strategic concerns? In *Organizational Behavior and Human Decision Processes* 90 (1), pp. 148–164.

Freeman, Jonathan B.; Ambady, Nalini (2010): MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. In *Behavior Research Methods* 42 (1), pp. 226–241.

Freeman, Jonathan B.; Dale, Rick; Farmer, Thomas A. (2011): Hand in motion reveals mind in motion. In *Frontiers in Psychology* 2 (1). DOI: 10.3389/fpsyg.2011.00059.

Gertzen, Heiner (1992): Component processes of phased decision strategies. In *Acta Psychologica* 80 (1-3), pp. 229–246. DOI: 10.1016/0001-6918(92)90049-J.

Gintis, Herbert (2009): Game Theory Evolving. A Problem-Centered Introduction to Modeling Strategic Interaction. 2nd ed. Princeton, NJ: Princeton Univ. Press.

Gnambs, T.; Strassnig, B. (2007): Experimentelle Online-Untersuchungen. In Martin Welker, Olaf Wenzel (Eds.): Online-Forschung 2007. Grundlagen und Fallstudien. Köln: von Halem (Neue Schriften zur Online-Forschung, 1), pp. 233–250.

Gonzalez, Cleotilde; Dana, Jason; Koshino, Hideya; Just, Marcel (2005): The framing effect and risky decisions. Examining cognitive functions with fMRI. In *Journal of Economic Psychology* 26 (1), pp. 1–20.

Göritz, Anja S. (2006): Incentives in web studies: Methodological issues and a review. In *International Journal of Internet Science* 1 (1), pp. 58–70.

Greiner, Ben (2004): The online recruitment system ORSEE 2.0 – a guide for the organization of experiments in economics. University of Colgne. Cologne (Working paper series in economics, 10(23)).

Hahn, P. Richard; Lum, Kristian; Mela, Carl (2010): Testing cognitive hierarchy theories of beauty contest games. Duke University (Working paper Google Scholar). Available online at http://faculty.chicagobooth.edu/workshops/econometrics/past/pdf/Hahn2.pdf, checked on 2018-10-11.

Han, Jiawei; Pei, Jian; Kamber, Micheline (2011): Data Mining. Concepts and Techniques. 3rd ed. New York: Elsevier (The Morgan Kaufmann Series in Data Management Systems).

Hastie, T.; Tibshirani, R.; Friedman, J. (2013): The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer.

Henrich, Joseph; Heine, Steven J.; Norenzayan, Ara (2010): The weirdest people in the world? In *The Behavioral and Brain Sciences* 33 (2–3), 6–83; discussion 83–135.

Hess, Eckhard H. (1965): Attitude and pupil size. In *Scientific American* 212 (4), pp. 46–55.

Ho, Teck-Hua; Weigelt, Keith (1996): Task complexity, equilibrium selection, and learning: An experimental study. In *Management Science* 42 (5), pp. 659–679. DOI: 10.1287/mnsc.42.5.659.

Hoffrage, Ulrich (1999): When do people use simple heuristics, and how can we tell? In Gerd Gigerenzer, Peter M. Todd, The ABC Research Group (Eds.): Simple Heuristics That Make us Smart: Oxford University Press, pp. 141–167.

Huang, Jeff; White, Ryen; Buscher, Georg (2012): User See, User Point: Gaze and Cursor Alignment in Web Search. In Joseph A. Konstan, Ed H. Chi, Kristina Höök (Eds.): Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12). New York, NY, USA: ACM, pp. 1341–1350.

Huang, Jeff; White, Ryen W.; Dumais, Susan (2011): No Clicks, No Problem: Using Cursor Movements to Understand and Improve Search. In Desney Tan, Geraldine Fitzpatrick, Carl Gutwin, Bo Begole, Wendy A. Kellogg (Eds.): Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11), vol. 2011. New York, NY, USA: ACM, pp. 1225–1234.

Huber, Oswald (1980): The influence of some task variables on cognitive operations in an information-processing decision model. In *Acta Psychologica* 45 (1), pp. 187–196.

Hwang, Mark I. (1994): Decision making under time pressure: A model for information systems research. In *Information & Management* 27 (4), pp. 197–203. DOI: 10.1016/0378-7206(94)90048-5.

Johnson, Eric J. (1990): The adaptive decision maker: Effort and accuracy in choice. In Robin M. Hogarth (Ed.): Insights in Decision Making. A Tribute to Hillel J. Einhorn. With assistance of Hillel J. Einhorn. Chicago: Chicago Univ. Press, pp. 129–153.

Johnson, Eric J.; Payne, John W. (1985): Effort and accuracy in choice. In *Management Science* 31 (4), pp. 395–414.

Johnson, Eric J.; Payne, John W.; Bettman, James R. (1993): Adapting to time constraints. In Ola Svenson, A. John Maule (Eds.): Time Pressure and Stress in Human Judgment and Decision Making. Boston, MA: Springer US, pp. 103–116.

Johnson, Eric J.; Payne, John W.; Schkade, David A.; Bettman, James R. (1989): Monitoring Information Processing and Decisions – The Mouselab System. Fuqua School of Business, Duke University, Durham. Center for Decision Studies.

Johnson, Eric J.; Schulte-Mecklenbeck, Michael; Willemsen, Martijn C. (2008): Process models deserve process data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). In *Psychological Review* 115 (1), pp. 263–273. DOI: 10.1037/0033-295X.115.1.263.

Just, Marcel Adam; Carpenter, Patricia A. (1976): Eye fixations and cognitive processes. In *Cognitive Psychology* 8 (4), pp. 441–480.

Kagel, John H.; Roth, Alvin E. (Eds.) (1995): The Handbook of Experimental Economics. Princeton, N.J.: Princeton University Press.

Kahneman, Daniel (1973): Attention and Effort. Englewood Cliffs, N.J.: Prentice-Hall (Prentice-Hall series in experimental psychology).

Kahneman, Daniel (2012): Schnelles Denken, Langsames Denken. 21. Aufl. München: Siedler.

Kahneman, Daniel; Tursky, Bernard; Shapiro, David; Crider, Andrew (1969): Pupillary, heart rate, and skin resistance changes during a mental task. In *Journal of Experimental Psychology* 79 (1, Pt.1), pp. 164–167.

Kaplan, Martin F.; Wanshula, L. Tatiana; Zanna, Mark P. (1993): Time pressure and information integration in social judgment. In Ola Svenson, A. John Maule (Eds.): Time Pressure and Stress in Human Judgment and Decision Making. Boston, MA: Springer US, pp. 255–267.

Kaufman, Leonard; Rousseeuw, Peter J. (2008): Finding Groups in Data: An Introduction to Cluster Analysis. Hoboken, NJ, USA: John Wiley & Sons, Inc (Wiley Series in Probability and Statistics, 344).

Kelton, W. David; Barton, Russell R. (2003): Experimental design for simulation: experimental design for simulation. In J. A. Joines, R. R. Barton, K. Kang, P. A. Fishwick (Eds.): Proceedings of the 35th conference on Winter simulation driving innovation. Winter Simulation Conference 2003; ACM Special Interest Group on Simulation and Modeling. s.l.: Winter Simulation Conference, pp. 59–65.

Kleinmuntz, Don N.; Schkade, David A. (1990): Cognitive processes and information displays in computer-supported decision making: implications for research. University of Illinois. Urbana-Champaign, Ill., USA (BEBR Faculty Working Paper, 1625).

Knaak, Manuela (2013): Untersuchungen zur Objektklassifikation in digitalen Bildfolgen des Straßenverkehrs. Studienarbeit. Technische Hochschule Dresden, Dresden. Institut für Verkehrstelematik. Available online at www.elib.dlr.de/83050/1/Studienarbeit_ManuelaKnaak.pdf, checked on 2017-07-06.

Kocher, Martin G.; Sutter, Matthias (2006): Time is money – Time pressure, incentives, and the quality of decision-making. In *Journal of Economic Behavior & Organization* 61 (3), pp. 375–392.

Kühberger, Anton; Schulte-Mecklenbeck, Michael; Ranyard, Rob (2011): Introduction: Windows for understanding the mind. In Michael Schulte-Mecklenbeck, Anton Kühberger, Rob Ranyard (Eds.): A Handbook of Process Tracing Methods for Decision Research. A Critical Review and User's Guide. New York: Psychology Press (Society for judgment and decision making series), pp. 1–17.

Lambertini, Luca (1998): Time consistency in games of timing. Università degli Studi di Bologna. Bologna (Quaderni-Working Paper, 302).

Leyton-Brown, Kevin; Shoham, Yoav (2008): Essentials of game theory: A concise multidisciplinary introduction. In *Synthesis Lectures on Artificial Intelligence and Machine Learning* 2 (1), pp. 1–88.

Lindner, Florian (2014): Decision time and steps of reasoning in a competitive market entry game. In *Economics Letters* 122 (1), pp. 7–11.

Lindner, Florian; Sutter, Matthias (2013): Level-*k* reasoning and time pressure in the 11–20 money request game. In *Economics Letters* 120 (3), pp. 542–545. DOI: 10.1016/j.econlet.2013.06.005.

MacGregor, Donald (1993): Time pressure and task adaptation. In Ola Svenson, A. John Maule (Eds.): Time Pressure and Stress in Human Judgment and Decision Making. Boston, MA: Springer US, pp. 73–82.

Manor, Barry R.; Gordon, Evian (2003): Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. In *Journal of Neuroscience Methods* 128 (1), pp. 85–93.

Matsumoto, Akio; Szidarovszky, Ferenc (2016): Game Theory and Its Applications. Tokyo: Springer Japan.

Matthews, Steven A.; Postlewaite, Andrew (1989): Pre-play communication in two-person sealed-bid double auctions. In *Journal of Economic Theory* 48 (1), pp. 238–263.

Maule, John A.; Hockey, G. Robert J.; Bdozla, L. (2000): Effects of time-pressure on decision-making under uncertainty changes in affective state and information processing strategy. In *Acta Psychologica* 104, pp. 283–301.

Mellers, Barbara A.; Ordoñez, Lisa D.; Birnbaum, Michael H. (1992): A change-of-process theory for contextual effects and preference reversals in risky decision making. In *Organizational Behavior and Human Decision Processes* 52 (3), pp. 331–369.

Miller, George A. (1956): The magical number seven, plus or minus two: Some limits on our capacity for processing information. In *Psychological Review* 63 (2), pp. 81–97. DOI: 10.1037/h0043158.

Miller, James G. (1960): Information input overload and psychopathology. In *American Journal of Psychiatry* 116 (8), pp. 695–704. DOI: 10.1176/ajp.116.8.695.

Montgomery, Henry; Svenson, Ola (1976): On decision rules and information processing strategies for choices among multiattribute alternatives. In *Scandinavian Journal of Psychology* 17 (1), pp. 283–291.

Moran, Thomas P.; Newell, Allen; Card, Stuart K. (1983): The Psychology of Human-Computer Interaction. Boca Raton: Chapman and Hall/CRC.

Musch, Jochen; Reips, Ulf-Dietrich (2000): A brief history of Web experimenting. In Michael H. Birnbaum (Ed.): Psychological Experiments on the Internet. San Diego: Elsevier, pp. 61–87.

Newell, Allen; Shaw, John Calman; Simon, Herbert A. (1958): Elements of a theory of human problem solving. In *Psychological Review* 65 (3), p. 151.

Newell, Allen; Simon, Herbert A. (1972): Human Problem Solving. 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall (104).

Nisbett, Richard E.; Wilson, Timothy D. (1977a): Telling more than we can know: Verbal reports on mental processes. In *Psychological Review* 84 (3), p. 231.

Nisbett, Richard E.; Wilson, Timothy D. (1977b): The halo effect: Evidence for unconscious alteration of judgments. In *Journal of Personality and Social Psychology* 35 (4), pp. 250–256. DOI: 10.1037/0022-3514.35.4.250.

Nutt, Paul C. (1998): Framing strategic decisions. In *Organization Science* 9 (2), pp. 195–216.

Ordóñez, Lisa D.; Benson, Lehman; Pittarello, Andrea (2015): Time-pressure Perception and Decision Making. In Gideon Keren, George Wu (Eds.): The Wiley-Blackwell Handbook of Judgment and Decision Making, vol. 80. Chichester, West Sussex: Wiley-Blackwell, pp. 517–542.

Oxford Dictionaries (2018a): effectiveness | Definition of effectiveness in English by Oxford Dictionaries. Available online at https://en.oxforddictionaries.com/definition/effectiveness, checked on 2018-05-27.

Oxford Dictionaries (2018b): efficiency | Definition of efficiency in English by Oxford Dictionaries. Available online at https://en.oxforddictionaries.com/definition/efficiency, checked on 2018-05-27.

Payne, John W. (1976): Task complexity and contingent processing in decision making: An information search and protocol analysis. In *Organizational Behavior and Human Performance* 16 (2), pp. 366–387. DOI: 10.1016/0030-5073(76)90022-2.

Payne, John W. (1982): Contingent decision behavior. In *Psychological Bulletin* 92 (2), pp. 382–402. DOI: 10.1037/0033-2909.92.2.382.

Payne, John W.; Bettman, James R.; Johnson, Eric J. (1988): Adaptive strategy selection in decision making. In *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14 (3), pp. 534–611.

Payne, John W.; Bettman, James R.; Johnson, Eric J. (1992): Behavioral decision research: A constructive processing perspective. In *Annual Review of Psychology* 43 (1), pp. 87–131.

Payne, John W.; Venkatraman, Vinod (2011): Opening the black box: conclusions to a handbook of process tracing methods for decision research. In Michael Schulte-Mecklenbeck, Anton Kühberger, Rob Ranyard (Eds.): A Handbook of Process Tracing Methods for Decision Research. A Critical Review and User's Guide. New York: Psychology Press (Society for judgment and decision making series), pp. 223–251.

Polonio, Luca; Di Guida, Sibilla; Coricelli, Giorgio (2015): Strategic sophistication and attention in games: An eye-tracking study. In *Games and Economic Behavior* 94, pp. 80–96. DOI: 10.1016/j.geb.2015.09.003.

Ranyard, R.; Svenson, O. (2011): Verbal data and decision process analysis. In Michael Schulte-Mecklenbeck, Anton Kühberger, Rob Ranyard (Eds.): A Handbook of Process Tracing Methods for Decision Research. A Critical Review and User's Guide. New York: Psychology Press (Society for judgment and decision making series), pp. 115–138.

Rastegary, Haleh; Landy, Frank J. (1993): The Interactions among Time Urgency, Uncertainty, and Time Pressure. In Ola Svenson, A. John Maule (Eds.): Time Pressure and Stress in Human Judgment and Decision Making. Boston, MA: Springer US, pp. 217–239.

Reutskaja, Elena; Nagel, Rosemarie; Camerer, Colin F.; Rangel, Antonio (2011): Search dynamics in consumer choice under time pressure: An eye-tracking study. In *The American Economic Review* 101 (2), pp. 900–926. DOI: 10.1257/aer.101.2.900.

Roberson, Brian (2006): The Colonel Blotto game. In *Economic Theory* 29 (1), pp. 1–24.

Robert, Christian; Casella, George (2010): Introducing Monte Carlo Methods with R. New York, NY: Springer Science+Business Media LLC (Use R). Available online at http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10355150.

Rousseeuw, Peter J. (1987): Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. In *Journal of Computational and Applied Mathematics* 20, pp. 53–65. DOI: 10.1016/0377-0427(87)90125-7.

Rumsey, Deborah J. (2010): Statistics Essentials for Dummies. Hoboken, NJ, USA: John Wiley & Sons.

Russo, J. Edward (1978): Adaptation of cognitive processes to the eye movement system. In John W. Senders (Ed.): Eye Movements and the Higher Psychological Functions. London: Taylor and Francis (Psychology Library Editions, v.26), pp. 89–112.

Russo, J. Edward; Dosher, Barbara A. (1983): Strategies for multiattribute binary choice. In *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9 (4), p. 676.

Russo, J. Edward; Johnson, Eric J.; Stephens, Debra L. (1989): The validity of verbal protocols. In *Memory & Cognition* 17 (6), pp. 759–769.

Saaty, Thomas L. (2008): Decision making with the analytic hierarchy process. In *International Journal of Services Sciences* 1 (1), pp. 83–98. DOI: 10.1504/IJSSCI.2008.017590.

Salthouse, Timothy A.; Ellis, Cecil L. (1980): Determinants of eye-fixation duration. In *The American Journal of Psychology*, pp. 207–234.

Schell, Jesse (2008): The Art of Game Design: A Book of Lenses. Burlington, USA: Morgan Kaufmann.

Schulte Mecklenbeck, Michael; Kuhberger, Anton; Ranyard, Rob (2011a): The role of process data in the development and testing of process models of judgment and decision making. In *Judgment and Decision Making* 6 (8), pp. 733–739. Available online at http://journal.sjdm.org/11/m36/m36.html.

Schulte-Mecklenbeck, Michael; Kühberger, Anton; Ranyard, Rob (Eds.) (2011b): A Handbook of Process Tracing Methods for Decision Research. A Critical Review and User's Guide. New York: Psychology Press (Society for judgment and decision making series).

Simon, Herbert A. (1974): How big is a chunk? In *Science* 183 (4124), pp. 482–488.

Simon, Herbert A. (1978): Rationality as process and as product of thought. In *The American Economic Review* 68 (2), pp. 1–16.

Simon, Herbert A.; Newell, Allen (1971): Human problem solving: The state of the theory in 1970. In *American Psychologist* 26 (2), p. 145.

Singer, Eleanor; van Hoewyk, John; Gebler, Nancy; McGonagle, Katherine (1999): The effect of incentives on response rates in interviewer-mediated surveys. In *Journal of Official Statistics* 15 (2), p. 217.

Stahl, Dale O.; Wilson, Paul W. (1995): On players′ models of other players: Theory and experimental evidence. In *Games and Economic Behavior* 10 (1), pp. 218–254.

Sutter, Matthias; Kocher, Martin; Strauß, Sabine (2003): Bargaining under time pressure in an experimental ultimatum game. In *Economics Letters* 81 (3), pp. 341–347.

Tan, Pang-Ning; Steinbach, Michael; Kumar, Vipin (2013): Introduction to Data Mining: Pearson New International Edition. 2nd ed. New York: Pearson Education Limited.

Thomas, Ewart A. C.; Weaver, Wanda B. (1975): Cognitive processing and time perception. In *Perception & Psychophysics* 17 (4), pp. 363–367.

Thorngate, Warren (1980): Efficient decision heuristics. In *Behavioral Science* 25 (3), pp. 219–225. DOI: 10.1002/bs.3830250306.

Tversky, Amos (1969): Intransitivity of preferences. In *Psychological Review* 76 (1), pp. 31–48.

Tversky, Amos (1972): Elimination by aspects: A theory of choice. In *Psychological Review* 79 (4), pp. 281–299. DOI: 10.1037/h0032955.

Tversky, Amos; Kahneman, Daniel (1981): The framing of decisions and the psychology of choice. In *Science* 211 (4481), pp. 453–458.

Tversky, Amos; Kahneman, Daniel (1986): Rational choice and the framing of decisions. In *The Journal of Business,* 59 (4 (Teil 2)), pp. 5251–5278.

VanRullen, Rufin; Thorpe, Simon J. (2001a): Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. In *Perception;* 30 (6), pp. 655–668. DOI: 10.1068/p3029.

VanRullen, Rufin; Thorpe, Simon J. (2001b): The time course of visual processing: From early perception to decision-making. In *Journal of Cognitive Neuroscience* 13 (4), pp. 454–461. DOI: 10.1162/08989290152001880.

Verband der Hochschullehrer für Betriebswirtschaft e.V. (2018): Teilrating Allgemeine Betriebswirtschaftslehre. Available online at http://vhbonline.org/vhb4you/jourqual/vhb-jourqual-3/teilrating-abwl/, checked on 2018-05-21.

Vogt, Herbert (1988): Methoden der Statistischen Qualitätskontrolle. Wiesbaden: Vieweg+Teubner Verlag (Mathematische Methoden in der Technik, 10).

Vorob′ev, N. N. (1972): Grundlagen der Spieltheorie und ihre Praktische Bedeutung. 2., verb. und erw. Aufl. Würzburg: Physica-Verlag (Physica-Paperback).

Wang, Joseph Tao-yi (2011): Pupil dilation and eye tracking. In Michael Schulte-Mecklenbeck, Anton Kühberger, Rob Ranyard (Eds.): A Handbook of Process Tracing Methods for Decision Research. A Critical Review and User's Guide. New York: Psychology Press (Society for judgment and decision making series), pp. 185–204.

Wickens, C. D. (1981): Processing Resources in Attention, Dual Task Performance, and Workload Assessment. University of Illinois, Urbana-Champaign, Ill., USA. Engineering-Psychology Research Laboratory.

Willemsen, Martijn C.; Johnson, Eric J. (2011): Visiting the decision factory: observing cognition with MouselabWEB and other information acquisition methods. In Michael Schulte-Mecklenbeck, Anton Kühberger, Rob Ranyard (Eds.): A Handbook of Process Tracing Methods for Decision Research. A Critical Review and User's Guide. New York: Psychology Press (Society for judgment and decision making series), pp. 21–42.

Woods, David D. (1993): Process tracing methods for the study of cognition outside of the experimental psychology laboratory. In Gary A. Klein, Judith Orasanu, Roberta Calderwood, Caroline E. Zsambok (Eds.): Decision Making in Action. Models and Methods. 1st ed. Norwood, NJ: Ablex, pp. 228–251.

Wright, Peter (1974): The harassed decision maker: Time pressures, distractions, and the use of evidence. In *Journal of Applied Psychology* 59 (5), pp. 555–561.

Yu, Julie; Cooper, Harris (1983): A quantitative review of research design effects on response rates to questionnaires. In *Journal of Marketing Research* 20 (02), pp. 36–44.

Zakay, Dan (1993): The impact of time perception processes on decision making under time stress. In Ola Svenson, A. John Maule (Eds.): Time Pressure and Stress in Human Judgment and Decision Making. Boston, MA: Springer US, pp. 59–72.

Zakay, Dan; Wooler, Stuart (1984): Time pressure, training and decision effectiveness. In *Ergonomics* 27 (3), pp. 273–284.

Appendix A Game-theoretic Glossary

The technical concepts of game theory and their corresponding definitions, which are presented below (Table 40), supplement the explanations given in Chapter 2 and are intended to help understand this treatise's Chapter 3. It is by no means a complete collection of game-theoretic vocabulary used in the variety of game-theoretic concepts and studies or an attempt to give a comprehensive terminology of game theory. The definitions of the terms presented here are all taken from Vorob′ev (1972) if no other source is given. Similar explanations can surely be found in any other introductive treatise of game theory. Note that the definitions characterize two person games (i.e. games that are played by interaction of two players) in normal-form.

TABLE 40 – GAME-THEORETIC TERMINOLOGY

Term	**Definition**
Accuracy in choice	Consistency with the normatively optimal choice (Johnson and Payne 1985, p. 397).
Anticipation of a game	Strategic reflection on the consequences of choosing one alternative over another in a game in order to identify the (subjective) optimal strategy (Brams 2000, p. 25, footnote 11).
Equilibrium	A set of strategies, one for each player, such that no player has an incentive to unilaterally change action.
Game	Strategic conflict situation, where the decision of one player affects the payoff of the other. It is fully described by its players, the players' strategies and the resulting payoffs for each outcome.
(Im-)Perfect Information	(At least one/) All player/s (do not) know the (/at least one) choices previously made by all others. (Since normal-form games are simultaneous, they are of Imperfect Information.)
(In-)Complete Information	(At least one/) Every player (do not fully) know/s the payoff- and strategy set of the other player/s. (Normal-form games are of Complete Information.)
Information Set	Entirety of choice alternatives and payoff information accessible to a player in a game.
Irrationality / bounded Rationality	Player's inability / limited ability to generate a strategic decision.
Non-/Cooperative Games	Purpose of player in game is competitive / cooperative payoff gain.

Term	Definition
Outcome	A player's payoff in a game, given a combination of all players' strategies.
Payoff Matrix	Entirety of payoff information for all players and all combinations of players' strategies.
Rationality	Player's ability to generate a strategic decision.
Player	Involved decision maker in a game; often regarded as aware of an opponent and thus somehow rational.
Strategic Decision-making	Generating a Decision that takes into account the opponent's alternatives. Mostly used in context of game-theoretic problem solution concepts.
Strategy	Choice alternative of a player.

Appendix B Payoff Matrices of Set of Game Tasks

This appendix provides the payoff matrices applied in the simulation and in the experiment (versions V_1, V_2 and V_3). In case of the simulation all presented matrices are played in all possible sizes (2 by 2 alternatives up to 5 by 5 alternatives). For a complexity level of n (with $n \in \{2,3,4,5\}$ the first n alternatives formed the corresponding matrix (beginning in the upper left corner). The only exception are the game tasks Costa 2A, 2B, 3A, 3B, which are adopted from Costa-Gomes et al. (2001) and exclusively have a size of 2 by 2. These four game tasks also represent the first round of the experiment. In case of the experiment only the matrices of games that are actually in use are presented. The level of complexity corresponds with the round the game tasks are used in, i.e. in each round one level of complexity is employed.

The payoff matrix shows the payoff values for both players, separated by a semicolon. The first value belongs to the row player and the second to the column player. Note that in case of the experiment all subjects act from a row player's perspective. Subjects that are recorded as column players get a transformed payoff matrix, where contents of row and column players are interchanged. The payoff matrices of the experimental tasks are presented in tables that provide additional information. The upper part contains the payoff matrices in the lower right cell. In the upper left is the name of the game task. The cell below, left to the payoff matrix, additionally depict the heuristics that would select the respective alternatives when applied in the task. Those entries are valid for both row and column player as long as no other entries are given for the column player in the cell above the payoff matrix. In the last row of the table one can find the four random numbers used to multiply with the payoff matrix entries in order to create seemingly different matrices within one round of experimenting. The numbers are depicted in order of occurrence from left in first task of the round to right in the last task of the round.

Simulation

TABLE 41 – SIMULATION/V_1/V_2/V_3 COSTA 2A, ROUND 1, 2 BY 2, TASK 1

Costa 2A	$\mathbf{A, P, O, N}$	$\mathbf{D1, D2, L2, E}$
A	$72; 93$	$31; 46$
$P, O, N, D1, D2, L2, E$	$84; 52$	$55; 79$

TABLE 42 – SIMULATION/V_1/V_2/V_3 COSTA 3A, ROUND 1, 2 BY 2, TASK 2

Costa 3A	$\mathbf{D1, D2, L2, E}$	$\mathbf{A, P, O, N}$
$D1, D2, L2, E$	$75; 51$	$42; 27$
A, P, O, N	$48; 80$	$89; 68$

TABLE 43 – SIMULATION/V_1/V_2/V_3 COSTA 2B, ROUND 1, 2 BY 2, TASK 3

Costa 2B	$\mathbf{A, P, O, N}$	$\mathbf{D1, D2, L2, E}$
$P, O, N, L2, D1, D2, E$	$94; 23$	$38; 57$
A	$45; 89$	$14; 18$

TABLE 44 – SIMULATION/V_1/V_2/V_3 COSTA 3B, ROUND 1, 2 BY 2, TASK 4

Costa 3B	$\mathbf{P, O, N, L2}$ $\mathbf{D1, D2, E}$	$\mathbf{A}$
A, P, O, N	$21; 92$	$87; 43$
$D1, D2, L2, E$	$55; 36$	$16; 12$

TABLE 45 – SIMULATION BATTLE OF SEXES

Battle of Sexes					
	2; 1	0; 0	0; 0,5	5; 0,75	5; 0,75
	0; 0	1; 2	0; 1,5	5; 0,75	5; 0,25
	0,5; 0	1,5; 0	1,5; 1,5	5; 0,75	5; 0
	0,75; 5	0,75; 5	0,75; 5	5; 5	5; 5
	1; 5	0,25; 5	0; 5	5; 5	4; 4

TABLE 46 – SIMULATION CHICKEN-GAME

Chicken					
	4; 4	2; 6	0; 3	10; 5	9; 3
	6; 2	0; 0	0; 3	5; 1	9; 2
	3; 0	3; 0	3; 3	10; 0	9; 1
	5; 10	1; 5	0; 10	7; 7	9; 10
	3; 9	2; 9	1; 9	10; 9	9; 9

TABLE 47 – SIMULATION/V_3 HAWK-DOVE-GAME, ROUND 4, 5 BY 5

Hawk Dove					
A, O	0; 0	1; 0	0; 2	1; 1	2; 2
A, O	0; 1	1,5; 1,5	0; 2	1; 2	2; 2
$A, P, O, N, D1, D2, E, L2$	2; 0	2; 0	2; 2	1; 2	2; 0
$O, P, L2$	1; 1	2; 1	2; 1	1; 1	2; 1
A, O	2; 2	2; 2	0; 2	1; 2	2; 2
Multiplier Task 1 to 4	62; 21; 69; 20				

TABLE 48 – SIMULATION/V_1 PRISONERS' DILEMMA, ROUND 4, 5 BY 5

Prisoners' Dilemma					
P, O	$-2; -2$	$0; -6$	$-4; -1{,}5$	$-4; -2{,}25$	$-4; -2$
A	$-6; 0$	$-1; -1$	$-6; -1{,}5$	$-4; -1{,}25$	$-4; -0{,}5$
$P, N, D1, D2, E, L2$	$-1{,}5; -4$	$-1{,}5; -6$	$-1{,}5; -1{,}5$	$-4; -2{,}5$	$-4; -5$
P	$-2{,}25; -4$	$-1{,}25; -4$	$-2{,}5; -4$	$-4; -4$	$-4; -4$
$-$	$-2; -4$	$-0{,}5; -4$	$-5; -4$	$-4; -4$	$-4; -4$
Multiplier Task 1 to 4	89; 62; 33; 43				

TABLE 49 – SIMULATION/V_2 STACKELBERG'S LEADERSHIP, ROUND 4, 5 BY 5

Stackelberg's Leadership					
O	$0; 2$	$3; 0$	$0; 1{,}5$	$5; 1{,}75$	$5; 1{,}75$
O	$2; 1$	$1; 3$	$0; 1{,}5$	$5; 1$	$5; 1$
$O, P, N, D1, D2, E, L2$	$1{,}5; 0$	$1{,}5; 0$	$1{,}5; 1{,}5$	$5; 0{,}5$	$5; 0{,}25$
A, O	$1{,}75; 5$	$1; 5$	$0{,}5; 5$	$5; 5$	$4; 5$
A, O	$1{,}75; 5$	$1; 5$	$0{,}25; 5$	$5; 4$	$5; 5$
Multiplier Task 1 to 4	28; 29; 59; 53				

TABLE 50 – SIMULATION THROWING FINGERS

Throwing Fingers					
	$1; -1$	$-1; 1$	$0; 0$	$5; -1$	$5; 0$
	$-1; 1$	$1; -1$	$0; 0$	$5; 0$	$5; 0$
	$0; 0$	$0; 0$	$0; 0$	$5; 0$	$5; 0$
	$-1; 5$	$0; 5$	$0; 5$	$5; 5$	$5; 0$
	$0; 5$	$-1; 5$	$0; 5$	$0; 5$	$5; 5$

V_1

For Round 1 see Table 41 to Table 44.

TABLE 51 – V_1 CHICKEN-GAME, ROUND 2, 3 BY 3

Chicken			
$A, D1, D2, E$	$4; 4$	$2; 6$	$0; 3$
$O, D1, D2, E$	$6; 2$	$0; 0$	$0; 3$
$P, N, D1, D2, E, L2$	$3; 0$	$3; 0$	$3; 3$
Multiplier Task 1 to 4	36; 32; 95; 85		

TABLE 52 – V_1 HAWK-DOVE-GAME, ROUND 3, 4 BY 4

Hawk-Dove				
$-$	$0; 0$	$1; 0$	$0; 2$	$1; 1$
$-$	$0; 1$	$1{,}5; 1{,}5$	$0; 2$	$1; 2$
$N, O, P, A, D1, D2, E, L2$	$2; 0$	$2; 0$	$2; 2$	$1; 2$
$O, P, L2$	$1; 1$	$2; 1$	$2; 1$	$1; 1$
Multiplier Task 1 to 4	54; 53; 93; 78			

For Round 4, see Table 48.

V_2

For Round 1 see Table 41 to Table 44.

TABLE 53 – V_2 HAWK-DOVE-GAME, ROUND 2, 3 BY 3

Hawk-Dove			
$-$	0; 0	1; 0	0; 2
$-$	0; 1	1,5; 1,5	0; 2
$N, O, P, A, D1, D2, E, L2$	2; 0	2; 0	2; 2
Multiplier Task 1 to 4	19; 66; 14; 98		

TABLE 54 – V_2 CHICKEN-GAME, ROUND 3, 4 BY 4

Chicken				
A, O	4; 4	2; 6	0; 3	10; 5
$-$	6; 2	0; 0	0; 3	5; 1
$O, P, N, D1, D2, E, L2$	3; 0	3; 0	3; 3	10; 0
A	5; 10	1; 5	0; 10	7; 7
Multiplier Task 1 to 4	62; 69; 40; 4			

For Round 4, see Table 49.

V_3

For Round 1 see Table 41 to Table 44.

TABLE 55 – V_3 CHICKEN-GAME, ROUND 2, 3 BY 3

Chicken			
$A, D1, D2, E$	$4; 4$	$2; 6$	$0; 3$
$A, O, D1, D2, E$	$6; 2$	$0; 0$	$0; 3$
$P, N, D1, D2, E, L2$	$3; 0$	$3; 0$	$3; 3$
Multiplier Task 1 to 4	36; 90; 72; 52		

TABLE 56 – V_3 PRISONERS' DILEMMA, ROUND 3, 4 BY 4

Prisoners' Dilemma				
P, O	$-2; -2$	$0; -6$	$-4; -1{,}5$	$-4; -2{,}25$
A	$-6; 0$	$-1; -1$	$-6; -1{,}5$	$-4; -1{,}25$
$P, N, D1, D2, E, L2$	$-1{,}5; -4$	$-1{,}5; -6$	$-1{,}5; -1{,}5$	$-4; -2{,}5$
P	$-2{,}25; -4$	$-1{,}25; -4$	$-2{,}5; -4$	$-4; -4$
Multiplier Task 1 to 4	40; 84; 45; 91			

For Round 4 see Table 47.

Appendix C Ranking Tables

Table 57 and Table 58 below represent the ranking results for the heuristics from simulation. In the first table, the development of the rankings under varying complexity is presented. The second table shows the developments of the rankings under changing time limits. The shades of gray correspond to the heuristics' ranks with rank one having the lightest shade and rank nine the darkest.

TABLE 57 – COMPLEXITY RANKING

Time Stage	Complexity	2x2				3x3				4x4				5x5				mean			
	Goal / Heuristic	G2	G3	G4	Ø	G2	G3	G4	Ø	G2	G3	G4	Ø	G2	G3	G4	Ø	ØG2	ØG3	ØG4	ØallG
early	*Altruist*	2	9	8	6,3	2	9	6	5,7	3	9	6	6,0	3	9	6	6,0	2,5	9,0	6,5	6,0
	Random	9	8	1	6,0	9	8	1	6,0	9	8	1	6,0	9	8	1	6,0	9,0	8,0	1,0	6,0
	Optimist	1	4	3	2,7	1	7	3	3,7	1	4	4	3,0	1	7	4	4,0	1,0	5,5	3,5	3,3
	Pessimist	5	7	9	7,0	4	2	5	3,7	4	3	5	4,0	4	3	5	4,0	4,3	3,8	6,0	4,7
	Naïve	2	6	2	3,3	3	1	2	2,0	2	1	2	1,7	2	1	2	1,7	2,3	2,3	2,0	2,2
	L2	4	1	4	3,0	5	3	4	4,0	5	2	3	3,3	5	2	3	3,3	4,8	2,0	3,5	3,4
	D1	6	3	6	5,0	6	4	7	5,7	6	5	7	6,0	6	4	7	5,7	6,0	4,0	6,8	5,6
	D2	7	5	6	6,0	7	5	9	7,0	7	7	9	7,7	8	6	9	7,7	7,3	5,8	8,3	7,1
	Equilibrium	8	2	6	5,3	8	6	8	7,3	8	6	8	7,3	7	5	8	6,7	7,8	4,8	7,5	6,7
middle	*Altruist*	3	9	5	5,7	3	9	6	6,0	3	9	6	6,0	3	9	6	6,0	3,0	9,0	5,8	5,9
	Random	9	8	1	6,0	9	8	1	6,0	9	8	1	6,0	9	8	1	6,0	9,0	8,0	1,0	6,0
	Optimist	3	7	3	4,3	3	7	3	4,3	3	6	4	4,3	3	7	4	4,7	3,0	6,8	3,5	4,4
	Pessimist	3	6	6	5,0	3	1	5	3,0	3	3	5	3,7	3	3	5	3,7	3,0	3,3	5,3	3,8
	Naïve	3	5	2	3,3	3	2	2	2,3	3	1	2	2,0	3	1	2	2,0	3,0	2,3	2,0	2,4
	L2	3	1	4	2,7	3	3	4	3,3	3	2	3	2,7	3	2	3	2,7	3,0	2,0	3,5	2,8
	D1	6	2	7	5,0	6	5	7	6,0	6	5	7	6,0	6	4	7	5,7	6,0	4,0	7,0	5,7
	D2	7	3	9	6,3	8	6	9	7,7	8	7	9	8,0	8	6	9	7,7	7,8	5,5	9,0	7,4
	Equilibrium	8	4	8	6,7	7	4	8	6,3	7	4	8	6,3	7	5	8	6,7	7,3	4,3	8,0	6,5
late	*Altruist*	3	9	5	5,7	3	9	6	6,0	3	9	6	6,0	3	9	6	6,0	3,0	9,0	5,8	5,9
	Random	9	8	1	6,0	9	8	1	6,0	9	8	1	6,0	9	8	1	6,0	9,0	8,0	1,0	6,0
	Optimist	3	7	3	4,3	3	7	3	4,3	3	7	4	4,7	3	7	4	4,7	3,0	7,0	3,5	4,5
	Pessimist	3	6	6	5,0	3	1	5	3,0	3	5	5	4,3	3	4	5	4,0	3,0	4,0	5,3	4,1
	Naïve	3	5	2	3,3	3	2	2	2,3	3	1	2	2,0	3	1	2	2,0	3,0	2,3	2,0	2,4
	L2	3	4	4	3,7	3	3	4	3,3	3	2	3	2,7	3	2	3	2,7	3,0	2,8	3,5	3,1
	D1	8	2	7	5,7	7	5	7	6,3	6	3	7	5,3	6	3	7	5,3	6,8	3,3	7,0	5,7
	D2	6	3	9	6,0	8	6	9	7,7	8	6	9	7,7	8	6	9	7,7	7,5	5,3	9,0	7,3
	Equilibrium	7	1	8	5,3	6	4	8	6,0	7	4	8	6,3	7	5	8	6,7	6,8	3,5	8,0	6,1

constant	*Altruist*	3	9	5	5,7	3	9	6	6,0	3	9	6	6,0	3	8	6	5,7	3,0	8,8	5,8	5,8
	Random	9	8	1	6,0	9	8	1	6,0	9	8	1	6,0	9	7	1	5,7	9,0	7,8	1,0	5,9
	Optimist	3	6	3	4,0	3	7	3	4,3	3	7	4	4,7	3	6	4	4,3	3,0	6,5	3,5	4,3
	Pessimist	3	5	6	4,7	3	1	5	3,0	3	4	5	4,0	3	5	5	4,3	3,0	3,8	5,3	4,0
	Naïve	3	7	2	4,0	3	2	2	2,3	3	1	2	2,0	3	1	2	2,0	3,0	2,8	2,0	2,6
	L2	3	2	4	3,0	3	3	4	3,3	3	2	3	2,7	3	1	3	2,3	3,0	2,0	3,5	2,8
	D1	6	4	7	5,7	7	5	7	6,3	6	5	7	6,0	6	2	7	5,0	6,3	4,0	7,0	5,8
	D2	7	3	9	6,3	6	6	9	7,0	7	6	9	7,3	7	4	9	6,7	6,8	4,8	9,0	6,8
	Equilibrium	8	1	8	5,7	8	4	8	6,7	8	3	8	6,3	8	3	8	6,3	8,0	2,8	8,0	6,3

TABLE 58 – HEURISTICS RANKING

Time Stage	Heuristic / Goal / Complexity	Nonstrategic Heuristics															Strategic Heuristics											
		Altruist			*Random*			*Optimist*			*Pessimist*			*Naïve*			*L2*			*D1*			*D2*			*Equilibrium*		
		G2	G3	G4	G2	G3	G4	G2	G3	G4	G2	G3	G4	G2	G3	G4	G2	G3	G4	G2	G3	G4	G2	G3	G4	G2	G3	G4
early	2x2	2	9	8	9	8	1	1	4	3	5	7	9	2	6	2	4	1	4	6	3	6	7	5	6	8	2	6
	3x3	2	9	6	9	8	1	1	7	3	4	2	5	3	1	2	5	3	4	6	4	7	7	5	9	8	6	8
	4x4	3	9	6	9	8	1	1	4	4	4	3	5	2	1	2	5	2	3	6	5	7	7	7	9	8	6	8
	5x5	3	9	6	9	8	1	1	7	4	4	3	5	2	1	2	5	2	3	6	4	7	8	6	9	7	5	8
	Ø	2,5	9	6,5	9	8	1	1	5,5	3,5	4,3	3,8	6	2,3	2,3	2	4,8	2	3,5	6	4	6,8	7,3	5,8	8,3	7,8	4,8	7,5
middle	2x2	3	9	5	9	8	1	3	7	3	3	6	6	3	5	2	3	1	4	6	2	7	7	3	9	8	4	8
	3x3	3	9	6	9	8	1	3	7	3	3	1	5	3	2	2	3	3	4	6	5	7	8	6	9	7	4	8
	4x4	3	9	6	9	8	1	3	6	4	3	3	5	3	1	2	3	2	3	6	5	7	8	7	9	7	4	8
	5x5	3	9	6	9	8	1	3	7	4	3	3	5	3	1	2	3	2	3	6	4	7	8	6	9	7	5	8
	Ø	3	9	5,8	9	8	1	3	6,8	3,5	3	3,3	5,3	3	2,3	2	3	2	3,5	6	4	7	7,8	5,5	9	7,3	4,3	8
late	2x2	3	9	5	9	8	1	3	7	3	3	6	6	3	5	2	3	4	4	8	2	7	6	3	9	7	1	8
	3x3	3	9	6	9	8	1	3	7	3	3	1	5	3	2	2	3	3	4	7	5	7	8	6	9	6	4	8
	4x4	3	9	6	9	8	1	3	7	4	3	5	5	3	1	2	3	2	3	6	3	7	8	6	9	7	4	8
	5x5	3	9	6	9	8	1	3	7	4	3	4	5	3	1	2	3	2	3	6	3	7	8	6	9	7	5	8
	Ø	3	9	5,8	9	8	1	3	7	3,5	3	4	5,3	3	2,3	2	3	2,8	3,5	6,8	3,3	7	7,5	5,3	9	6,8	3,5	8
constant	2x2	3	9	5	9	8	1	3	6	3	3	5	6	3	7	2	3	2	4	6	4	7	7	3	9	8	1	8
	3x3	3	9	6	9	8	1	3	7	3	3	1	5	3	2	2	3	3	4	7	5	7	6	6	9	8	4	8
	4x4	3	9	6	9	8	1	3	7	4	3	4	5	3	1	2	3	2	3	6	5	7	7	6	9	8	3	8
	5x5	3	8	6	9	7	1	3	6	4	3	5	5	3	1	2	3	1	3	6	2	7	7	4	9	8	3	8
	Ø	3	8,8	5,8	9	7,8	1	3	6,5	3,5	3	3,8	5,3	3	2,8	2	3	2	3,5	6,3	4	7	6,8	4,8	9	8	2,8	8
	ØØ	2,9	8,9	5,9	9	7,9	1	2,5	6,4	3,5	3,3	3,7	5,4	2,8	2,4	2	3,4	2,2	3,5	6,3	3,8	6,9	7,3	5,3	8,8	7,4	3,8	7,9
	ØØØ	5,9			6,0			4,1			4,1			2,4			3,0			5,7			7,1			6,4		

Appendix D Technical Documentation and Experiment Web Pages

The technical documentation is available only in German language.

Appendix D Technical Documentation (DEU)

Appendix D Experiment Web Page

Data is available in this treatise's digital version in *Microsoft Word* format. It is also part of the digital supplement to this treatise.

Appendix E Procedure of Hypothesis Testing

Confidence Intervals

Data evaluation with the aid of confidence intervals (CI) in the context of this analysis is applied once – namely to test Hypothesis VII. For developing the corresponding CI, Eq. (10) is used:

$$CI = \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \tag{10}$$

with $\hat{p}$ as sample proportion[183], z^* as the value from the standard normal distribution that corresponds to the sought confidence level. The square root term depicts the standard error. Together with the sample proportion, it is applied in hypothesis testing, too.

Hypothesis testing

Hypotheses are tested by applying the standard statistical procedure of hypothesis testing. The following list[184] illustrates the proceeding in five steps:

1. Set up the null and alternative hypotheses H_0 and H_1.

2. Take a random sample of individuals from the population and calculate the sample statistics according to Eq. (11) together with Eq. (12):

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{11}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \tag{12}$$

with x_i as the relevant value from dataset i of n datasets. When interested in population proportions, apply Eq. (13) as sample statistics:

[183] See Eq. (13) for a description.
[184] The proceeding of statistical analysis and hypothesis testing is often discussed in scholar books. An introduction offers Rumsey (2010, p. 94) for instance where the above description of proceeding is derived from to a great extent.

$$\hat{p} = \frac{n_i}{n} \tag{13}$$

with n_i as the number of subjects who carry the relevant characteristic(s) i within a population sized n.

3. Convert the sample statistic to a test statistic T by changing it to a standard score, normalized on standard deviation. Three variations of test statistics are applied here – depending on the hypothesis' purpose. Equation (14) is applied to testing Hypotheses II to VI and VIII:

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n_1} - \frac{s_y^2}{n_2}}} \tag{14}$$

with $\bar{x}$ and $\bar{y}$ as sample means, n_1 and n_2 as sample sizes and $\sqrt{\frac{s_x^2}{n_1} - \frac{s_y^2}{n_2}}$ as the standard error of the difference between both samples' arithmetic means. The x-component serves as the reference value to which the y-component is compared.

Equation (15) is used in the test of Hypothesis VII:

$$T = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} \tag{15}$$

with the population proportion p_0 as reference value suggested by H_0 and $\hat{p}$ as the detected proportion in the population carrying the relevant characteristic.

In case of comparing two proportions (Hypothesis IX and X), the following test statistic is used (Eq. (16)):

$$T = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_{12}(1 - \hat{p}_{12})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{16}$$

with $\hat{p}_1$ and $\hat{p}_2$ being the sample proportions to be compared, $\hat{p}_{12}$ as pooled sample proportion[185] and n_1 and n_2 as sample sizes.

4. Find the p-value for the test statistic. In cases of examining two means, one can apply Student's t-distribution since the standard deviation of the population's characteristic of interest is unknown

[185] The pooled sample proportion $\hat{p}_{12}$ is the proportion of the combined datasets of size $n_1 + n_2$ carrying the examined characteristic.

(Hypotheses II to VI and VIII). The degrees of freedom are $n_1 + n_2 - 1$. When examining population proportions, the standardized normal distribution is used to determine the p-value (Hypotheses I, VII, IX, and X).

5. Examine the p-value and make a decision. For evaluating two means or two proportions, the following holds: If the reference value is smaller than the measured value of interest, T is negative, and the corresponding p-values are smaller than 0.5. If T is zero, p is equal to 0.5 and if T is positive, p is larger 0.5. That means if the alternative hypothesis expects the value of interest to be significantly higher than the reference value, the corresponding p-values need to be reasonably small and vice versa. For evaluating one proportion, the opposite is true which follows directly from the different sign of Eq. (15) in comparison to the other two cases. One can reject the null-hypothesis H_0 if the p-value lies in a certain tail of the distribution. Which tail matters depends on the alternative Hypothesis H_1: In the case of a two-sided examination, both tails are relevant and H_0 needs to be rejected if $p < \frac{\propto}{2}$ or $p > 1 - \frac{\propto}{2}$. In case of one-sided examinations, only one tail is relevant, depending on the orientation of the H_0 inequity. If either $p < \propto$ or $p > 1 - \propto$, H_0 has to be rejected (Hypotheses II to X). In those cases, one can regard the difference between the observed x and y as unlikely enough to reject H_0.

Appendix F Experimental Results – Hypotheses Test Data

The complete dataset of experimental results is available as embedded *Excel* files in Table 59.[186]

TABLE 59 – DATASET OF EXPERIMENTAL RESULTS

Appendix F Test Results	Appendix F Sample Size	Appendix F Social Aspects

The following tables provide the p-values of hypothesis test. The level of complexity is noted as number of alternatives per player. Squaring this number gives the number of cells in the payoff matrix. Light-gray colored entries depict the case of falsifying H_0, whereas values are dark-gray colored when H_0 is retained. Entries in black color represent cases of rejecting H_0 without supporting the predicted alternative, but one with opposite orientation. In case of focussing on time limitation, the data corresponding to the maximum time limit per round represent the reference value of H_0 to which data of the other tasks are compared. When dealing with complexity, the reference value is provided from the data corresponding to the smallest complexity. The reference value is given as H_0 in the tables. Horizontal black lines in the table group data according to their reference value. The so derived sections are similar to ones separated by dashed lines in the corresponding figures (see Chapter 10).

TABLE 60 – p-VALUES OF HYPOTHESIS I: INFLUENCE OF TIME LIMITATION ON CHOICE RATE

Level of complexity	**Time limit [s]** V_1/V_3	V_2	p**-value** V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	0.66	0.73	0.98
2	7	9	0.98	0.86	0.99
2	6	8	0.63	0.97	0.97
3	75	75	H_0	H_0	H_0

[186] The *Excel* files are available in this treatise's digital version in *Microsoft Word* format. It is also part of the digital supplement to this treatise.

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
3	35	56	0.70	0.68	0.51
3	20	13	0.72	0.99	0.65
3	10	9	0.84	0.96	0.84
4	150	150	H_0	H_0	H_0
4	54	104	0.73	0.64	0.82
4	22	24	0.92	0.64	0.69
4	12	8	0.96	0.77	0.74
5	200	200	H_0	H_0	H_0
5	53	120	0.50	0.68	0.98
5	31	26	0.84	0.80	0.96
5	15	7	0.51	0.93	0.96

TABLE 61 – p-VALUES OF HYPOTHESIS I: INFLUENCE OF COMPLEXITY ON CHOICE RATE

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
3	75	75	0.85	0.96	0.54
4	150	150	0.92	0.90	0.94
5	200	200	0.91	0.98	1.00
2	13	19	H_0	H_0	H_0
3	35	56	0.88	0.96	0.95
4	54	104	0.68	0.98	0.91
5	53	120	0.86	0.98	0.80
2	7	9	H_0	H_0	H_0
3	20	13	0.98	0.57	0.96
4	22	24	0.94	0.99	0.90
5	31	26	1.00	0.98	0.82
2	6	8	H_0	H_0	H_0
3	10	9	0.59	0.95	0.71
4	12	8	0.59	0.98	0.67
5	15	7	0.94	0.99	0.65

TABLE 62 – p-VALUES OF HYPOTHESIS II: INFLUENCE OF TIME LIMITATION ON GENERATING PAYOFF

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	1.00	0.84	1.00
2	7	9	1.00	1.00	0.99
2	6	8	1.00	1.00	1.00
3	75	75	H_0	H_0	H_0
3	35	56	0.23	0.56	0.58
3	20	13	0.24	0.54	0.76
3	10	9	0.23	0.70	0.51
4	150	150	H_0	H_0	H_0
4	54	104	0.77	0.90	0.61
4	22	24	0.74	1.00	0.58
4	12	8	0.94	0.96	0.60
5	200	200	H_0	H_0	H_0
5	53	120	0.06	1.00	0.93
5	31	26	0.02	0.00	0.59
5	15	7	0.02	0.87	0.79

TABLE 63 – p-VALUES OF HYPOTHESIS II: INFLUENCE OF COMPLEXITY ON GENERATING PAYOFF

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
3	75	75	1.00	0.99	1.00
4	150	150	0.04	1.00	0.96
5	200	200	1.00	1.00	0.00
2	13	19	H_0	H_0	H_0
3	35	56	1.00	0.94	1.00
4	54	104	0.00	1.00	0.49
5	53	120	1.00	1.00	0.01
2	7	9	H_0	H_0	H_0

3	20	13	0.93	0.53	1.00
4	22	24	0.00	0.99	0.56
5	31	26	0.97	0.00	0.00
2	6	8	H_0	H_0	H_0
3	10	9	0.92	0.24	0.98
4	12	8	0.00	0.13	0.23
5	15	7	0.97	0.03	0.00

TABLE 64 – p-VALUES OF HYPOTHESIS III: INFLUENCE OF TIME LIMITATION

Level of complexity	**Time limit [s]** V_1/V_3	V_2	**p-value** V_1	V_2
2	30	30	H_0	H_0
2	13	19	0.08	0.96
2	7	9	0.29	0.97
2	6	8	0.23	0.89
3	75	75	H_0	H_0
3	35	56	0.98	0.78
3	20	13	0.92	0.86
3	10	9	0.99	0.59
4	150	150	H_0	H_0
4	54	104	0.94	0.47
4	22	24	0.96	0.75
4	12	8	1.00	0.73
5	200	200	H_0	H_0
5	53	120	0.98	0.31
5	31	26	1.00	0.34
5	15	7	1.00	0.55

TABLE 65 – p-VALUES OF HYPOTHESIS III: INFLUENCE OF COMPLEXITY

Level of complexity	**Time limit [s]** $\boldsymbol{V_1/V_3}$	$\boldsymbol{V_2}$	**p-value** $\boldsymbol{V_1}$	$\boldsymbol{V_2}$
2	30	30	H_0	H_0

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2
3	75	75	0.00	0.02
4	150	150	0.00	0.10
5	200	200	0.00	0.48
2	13	19	H_0	H_0
3	35	56	0.04	0.01
4	54	104	0.00	0.03
5	53	120	0.00	0.07
2	7	9	H_0	H_0
3	20	13	0.00	0.01
4	22	24	0.00	0.09
5	31	26	0.00	0.08
2	6	8	H_0	H_0
3	10	9	0.01	0.01
4	12	8	0.00	0.14
5	15	7	0.00	0.31

TABLE 66 – p-VALUES OF HYPOTHESIS IV: INFLUENCE OF TIME LIMITATION

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	1.00	0.42	0.99
2	7	9	1.00	0.80	1.00
2	6	8	1.00	0.99	1.00
3	75	75	H_0	H_0	H_0
3	35	56	0.15	0.23	0.52
3	20	13	0.25	1.00	0.84
3	10	9	0.88	0.27	0.62
4	150	150	H_0	H_0	H_0
4	54	104	0.43	0.25	0.22
4	22	24	0.88	0.64	0.42
4	12	8	0.99	0.77	0.53

Level of complexity	**Time limit [s]** V_1/V_3	V_2	***p*-value** V_1	V_2	V_3
5	200	200	H_0	H_0	H_0
5	53	120	0.59	0.52	0.22
5	31	26	0.09	0.71	0.07
5	15	7	0.17	0.64	0.08

TABLE 67 – p-VALUES OF HYPOTHESIS IV: INFLUENCE OF COMPLEXITY

Level of complexity	**Time limit [s]** V_1/V_3	V_2	***p*-value** V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
3	75	75	1.00	0.56	1.00
4	150	150	0.12	0.77	1.00
5	200	200	1.00	0.34	0.82
2	13	19	H_0	H_0	H_0
3	35	56	0.91	0.29	0.93
4	54	104	0.00	0.69	0.35
5	53	120	0.99	0.43	0.03
2	7	9	H_0	H_0	H_0
3	20	13	0.18	0.99	0.85
4	22	24	0.00	0.55	0.27
5	31	26	0.11	0.25	0.00
2	6	8	H_0	H_0	H_0
3	10	9	0.93	0.01	0.63
4	12	8	0.06	0.19	0.33
5	15	7	0.57	0.02	0.00

TABLE 68 – p-VALUES OF HYPOTHESIS V: INFLUENCE OF TIME LIMITATION

Level of complexity	**Time limit [s]** V_1/V_3	V_2	***p*-value** V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	0.81	0.91	0.99

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	7	9	0.99	1.00	0.46
2	6	8	0.87	0.97	0.67
3	75	75	H_0	H_0	H_0
3	35	56	0.51	0.96	0.50
3	20	13	0.72	0.53	0.00
3	10	9	0.12	0.87	0.67
4	150	150	H_0	H_0	H_0
4	54	104	0.41	0.13	0.81
4	22	24	0.14	0.81	0.82
4	12	8	0.40	0.61	0.79
5	200	200	H_0	H_0	H_0
5	53	120	0.13	0.96	0.45
5	31	26	0.07	0.29	0.35
5	15	7	0.04	1.00	0.39

TABLE 69 – p-VALUES OF HYPOTHESIS V: INFLUENCE OF COMPLEXITY

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
3	75	75	0.87	0.48	0.82
4	150	150	0.59	0.88	0.36
5	200	200	0.97	0.71	0.76
2	13	19	H_0	H_0	H_0
3	35	56	0.95	0.84	0.10
4	54	104	0.21	0.33	0.09
5	53	120	0.82	0.35	0.17
2	7	9	H_0	H_0	H_0
3	20	13	0.65	0.05	0.73
4	22	24	0.00	0.13	0.86
5	31	26	0.20	0.02	0.63
2	6	8	H_0	H_0	H_0

Level of complexity	Time limit [s]		p-value		
	V_1/V_3	V_2	V_1	V_2	V_3
3	10	9	0.46	0.41	0.93
4	12	8	0.16	0.08	0.67
5	15	7	0.74	0.19	0.53

TABLE 70 – p-VALUES OF HYPOTHESIS VI: INFLUENCE OF TIME LIMITATION

Level of complexity	Time limit [s]		p-value		
	V_1/V_3	V_2	V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	0.72	0.54	1.00
2	7	9	1.00	0.98	1.00
2	6	8	0.97	0.98	1.00
3	75	75	H_0	H_0	H_0
3	35	56	1.00	0.87	1.00
3	20	13	1.00	0.89	1.00
3	10	9	1.00	0.97	1.00
4	150	150	H_0	H_0	H_0
4	54	104	1.00	0.07	0.99
4	22	24	1.00	0.80	1.00
4	12	8	1.00	0.98	1.00
5	200	200	H_0	H_0	H_0
5	53	120	1.00	0.83	1.00
5	31	26	1.00	0.86	1.00
5	15	7	1.00	0.99	1.00

TABLE 71 – p-VALUES OF HYPOTHESIS VI: INFLUENCE OF COMPLEXITY

Level of complexity	Time limit [s]		p-value		
	V_1/V_3	V_2	V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
3	75	75	0.00	0.04	0.01
4	150	150	0.00	0.32	0.00

Level of complexity	Time limit [s]		p-value		
	V_1/V_3	V_2	V_1	V_2	V_3
5	200	200	0.00	0.12	0.00
2	13	19	H_0	H_0	H_0
3	35	56	0.01	0.28	0.00
4	54	104	0.00	0.02	0.00
5	53	120	0.00	0.53	0.00
2	7	9	H_0	H_0	H_0
3	20	13	0.00	0.02	0.00
4	22	24	0.00	0.12	0.00
5	31	26	0.00	0.04	0.00
2	6	8	H_0	H_0	H_0
3	10	9	0.02	0.09	0.00
4	12	8	0.00	0.64	0.00
5	15	7	0.00	0.78	0.00

TABLE 72 – p-VALUES OF HYPOTHESIS VII: INFLUENCE OF TIME LIMITATION

Level of complexity	Time limit [s]		p-value		
	V_1/V_3	V_2	V_1	V_2	V_3
2	30	30	0.92	1.00	0.30
2	13	19	0.08	0.74	0.07
2	7	9	0.10	0.95	0.24
2	6	8	0.10	0.41	0.26
3	75	75	0.42	0.61	0.05
3	35	56	0.09	0.46	0.08
3	20	13	0.10	0.17	0.08
3	10	9	0.31	0.16	0.08
4	150	150	0.74	0.99	0.41
4	54	104	0.09	0.50	0.08
4	22	24	0.10	0.94	0.56
4	12	8	0.10	0.18	0.09
5	200	200	0.26	0.95	0.43
5	53	120	0.09	0.19	0.24

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
5	31	26	0.10	0.64	0.27
5	15	7	0.10	0.17	0.61

TABLE 73 – p-VALUES OF HYPOTHESIS VIII: INFLUENCE OF TIME LIMITATION

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	0.15	0.60	0.99
2	7	9	0.83	0.99	1.00
2	6	8	0.34	0.97	1.00
3	75	75	H_0	H_0	H_0
3	35	56	1.00	0.66	1.00
3	20	13	1.00	0.91	1.00
3	10	9	1.00	0.95	1.00
4	150	150	H_0	H_0	H_0
4	54	104	1.00	0.16	0.96
4	22	24	1.00	0.70	1.00
4	12	8	1.00	0.94	1.00
5	200	200	H_0	H_0	H_0
5	53	120	1.00	0.80	1.00
5	31	26	1.00	0.79	1.00
5	15	7	1.00	0.98	1.00

TABLE 74 – p-VALUES OF HYPOTHESIS VIII: INFLUENCE OF COMPLEXITY

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
3	75	75	0.00	0.09	0.11

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
4	150	150	0.00	0.31	0.00
5	200	200	0.00	0.10	0.01
2	13	19	H_0	H_0	H_0
3	35	56	0.18	0.13	0.57
4	54	104	0.00	0.04	0.01
5	53	120	0.00	0.42	0.44
2	7	9	H_0	H_0	H_0
3	20	13	0.00	0.00	0.03
4	22	24	0.00	0.04	0.00
5	31	26	0.00	0.02	0.05
2	6	8	H_0	H_0	H_0
3	10	9	0.56	0.15	0.23
4	12	8	0.11	0.54	0.55
5	15	7	0.44	0.87	0.48

TABLE 75 – p-VALUES OF HYPOTHESIS IX: INFLUENCE OF TIME LIMITATION

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	0.44	n.a.[187]	
2	7	9	0.36		
2	6	8	0.32		
3	75	75	H_0	H_0	H_0
3	35	56	0.06		

[187] No random decision examined in case of empty cells for neither of the two versions V2 and V3.

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
3	20	13	0.17	0.12	
3	10	9	0.01		
4	150	150	H_0	H_0	H_0
4	54	104	0.09	0.76	
4	22	24	0.16	0.44	
4	12	8	0.00	0.23	
5	200	200	H_0	H_0	H_0
5	53	120	0.00	0.43	
5	31	26	0.00	0.82	
5	15	7	0.00	0.00	

TABLE 76 – p-VALUES OF HYPOTHESIS IX: INFLUENCE OF COMPLEXITY

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
3	75	75	0.57	n.a.[188]	|
4	150	150	0.37	0.02	n.a.
5	200	200	0.48	0.00	|
2	13	19	H_0	H_0	H_0
3	35	56	0.09	n.a.	|
4	54	104	0.05	0.08	|
5	53	120	0.00	0.00	|
2	7	9	H_0	H_0	H_0
3	20	13	0.30	0.11	|
4	22	24	0.14	0.02	|
5	31	26	0.00	0.03	|
2	6	8	H_0	H_0	H_0

[188] See footnote 187.

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
3	10	9	0.05	n.a.	|
4	12	8	0.00	0.01	|
5	15	7	0.00	0.00	|

TABLE 77 – RANDOMNESS OF EQUILIBRIUM CHOICE

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	30	30	1.00	0.69	0.65
2	13	19	0.34	0.93	0.95
2	7	9	0.17	0.87	0.99
2	6	8	0.06	0.50	1.00
3	75	75	|	0.77	|
3	35	56	n.a.[189]	0.50	n.a.[190]
3	20	13	|	0.81	|
3	10	9	|	0.87	|
4	150	150	0.00	0.99	0.99
4	54	104	0.00	1.00	0.75
4	22	24	0.00	0.97	0.96
4	12	8	0.01	0.89	0.85
5	200	200	1.00	0.92	1.00
5	53	120	0.99	0.68	1.00
5	31	26	0.99	1.00	1.00
5	15	7	0.95	0.89	1.00

[189] In Versions V1 and V3 at this level of complexity the Chicken Game is played (see Appendix B). Here, all alternatives are potential choices of heuristic Equilibrium. Thus, in this round any decision is labeled as equilibrium choice. With this, neither a discrimination based on choice nor a hypothesis-proportion test is possible here, since the equation of the corresponding test statistic is invalid.
[190] See footnote 189.

TABLE 78 – p-VALUES OF HYPOTHESIS X: INFLUENCE OF TIME LIMITATION

Level of complexity	Time limit [s]		p-value		
	V_1/V_3	V_2	V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	0.00	0.83	0.29
2	7	9	0.00	0.31	0.99
2	6	8	0.00	0.73	0.09
3	75	75	H_0	H_0	H_0
3	35	56	0.30	0.34	0.51
3	20	13	0.72	0.79	0.65
3	10	9	0.84	0.77	0.84
4	150	150	H_0	H_0	H_0
4	54	104	0.64	0.71	0.83
4	22	24	0.72	0.31	0.62
4	12	8	0.94	0.19	0.85
5	200	200	H_0	H_0	H_0
5	53	120	0.16	0.23	0.56
5	31	26	0.16	0.96	0.16
5	15	7	0.15	0.52	0.37

TABLE 79 – p-VALUES OF HYPOTHESIS X: INFLUENCE OF COMPLEXITY

Level of complexity	Time limit [s]		p-value		
	$\boldsymbol{V_1/V_3}$	$\boldsymbol{V_2}$	$\boldsymbol{V_1}$	$\boldsymbol{V_2}$	$\boldsymbol{V_3}$
2	30	30	H_0	H_0	H_0
3	75	75	0.00	0.85	0.00
4	150	150	0.00	1.00	0.79
5	200	200	1.00	1.00	0.81
2	13	19	H_0	H_0	H_0
3	35	56	0.00	0.38	0.00
4	54	104	0.39	1.00	0.99
5	53	120	1.00	0.88	0.93
2	7	9	H_0	H_0	H_0
3	20	13	0.00	0.99	0.00

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
4	22	24	0.32	1.00	0.13
5	31	26	1.00	1.00	0.01
2	6	8	H_0	H_0	H_0
3	10	9	0.00	0.88	0.00
4	12	8	0.94	0.96	1.00
5	15	7	1.00	0.99	0.95

TABLE 80 – p-VALUES OF HYPOTHESIS XI: INFLUENCE OF TIME LIMITATION

Level of complexity	Time limit [s] V_1/V_3	V_2	p-value V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	0.99	0.97	0.72
2	7	9	1.00	0.97	0.74
2	6	8	1.00	0.96	0.36
3	75	75	H_0	H_0	H_0
3	35	56	0.09	0.65	0.41
3	20	13	0.54	0.89	0.48
3	10	9	0.27	0.87	0.07
4	150	150	H_0	H_0	H_0
4	54	104	0.67	0.87	0.62
4	22	24	0.75	0.75	0.26
4	12	8	0.65	0.77	0.08
5	200	200	H_0	H_0	H_0
5	53	120	0.76	0.32	0.61
5	31	26	0.94	0.56	0.50
5	15	7	0.87	0.23	0.19

TABLE 81 – p-VALUES OF HYPOTHESIS XI: INFLUENCE OF COMPLEXITY

Level of complexity	**Time limit [s]** V_1/V_3	V_2	**p-value** V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
3	75	75	1.00	0.85	0.98
4	150	150	1.00	0.92	0.98
5	200	200	1.00	0.87	0.99
2	13	19	H_0	H_0	H_0
3	35	56	1.00	0.47	0.98
4	54	104	1.00	0.99	0.99
5	53	120	1.00	0.10	1.00
2	7	9	H_0	H_0	H_0
3	20	13	1.00	0.91	0.99
4	22	24	1.00	0.77	0.95
5	31	26	1.00	0.31	1.00
2	6	8	H_0	H_0	H_0
3	10	9	0.99	0.91	0.93
4	12	8	0.98	0.82	0.90
5	15	7	0.99	0.06	0.98

TABLE 82 – MOUSE ACTIVITY

Level of complexity	**Time limit [s]** $\boldsymbol{V_1/V_3}$	$\boldsymbol{V_2}$	**p-value** $\boldsymbol{V_1}$	$\boldsymbol{V_2}$	$\boldsymbol{V_3}$
2	30	30	H_0	H_0	H_0
2	13	19	0.39	0.47	0.13
2	7	9	0.02	0.22	0.12
2	6	8	0.00	0.19	0.15
3	75	75	H_0	H_0	H_0
3	35	56	0.27	0.22	0.65
3	20	13	0.11	0.24	0.26
3	10	9	0.02	0.63	0.49
4	150	150	H_0	H_0	H_0
4	54	104	0.10	0.60	0.25

Level of complexity	Time limit [s]		p-value		
	V_1/V_3	V_2	V_1	V_2	V_3
4	22	24	0.03	0.70	0.28
4	12	8	0.00	0.56	0.66
5	200	200	H_0	H_0	H_0
5	53	120	0.00	0.93	0.00
5	31	26	0.00	0.77	0.00
5	15	7	0.29	0.49	0.01

TABLE 83 – p-VALUES OF HYPOTHESIS XII: INFLUENCE OF TIME LIMITATION

Level of complexity	Time limit [s]		p-value		
	V_1/V_3	V_2	V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	0.91	0.59	0.85
2	7	9	0.99	0.30	0.95
2	6	8	1.00	0.94	1.00
3	75	75	H_0	H_0	H_0
3	35	56	0.63	0.43	0.32
3	20	13	0.65	0.20	0.56
3	10	9	0.95	0.57	0.94
4	150	150	H_0	H_0	H_0
4	54	104	0.80	0.89	0.51
4	22	24	1.00	0.80	0.33
4	12	8	1.00	0.45	0.79
5	200	200	H_0	H_0	H_0
5	53	120	0.99	0.50	0.25
5	31	26	0.98	0.43	0.25
5	15	7	1.00	0.36	0.44

TABLE 84 – p-VALUES OF HYPOTHESIS XII: INFLUENCE OF COMPLEXITY

Level of complexity	**Time limit [s]** V_1/V_3	V_2	***p*-value** V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
3	75	75	0.12	0.91	0.10
4	150	150	0.00	0.89	0.26
5	200	200	0.00	1.00	0.18
2	13	19	H_0	H_0	H_0
3	35	56	0.02	0.78	0.00
4	54	104	0.00	0.98	0.08
5	53	120	0.00	0.99	0.00
2	7	9	H_0	H_0	H_0
3	20	13	0.00	0.73	0.01
4	22	24	0.00	0.99	0.01
5	31	26	0.00	0.99	0.00
2	6	8	H_0	H_0	H_0
3	10	9	0.00	0.43	0.00
4	12	8	0.00	0.33	0.01
5	15	7	0.00	0.66	0.00

TABLE 85 – p-VALUES OF HYPOTHESIS XIII

Level of complexity	**Time limit [s]** $\boldsymbol{V_1/V_3}$	$\boldsymbol{V_2}$	***p*-value** $\boldsymbol{V_1}$	$\boldsymbol{V_2}$	$\boldsymbol{V_3}$
2	30	30	0.45	0.75	0.75
2	13	19	0.87	0.44	0.24
2	7	9	0.85	0.71	0.37
2	6	8	0.59	0.84	0.48
3	75	75	1.00	0.88	0.56
3	35	56	0.25	0.33	0.15
3	20	13	0.75	0.92	0.41
3	10	9	0.60	0.50	0.24
4	150	150	0.09	0.41	0.02
4	54	104	0.17	0.06	0.41

Level of complexity	Time limit [s]		p-value		
	V_1/V_3	V_2	V_1	V_2	V_3
4	22	24	0.14	0.65	0.05
4	12	8	0.93	0.18	0.81
5	200	200	n.a.[191]	0.23	0.38
5	53	120	0.63	0.47	0.04
5	31	26	0.19	n.a.[191]	0.12
5	15	7	0.09	0.10	0.27

TABLE 86 – p-VALUES OF HYPOTHESIS XIV: INFLUENCE OF TIME LIMITATION

Level of complexity	Time limit [s]		p-value		
	V_1/V_3	V_2	V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
2	13	19	0.11	0.35	0.10
2	7	9	0.08	0.76	0.37
2	6	8	0.05	0.91	0.41
3	75	75	H_0	H_0	H_0
3	35	56	0.88	0.24	0.05
3	20	13	0.76	0.53	0.21
3	10	9	0.94	0.76	0.18
4	150	150	H_0	H_0	H_0
4	54	104	0.42	0.00	0.41
4	22	24	0.17	0.21	0.61
4	12	8	0.20	0.11	0.47
5	200	200	H_0	H_0	H_0
5	53	120	0.58	0.57	0.99
5	31	26	0.48	0.63	0.33
5	15	7	0.71	0.76	0.95

[191] Testing is not possible due to data-set-size of zero for comparative values.

TABLE 87 – p-VALUES OF HYPOTHESIS XIV: INFLUENCE OF COMPLEXITY

Level of complexity	**Time limit [s]**		**p-value**		
	V_1/V_3	V_2	V_1	V_2	V_3
2	30	30	H_0	H_0	H_0
3	75	75	0.00	0.16	0.07
4	150	150	0.09	0.72	0.01
5	200	200	0.00	0.11	0.02
2	13	19	H_0	H_0	H_0
3	35	56	0.38	0.10	0.08
4	54	104	0.42	0.02	0.19
5	53	120	0.09	0.29	0.85
2	7	9	H_0	H_0	H_0
3	20	13	0.33	0.07	0.04
4	22	24	0.32	0.19	0.07
5	31	26	0.12	0.10	0.03
2	6	8	H_0	H_0	H_0
3	10	9	0.70	0.08	0.04
4	12	8	0.42	0.03	0.05
5	15	7	0.33	0.07	0.55

TABLE 88 – p-VALUES OF HYPOTHESIS XV: INFLUENCE OF TIME LIMITATION

Level of complexity	**Time limit [s]**	***p*-value**
2	30	H_0
2	13	0.02
2	7	0.18
2	6	0.12
3	75	H_0
3	35	0.97
3	20	0.97
3	10	1.00
4	150	H_0
4	54	0.68
4	22	0.96
4	12	1.00
5	200	H_0
5	53	1.00
5	31	1.00
5	15	1.00

TABLE 89 – p-VALUES OF HYPOTHESIS XV: INFLUENCE OF COMPLEXITY

Level of complexity	**Time limit [s]**	***p*-value**
2	30	H_0
3	75	0.02
4	150	0.68
5	200	0.68
2	13	H_0
3	35	0.96
4	54	1.00
5	53	1.00
2	7	H_0
3	20	0.79
4	22	1.00
5	31	1.00
2	6	H_0
3	10	1.00
4	12	1.00
5	15	1.00