TU Dortmund University

Faculty of Education, Psychology and Sociology

The influence of mental representations on eye movement patterns under uncertainty

Doctoral Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy (Dr. phil.)

by

Johanna Renker

born in Herne

Reviewer:

PD Dr. phil. Gerhard Rinkenauer

Prof. Dr. Josef F. Krems

December 2017

Dortmund, Germany

Acknowledgments

At first, I would like to thank Gerhard Rinkenauer for his great support and engagement all the time. He inspired me as a supervisor as well as a person, and his constantly friendly manner always calmed me down. I would also like to thank my Canadian supervisor Alan Kingstone for his warm welcome at UBC and helping me to have a great experience abroad, which had a formative influence on my personal development. Another thank goes to Josef Krems for being my second assessor. My thesis would not have been successful without the technical support by Hanno and Aswin. Both helped me with programming the experiments, data preparation and analysis. Many special thanks go to Hanno for his continuous support and openness to new ideas to deal with all the eye movement data. More valuable support was done by Martin, Tanja and Carola for the data collection. I would like to express special thanks to Martin for his thorough and conscientious work. I also would like to thank all the participants taking part in the experiments who tried their best even if the experiments became annoying after a while. Special thanks also go to all the colleagues who made me feel comfortable at the institute and who contributed to a warm atmosphere. In particular, I would like to mention Oliver and Tobias who provided me helpful feedback on the thesis and Christine, who always had an open door to listen to me and became a friend for life. Finally, Magali was the best new colleague I could imagine and she gave me an incredible support to finish the thesis.

Outside of the institute, another special thank goes to Nele Rußwinkel for being my mentor, spending time with me to discuss all my issues of the scientific world and finally, integrating me into her working group by offering me time for the presentation of my work. I would also like to thank Agnes Scholz for showing me the work at the TU Chemnitz and giving me useful tips for my thesis. At a personal level, I would like to thank for the support of my family. A very special thank goes to my brother for providing me helpful feedback with regard to the English language and my parents for trying to understand what I'm doing and always being there for me no matter what the situation. They always felt solidarity with me. Further, I would like to thank my friends from school, university, improvisation theatre, canoeing, orienteering and Swedish Tandem class. They all helped me to relax after work, gave me inspiration and motivated me to keep going. A warm thank goes to Knut as well as to Nicolas who always stand behind me and take care for my mental well-being.

The Research Training Group 1855 "Discrete Optimization of technical systems under uncertainty" funded this thesis and would not have been possible without this funding. Furthermore, it was a pleasure to get to know all the people from the Research Training Group and to work interdisciplinarily even if it was not always a simple matter. Special thanks go to Stephan for initiating our cooperation project and being an open-minded and valuable communication partner. Last but not least, I would like to thank the "Leistungszentrum für Logistik und IT" who supported the final phase of my work.

Summary

This thesis investigates the acquisition of mental representations under uncertainty. Five experiments were conducted. The aim of the experiments was to assess if visual search behavior reflects the accuracy of the mental representations and the degree of uncertainty during the acquisition process. In order to induce uncertainty, a visual spatial search task was developed with likely and unlikely target locations. Each of three targets was associated to one of the three target locations with a higher probability and with a lower probability to the other locations. The probability distribution remained constant for all experiments except for Experiment 5. Participants' task was twofold: First, they had to predict the location of target object presentation in a *prediction task*. Second, they had to respond to color changes of the target object during the presentation of the target object in a subsequent *reaction task*. It is expected that participants develop a mental representation about the likelihoods of target locations that allows them to predict the location of the target objects more accurately. During the development of the mental representation, the degree of the participants' uncertainty should be indicated by eye movements which reflect searching for relevant information and inhibiting irrelevant or ambiguous information.

In Experiment 1, a new experimental paradigm was introduced to manipulate the probability of target locations. Participants were instructed to predict the locations of three different targets and to respond when they appeared on the screen. Prediction time, reaction time and the time course of eye movement patterns (e.g., fixation frequency, fixation duration, gaze shifts) were analyzed. Eye movement patterns changed over the time course of the task with increasing learning and reduced task uncertainty. The extensive visual search at the beginning of the experimental session became more focused towards the end of the session. Further, behavioral data suggests that participants developed a "Take The Best" decision strategy, i.e. lower probabilities were ignored and the most probable option was chosen. The new experimental design appeared to be useful for the purpose of the study and was therefore used in further experiments.

The aim of Experiment 2 was to assess to what extent search difficulty affects the development of the mental representation. Thus, targets were presented at an unstructured white-gray patterned background in order to degrade the target stimuli. The fixation frequency in Experiment 2 was significantly higher compared to the first experiment. Contrary to the former assumptions the developed mental representation was equally accurate.

Experiment 3 was designed as a relearning experiment to investigate if eye movement patterns indicate changes of probability concepts during the development of a mental representation. First,

participants had to learn a probability concept. Then, they had to relearn another concept which required the adaptation of the initial mental representation. The results of Experiment 3 suggest that eye movement patterns indicate different phases during relearning. The beginning of the relearning phase was immediately signalized by an increase of fixation frequency whereas the increase in fixation duration was temporally delayed. However, performance differences between learning and relearning were smaller than expected.

In Experiment 4, a *prediction* and a *reaction task* were assessed separately. The aim of this experiment was to elucidate which task is dominating the development of the mental representation. Participants developed an almost accurate mental representation of the probability concept in the *prediction task*, but not in the *reaction task*. Thus, the mental representation acquired during the performance of the visual spatial search task seems to rely mainly on the prediction task. Interestingly, this difference in the accuracy of the mental representation was only indicated by gaze shift activity.

In Experiment 5, two probability distributions were employed to manipulate the degree of uncertainty. Participants had to learn a higher and a lower probability distribution of the object-exit associations. However, eye movement patterns did not differ between the performances of both probability distributions. Further, behavioral data suggests that the “Take The Best” decision strategy was employed in both distribution conditions.

The results of this thesis demonstrate that eye movements give insights into the development of mental representations under uncertainty and thus, inform about the state of the participant. The results suggest that eye movement patterns reflect the learning state as well as the subjective uncertainty of the participant, viz. the usage of decision strategies and strategies to cope with uncertainty. Further, it was demonstrated that eye movements dependent on the design of the task (Experiment 1-2) and the initial knowledge (Experiment 3). However, processing depth was not reflected by eye movement parameters (Experiment 4). In addition, the manipulation of the objective uncertainty by varying the probability distribution did not seem to affect the degree of subjective uncertainty as intended (Experiment 5).

Zusammenfassung

Diese Dissertation beschäftigt sich mit der Entwicklung mentaler Repräsentationen unter Unsicherheit. Fünf Experimente wurden dafür konzipiert und durchgeführt. Ziel der Experimente war es, zu untersuchen, ob visuelles Suchverhalten die Genauigkeit der mentalen Repräsentation sowie den Grad der Unsicherheit während des Entwicklungsprozesses wiederspiegelt. Um Unsicherheit zu induzieren wurde eine visuell-räumliche Suchaufgabe entwickelt, mit wahrscheinlichen und unwahrscheinlichen Zielpositionen. Jedes der drei Zielobjekte erschien an einer von drei Zielpositionen mit einer höheren Wahrscheinlichkeit und an den anderen mit einer niedrigeren Wahrscheinlichkeit. Die Wahrscheinlichkeitsverteilung blieb über alle Experimente hinweg konstant mit Ausnahme von Experiment 5. Die Aufgabe der Versuchspersonen war zweigeteilt: Zunächst mussten sie vorhersagen an welcher Position die Zielobjekte erscheinen (*Vorhersageaufgabe*). Nachfolgend mussten sie während des Auftretens des Zielobjekts auf Farbveränderungen reagieren (*Reaktionsaufgabe*). Es wurde erwartet, dass Versuchspersonen nach und nach eine mentale Repräsentation über die Wahrscheinlichkeiten der Zielpositionen entwickeln, die es ermöglicht das Auftreten der Zielobjekte genauer vorherzusagen. Dabei sollte das Ausmaß an Unsicherheit, das von den Versuchspersonen empfunden wurde, anhand von Augenbewegungsmustern sichtbar werden, da diese die Suche nach relevanten Informationen und die Inhibition von irrelevanten oder mehrdeutigen Informationen wiederspiegeln.

In Experiment 1 wurde ein neues Paradigma eingeführt, um die Wahrscheinlichkeit der Zielpositionen zu manipulieren. Die Versuchspersonen wurden instruiert die Positionen von drei verschiedenen Zielobjekten vorherzusagen und gegebenenfalls zu reagieren, wenn diese auftreten. Die Vorhersagezeit, die Reaktionszeit sowie der zeitliche Verlauf der Augenbewegungsparameter (u.a. Fixationshäufigkeit, Fixationsdauer, Blickwechsel) wurden analysiert. Augenbewegungsmuster veränderten sich über den zeitlichen Verlauf der Aufgabe mit gesteigertem Wissen und reduzierter Aufgabenunsicherheit. Das extensive visuelle Suchverhalten wurde zum Ende des Experiments fokussierter. Zusätzlich deuten Verhaltensdaten darauf hin, dass Versuchspersonen eine „Take The Best“ Entscheidungsstrategie entwickeln, d.h. dass niedrige Wahrscheinlichkeiten ignoriert werden und eher die wahrscheinlichere Option gewählt wird. Das neue Paradigma schien für den Zweck der Studie angemessen zu sein und wurde somit auch für die weiteren Experimente benutzt.

Das Ziel von Experiment 2 war es, zu untersuchen inwieweit Schwierigkeiten bei der visuellen Suche die Entwicklung der mentalen Repräsentation beeinträchtigen. Dazu wurden die Zielobjekte auf einem gemusterten Hintergrund präsentiert und waren somit schwer erkenntlich. Die Anzahl der Fixationen war in Experiment 2 signifikant hört als in Experiment 1. Jedoch wurde konträr zu der anfänglichen Annahme eine mentale Repräsentation entwickelt, die ähnlich akkurat war wie in Experiment 1.

Experiment 3 wurde als Umlernexperiment konzipiert, um herauszufinden inwieweit Augenbewegungen auf Veränderungen von Wahrscheinlichkeitskonzepten bei der Entwicklung einer mentalen Repräsentation hindeuten. Zunächst mussten Versuchspersonen ein Wahrscheinlichkeitskonzept erlernen. Danach musste ein bestehendes Konzept umgelernt und die vorherige mentale Repräsentation angepasst werden. Die Ergebnisse des Experiments zeigten, dass Augenbewegungsmuster auf verschiedene Phasen während des Umlernprozess hinweisen. Der Beginn der Umlernphase wurde durch einen sofortigen Anstieg der Fixationshäufigkeit gekennzeichnet, wohingegen der Anstieg der Fixationsdauer zeitverzögert eintrat.

In Experiment 4 wurden die *Vorhersageaufgabe* und die *Reaktionsaufgabe* getrennt betrachtet. Ziel des Experiments war es, aufzuklären welche der Aufgaben hauptsächlich den Aufbau der mentalen Repräsentation beeinflusst. Versuchspersonen entwickelten eine nahezu akkurate mentale Repräsentation des Wahrscheinlichkeitskonzepts in der *Vorhersageaufgabe*, nicht aber in der *Reaktionsaufgabe*. Daher scheint die mentale Repräsentation, die während der Durchführung der visuell-räumlichen Aufgabe entsteht, hauptsächlich auf der *Vorhersageaufgabe* zu basieren. Interessanterweise spiegelten Blickwechsel diesen Unterschied der Genauigkeit der mentalen Repräsentation wieder.

In Experiment 5 wurden zwei Wahrscheinlichkeitsverteilungen verwendet, um den Grad der Unsicherheit zu manipulieren. Versuchspersonen mussten eine höhere und eine niedrigere Wahrscheinlichkeitsverteilung lernen. Augenbewegungsmuster zeigten jedoch keine Unterschiede zwischen dem Lernen der beiden Wahrscheinlichkeitsverteilungen. Weiterhin deuten die Verhaltensdaten darauf hin, dass die „Take The Best“ Entscheidungsstrategie bei beiden Bedingungen benutzt wurde und es somit auch bei der Wahl der Strategie keine Unterschiede gab.

Die Ergebnisse der Arbeit verdeutlichen, dass Augenbewegungen Einblicke in den Entwicklungsprozess von mentalen Repräsentationen unter Unsicherheit geben und somit über den Zustand der Versuchsperson informieren. Augenbewegungsmuster spiegelten den Lernstand sowie die subjektive Unsicherheit der Versuchsperson wieder, d.h. die Nutzung von Entscheidungsstrategien und die Nutzung von Strategien, um mit Unsicherheit umzugehen. Darüber hinaus zeigte sich, dass Augenbewegungen sowohl abhängig vom Aufgabendesign (Experiment 1-2) als auch vom Vorwissen sind (Experiment 3). Jedoch können Augenbewegungsmuster die Tiefe der kognitiven Verarbeitung nicht widerspiegeln (Experiment 4). Zudem wirkte sich die Manipulation der objektiven Unsicherheit durch die Variierung der Wahrscheinlichkeitsverteilung nicht auf die subjektive Unsicherheit aus (Experiment 5).

Table of Contents

1 Introduction 1

1.1 Mental Model Theory and Probabilistic Mental Models (PMM) 4

1.2 Decision Making under Uncertainty 8

1.3 Eye Movements, Attention and Learning 12

1.4 Further Evidence: Eye Movements and Mental Representations 15

1.5 Overview and Research Objectives 17

1.5.1 Experiment I: Predicting the Appearance of Geometric Objects 19

1.5.2 Experiment II: Degradation of Search Targets 20

1.5.3 Experiment III: Dynamic Relearning 20

1.5.3 Experiment IV: Separate Tasks 20

1.5.4 Experiment V: Learning Different Probabilities 20

2 General Method 22

2.1 Participants 22

2.2 Occluded Visual Spatial Search Task (OVSST) 22

2.3 Pre- and post-tests 24

2.4 Design and Procedure 25

2.5 Eye Tracking 26

2.7 Data Analysis 27

2.7.1 Data Preparation 27

2.7.2 Independent and Dependent Variables 28

2.7.3 Recurrence Quantification Analysis 29

2.7.4 Statistical Methods 31

3 Experiment I – Predicting the Appearance of Geometric Objects 32

3.1 Introduction 32

3.1.1 Research Question and Hypotheses 34

3.2 Method 35

3.2.1 Participants 35

3.2.2 Procedure 35

3.2.3 Data Analysis 36

3.3 Results 38

3.4 Discussion 48

3.5 Conclusion 53

4 Experiment II – Degradation of Search Targets 54
4.1 Introduction.. 54
4.1.1 Research Question and Hypotheses... 55
4.2 Method.. 56
4.2.1 Participants.. 56
4.2.2 Procedure 56
4.2.3 Data Analysis 57
4.3 Results 58
4.4 Discussion... 62
4.5 Conclusion 66
5 Experiment III – Dynamic Relearning 67
5.1 Introduction.. 67
5.1.1 Research Question and Hypotheses... 69
5.2 Method.. 69
5.2.1 Participants.. 69
5.2.2 Procedure 70
5.2.3 Data Analysis 70
5.3 Results 71
5.4 Discussion... 78
5.5 Conclusion 81
6 Experiment IV - Separate Tasks 82
6.1 Introduction.. 82
6.1.1 Research Question and Hypotheses... 84
6.2 Method.. 85
6.2.1 Participants.. 85
6.2.2 Procedure 85
6.2.3 Data Analysis 87
6.3 Results 88
6.3.1 Prediction Task 88
6.3.3 Reaction Task... 91
6.3.2 Comparison of the Prediction and the Reaction Task 91
6.4 Discussion... 94
6.5 Conclusion 98

7 Experiment V – Learning Different Probabilities 99
7.1 Introduction.. 99
7.1.1 Research Question and Hypotheses.. 100
7.2 Pilot Studies.. 101
7.2.1 Participants.. 101
7.2.2 Procedure 101
7.2.3 Data Analysis 102
7.2.4 Results 102
7.2.5 Discussion & Conclusion of the pilot studies... 104
7.3 Method.. 105
7.3.1 Participants.. 105
7.3.2 Procedure 105
7.3.3 Data Analysis 106
7.4 Results 107
7.4.1 High Probability Concept... 107
7.4.2 Low Probability Concept.. 111
7.4.3 Comparison of the Low and High Probability Concept 114
7.4.4 Comparison of Different Stimulus Presentations... 116
7.4.5 Analysis of Control Variables... 117
7.5 Discussion... 117
7.5 Conclusion 121
8 General Discussion...122
8.1 Summary 122
8.2 Practical Relevance... 129
8.3 Limitations.. 130
8.4 Directions for Future Research... 132
8.5 Conclusion 134
Reference List .... Fehler! Textmarke nicht definiert.
Appendix A: Concept Awareness Questionnaire i
Appendix B: Procedure and Results of the Eye Tracking Validation Study ii
Appendix C: Means and Standard Deviations for Trends across Blocks..............................iv

List of Figures

Figure 1.1:	Simplified depiction of the OVSST	3
Figure 1.2:	Flow diagram for the 5 steps of the Take The Best heuristic	7
Figure 1.3:	Task uncertainty consisting of subjective and objective uncertainty	8
Figure 1.4:	Framework for the 5 processes during decision making	10
Figure 1.5:	The three-stage models in comparison	18
Figure 2.1:	Schematic description of one trial of the OVSST	24
Figure 2.2:	Experimental setup	26
Figure 2.3:	The four AOIs of the display	29
Figure 2.4:	Example of a recurrence plot for the current task (OVSST)	30
Figure 3.1:	Box plot for the number of correct predictions in the 100% condition of Experiment I	36
Figure 3.2:	Trial analysis separated into two intervals	37
Figure 3.3:	Judgment Times in Experiment I	39
Figure 3.4:	Results of the variables in Experiment I	41
Figure 3.5:	Response frequencies in Experiment I	44
Figure 3.6:	Blinks in Experiment I	47
Figure 4.1:	Unstructured background of the screen	57
Figure 4.2:	Box plot for the number of correct predictions in the 100% condition of Experiment II	57
Figure 4.3:	Judgment Times in Experiment II	59
Figure 4.4:	Results of the variables in Experiment II	60
Figure 5.1:	Box plot for the number of correct predictions in the 100% condition of Experiment III	71
Figure 5.2:	Judgment Times in Experiment III	72
Figure 5.3:	Results of the variables in Experiment III	74
Figure 6.1:	The OVSST with fuzzy Gabor figures in Experiment IV	85
Figure 6.2:	Temporal sequence of Experiment IV	86
Figure 6.3:	Judgment Times in Experiment IV	89
Figure 6.4:	Results of the variables in the prediction task of Experiment IV	90
Figure 6.5:	Comparison of eye movement parameters in the prediction and the and the reaction task in Experiment V	93
Figure 7.1:	The OVSST with a new set of Gabor figures in Experiment V	102
Figure 7.2:	Judgment Times in Experiment V in the high probability condition	108
Figure 7.3:	Results of variables of the high-probability condition in Experiment V	110
Figure 7.4:	Judgment Times in Experiment V in the low probability condition	112
Figure 7.5:	Results of variables of the low probability condition in Experiment V	113
Figure 7.6:	Comparison of predictions in the low and high-probability condition in Experiment V	115
Figure 8.1:	An integration of the three-stage models	128

List of Tables

Table 3.1:	Memory representation of the probability concept and behavioral probabilities in Experiment I	43
Table 3.2:	Trial analysis in Experiment I	45
Table 3.3:	Descriptive statistics of the Performance Cluster in Experiment I	46
Table 4.1:	Memory representation of the probability concept and behavioral probabilities in Experiment I	62
Table 5.1:	Probability concept I in Experiment III - Memory representation and behavioral probabilities during learning and relearning the concept	77
Table 5.2:	Probability concept II in Experiment III - Memory representation and behavioral probabilities during learning and relearning the concept	77
Table 6.1:	Statistics of the Performance Cluster and Number of Gaze Shifts in Experiment IV	91
Table 6.2:	Memory representation of the probability concept of the prediction and reaction task in Experiment IV	92
Table 7.1:	Memory representation of the probability concept of two lower probability distributions in Pilot Study I	103
Table 7.2:	Memory representation of the probability concept of two lower probability distributions in Pilot Study II	104
Table 7.3:	AOI classifications and definitions in Experiment V	106
Table 7.4:	Results of eye movements in the AOIs in the high probability condition for correctly in comparison to incorrectly predicted trials of Experiment V	111
Table 7.5:	Results of eye movements in the AOIs in the low probability condition for correctly in comparison to incorrectly predicted trials of Experiment V	114
Table 7.6:	Subjective probability concept in the high and low probability condition in Experiment V	116

1 Introduction

What are the mental mechanisms underlying the interaction between users and technical systems? And how behave users if working with a technical system does not proceed as expected? If users interact with technical systems like a computer program or a ticket machine, they develop a mental representation or also called mental model of functioning of the system. Based on their mental model, users establish expectations concerning the technical system and behave accordingly. The technical system either strengthens these expectations or contradicts them. If it contradicts them, users have to adapt their mental model by building new associations (Preim & Dachselt, 2010, p. 94). The mental model becomes more detailed with increasing expertise. However, as mental models are only a reduced representation of the reality and complex systems cannot provide a complete set of information, errors still occur during the interaction. Thus, users have to handle a certain degree of task uncertainty due to a lack of knowledge concerning the cause-and-effect relationship, especially when expertise with a system is still low (Thompson, 1967). The task uncertainty involves the subjective uncertainty perceived by the participants as well as the objective uncertainty which is provided by the system, for instance, by the degree of complexity. Task uncertainty can be manipulated and be diminished by supporting the development of the mental model, for example, by introducing a help system, so that a more accurate mental model can be established in a shorter time period. Generally, user's system performance improve with more accurate mental models (Donnell, 1996).

The acquisition of accurate mental models can be supported by designing user friendly interfaces because an improved interaction environment reduces task uncertainty (e.g., Bennett & Flach, 2011; Brown, 1999; Galitz, 2007; Johnson, 2014). For instance, Jipp (2016) studied expertise development while participants had to control air traffic. Two different types of automation design were used within this scenario, information automation and decision automation. Information automation refers to the automation of the information acquisition and analysis, for example, by highlighting relevant information. During decision automation, the selection of the decision was additionally automated, for example, the system advises the participants of the ideal speed of the aircraft which is relevant for the task. Results of the study showed that decision automation leads to more accurate mental models than information automation because of an increased need to process information. Thus, the interaction automation also affects the acquisition of the mental model. However, in addition to a user-friendly interface and automation, the user might be also supported more individually during the interaction with the system. For example, determining at which point the users' individual mental model is incomplete and task uncertainty is high, is essential for further smooth processing and the main focus of the thesis.

This thesis seeks to find appropriate mechanisms and indicators for task uncertainty. Lipshitz and Strauss (1997) report different strategies people apply to cope with uncertainty. Humans might reduce uncertainty by searching for additional information or by waiting until additional information is available. Ignoring or distorting ambiguous information, in the way that they fit to the existing information, are other strategies known to suppress uncertainty. All these strategies deal with information search which is connected to attentional processes and visual search behavior. Consequently, visual search behavior and thus eye movement patterns might be a suitable indicator for the degree of task uncertainty. Eye movements are fast, frequent and automatic actions that reflect to which location users pay attention and thereby may allow insights into strategies of information accumulation (Hoffman & Rehder, 2010; Spivey & Dale, 2011). However, not only is information accumulation important for goal-directed behavior but also information processing. Findings of research on reading already provided evidence that stimuli are cognitively processed about the same amount of time the person has fixated them, for example the gaze duration for infrequent words was longer than for frequent (more familiar) words (Just & Carpenter, 1980). On the basis of such findings, Just and Carpenter (1980) proposed the eye-mind hypothesis assuming a strong causal relationship between attention and information processing. There should be “no appreciable lag between what is being fixated and what is being processed” (Just & Carpenter, 1980, p. 331). Transferring this idea to the present thesis, eye movement patterns might inform about an inadequate information processing or misunderstanding during the acquisition of mental models and thus, also about the degree of subjective uncertainty. The fixation on rather relevant than irrelevant information combined with less visual search, for example, may imply more accurate mental models and less subjective uncertainty. There is a considerable amount of studies in basic and applied research investigating eye movements to draw conclusions about visual information processing (Jacob & Karn, 2003, for review). This thesis tries to advance the present research by focusing on eye movements under uncertainty in the context of human-computer interaction (HCI). The purpose of the experimental research is to provide a first foundation for further studies in this context. Liversedge et al. (2011) already mentioned that knowledge about uncertainty and eye movements is limited and thus, stresses the importance for further research. In this thesis, I use input from different research fields, namely decision making, learning, working memory, information processing and eye movement research, to gain deeper insights into the described research topic.

In the following, the theoretical background of the thesis is summarized and a research framework is developed. At the end of Chapter 1 an overview of the relevant research questions and an outlook about the experimental setup is given. In Chapter 2 the experimental paradigm is outlined together with the General Method of the studies. In Chapters 3-7 the motivations of five successively developed experiments are described in detail as well as their methods and results. At the end of each chapter, the findings and implications are discussed. Finally, in Chapter 8 the research findings of all experiments are integrated and discussed in the context of the proposed research model. In addition, limitations of the studies are considered and an outlook for future research is presented.

For the purpose of the study a new experimental paradigm is developed: the Occluded Visual Spatial Search Task (OVSST). The experimental paradigm will be shortly introduced already at this point due to its high relevance for the comprehension of the introduction. The experimental task, however, will be explained in more detail later in the General Method section. Participants have to observe target objects that appear at one of three target locations with a certain probability (Fig. 1.1). They are asked to perform two tasks: predicting the target location of the objects and, subsequently, reacting to changes of the color intensity when the object appears. After the experiment, participants are asked to estimate the probabilities via questionnaire. The OVSST allows to investigate the acquisition of a new mental model of a simplified decision task inhering spatial uncertainty. Further, the conscious as well as unconscious understanding of the underlying probability concept is assessed via subjective as well as objective measurements. Furthermore, the OVSST allows to vary the degree of objective uncertainty by varying the probabilities of the target locations.

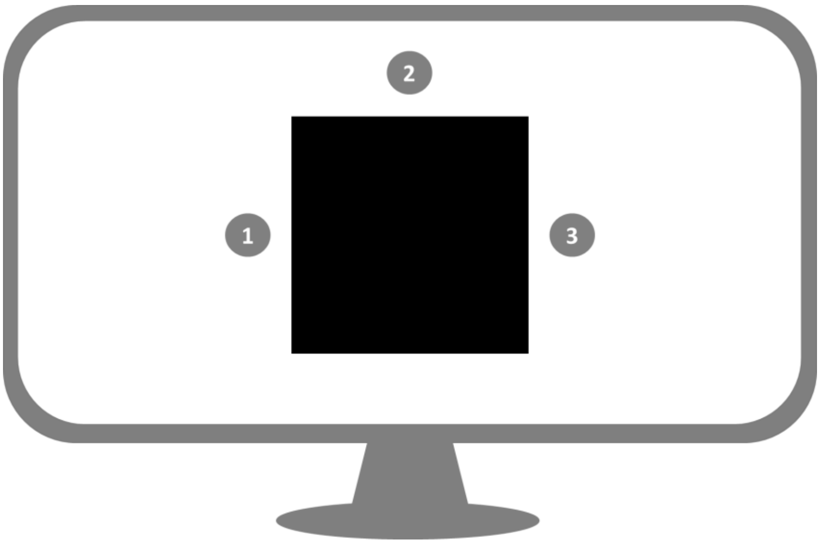


Figure 1.1: Simplified depiction of the OVSST. Target objects can appear at three different target locations presented on a computer screen.

1.1 Mental Model Theory and Probabilistic Mental Models (PMM)

Originally, the idea of mental models as cognitive representations of the real world was developed by Craik (1943) and later revived by Johnson-Laird (1983) who defined a mental model as a reasoning mechanism that is integrated in working memory. Nowadays, this vague definition of mental models is only one of many different definitions depending on the context. Mental models can be considered, for example, in the context of business, education or HCI and on an individual level or in teams. Mental models associated with teams, for example the context of sports, "[…] are organized mental representations of the key elements within a team's relevant environment that are shared across team members." (Mohammed, Ferzandi, & Hamilton, 2010, p. 1) and part of collective strategies in decision making. On the individual level Young and Veen (2008) emphasize the usefulness of mental models in business decisions. The authors describe methods like interview techniques to assess mental models which help to understand people's motivation and thought-processes.

In this thesis, mental models are considered in the context of HCI as long-term knowledge structures which represent the user's understanding of situation-specific system functioning (Durso & Gronlund, 1999). At the beginning of the mental model development, mental models are often incomplete and systematic faults may occur due to the abstract and schematic knowledge structure of mental models (Chalmers, 2003; Durso & Gronlund, 1999; Johnson-Laird, 2010). Chi and Roscoe (2002) report in their study about conceptual changes, for instance, that over half of the tested students in their study had coherent, but flawed mental models concerning the circulatory system of the human body. They assumed incorrectly that the circulatory system is a "single-loop" in which the blood flows from the heart to all parts of the body and not a "double-loop" from the heart to the body and from the heart to the lungs. Thus, systematic errors occurred which can be fixed with greater knowledge and adaption of the mental model. However, mental models seem to be resistant to change once proved useful. If information is once stored and an understanding is built up (e.g., when playing cards and having an understanding of the specific rules), it is difficult to change the model and adapt to new situations. This is comparable to the accommodation process described earlier by Piaget (1952), who states that existing knowledge has to be changed when no longer applicable.

Mental models become more detailed over time when users are more experienced. However, as already indicated by the word "model", mental models do not contain all available information, since perception and memory capacities are limited. Thus, mental models are characterized by a reduced copy of the reality. Further, a range of interindividual differences regarding the mental model development are based on different experiences and cognitive abilities of people (Jipp, 2016). In addition, mental models are hypothetic, imaginary and represent only what is true, but not what is false (Dutke, 1994; Johnson-Laird & Savary, 1996) probably to minimize the load in working memory. If someone, for instance, acquires a mental model about the electricity flow that is similar to how water is flowing, then this mental model does not contain the fact that water cannot flow upwards like electricity. In contrast, it implies that there is some time between the starting and the end point of the flow which is true for electricity as well as for water flows. This phenomenon of information accumulation is comparable to the confirmation bias describing the search for evident information that fits the existing beliefs and expectations (Nickerson, 1998). Summarizing the above, mental models share some specific characteristics and have to be differentiated from conceptual models and situation models. Conceptual models of an application describe the application on a high-level, viz. it describes the functioning of the application and concepts users need to understand and use this application (Gentner & Stevens, 1983; Johnson & Henderson, 2012). Mental models also have to be differentiated from a situation model that is build up in a particular situation and contains information about the environment of this situation (Durso & Gronlund, 1999; van Dijk & Kintsch, 1992). Thus, one mental model can be evoked by several situations with particular characteristics.

In contrast to the above mentioned aspects of mental model theory, the theory of Probabilistic Mental Models (PMM) by Gigerenzer et al. (1991) proposes a more specific view on mental models with concrete behavioral predictions in the context of uncertainty. This theory assumes that during problem solving, people build up a mental model about a relevant aspect of the reality. The uncertainty arising from the process of representation is also part of the model, for instance, if the problem solver clearly knows the solution for the problem, then uncertainty is not present. However, if uncertainty about the solution exists, then a PPM is established (Jungermann, Fischer, & Pfister, 2010, p. 177). A PMM considers the problem to be solved in a larger context by using a network of variables and "[…] connects the specific structure of the task with a probability structure of a corresponding natural environment (stored in long-term memory)." (Gigerenzer et al., 1991, p. 4). If, for instance, a decision between two options has to be made, a reference class activates all information relevant for the decision and enables to elicit valid probability cues which work as a predictor for the correct decision based on a probability estimation. Gigerenzer assumes that these probabilistic inferences are purely cognitively processed (Gigerenzer et al., 1991). PMM is based on different strategies like Take The Best (TTB), Take The Last (TTL) and minimalist strategy, assuming a random generation of cues until a discriminatory cue is found that allows to distinguish between different options (Dougherty, Franco-Watkins, & Thomas, 2008). The first strategy seems to be most relevant for this work, as studies showed that non-compensatory strategies, especially TTB, enable to predict the inference of decision makers under time pressure and when cue information has to be retrieved from memory (Rieskamp

& Hoffrage, 2005). The experimental task used in this thesis comprises of similar characteristics. In the following, TTB is presented in detail (cf. Bergert & Nosofsky, 2007).

TTB can be divided into five-steps in order to decide, for example, between two options as shown in Figure 1.2. A possible task might be to decide which of two football clubs has more supporters. The first step entails the recognition processes. If only one of two options is recognized, then the recognized option is chosen, for example, if only one football club is known, then this club is chosen. If both events are not recognized, then the decision maker guesses and chooses the football club with the most supporters by chance. If both football clubs are recognized, then the decision maker follows Step 2: searching for validity cues that provide information about the environment of the options, for example, whether the football club is located in a city that is a state capital or not. Thereafter, these cues are ranked from memory with regard to the relevance for the decision. After finishing this procedure, the decision maker chooses the highest ranked cue that seems to be most relevant for the decision and finds out if the chosen cue discriminates between the options (Step 3), for example, the size of population the city the football club is located in might be a cue that discriminates between the two football clubs. Step 4 stops the cue search if the cue discriminates between the options and the decision maker proceeds to Step 5. If there is no cue discrimination, then the decision maker returns to Step 2 and takes the next best cue which seems to be relevant for the decision and so on. Finally, Step 5 leads the decision maker to choose the alternative with the best cue value. If, for example, one of the football clubs is located in a bigger city than the other, this club is chosen. If there is no validity cue that discriminates between the options, then the decision maker chooses the event randomly, viz. they guess one of the options (Dougherty et al., 2008).

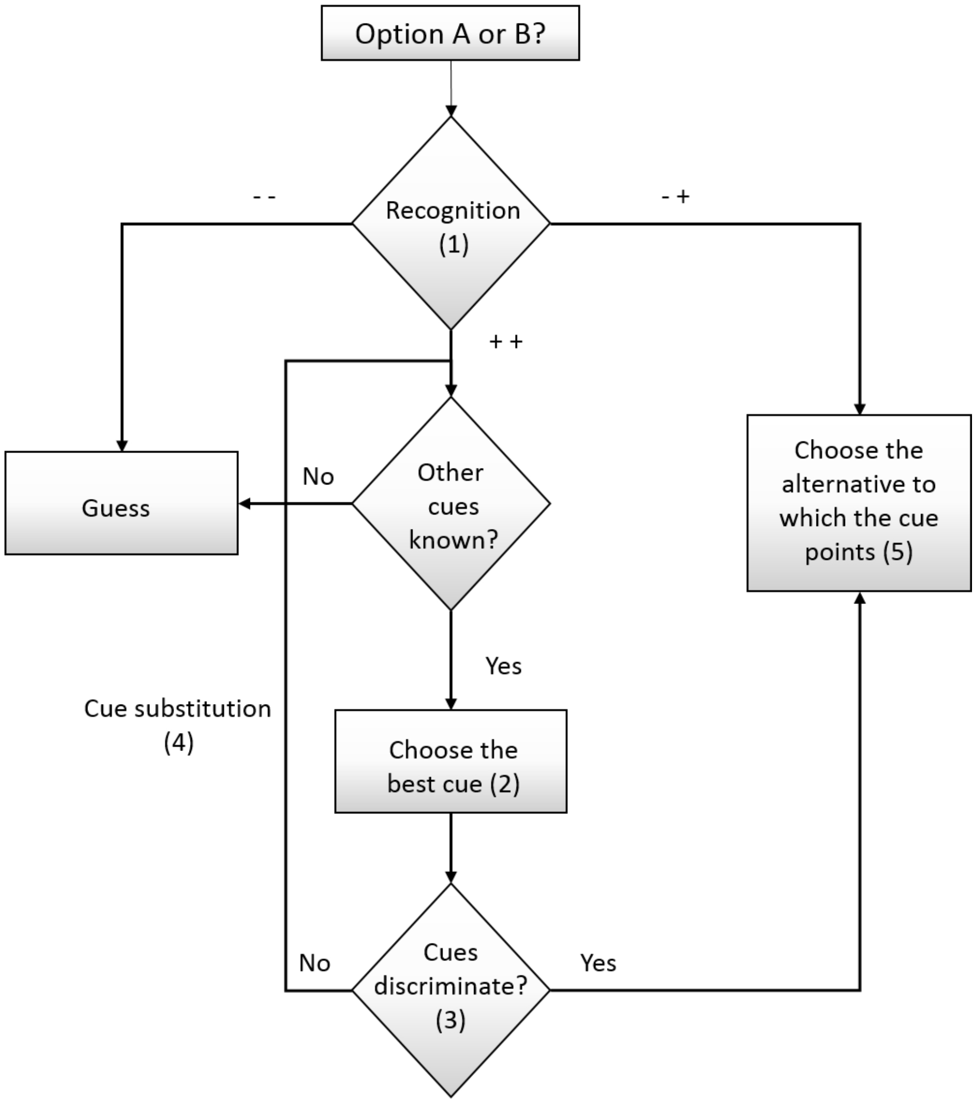


Figure 1.2: Flow diagram for the 5 steps of the Take The Best heuristic for choosing one of two options. Either no options (--), one of the options (-+) or both options (--) are known in advance influencing the further decision process. Adapted from “Psychological Plausibility of the Theory of Probabilistic Mental Models and the Fast and Frugal Heuristics”, by Dougherty et al. (2008, p.201).

1.2 Decision Making under Uncertainty

Mental models develop during decision making about relevant information as aforementioned and contain to a degree uncertainty about the outcome as described by PPM (Gigerenzer et al., 1991). Uncertainty is a broad concept and requires defining before it can be operationalized in scientific research. In the following, theories about uncertainty in the context of decision making are described in more detail. Lipshitz and Strauss (1997) provide distinct conceptualizations of uncertainty, for example, uncertainty related to ambiguity, risk or conflict. As mentioned earlier, task uncertainty is based on a lack of knowledge concerning the cause-and-effect relationship and leads to an inability to make a decision deterministically (Lipshitz & Strauss, 1997). As a working model, we assumed that uncertainty contains subjective uncertainty as well as objective uncertainty (Fig. 1.3). Subjective uncertainty describes the perceived uncertainty by the individual. In contrast, objective uncertainty is provided by the system and directly manipulable

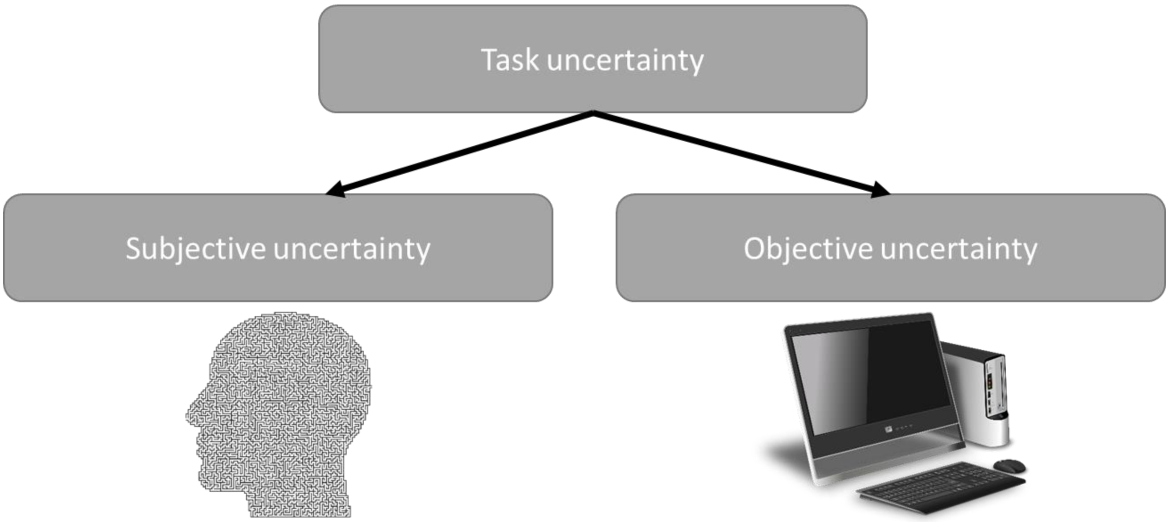


Figure 1.3: Task uncertainty is defined as the combination of subjective and objective uncertainty. Subjective uncertainty is perceived by the human and objective uncertainty is provided by the system.

Generally, humans try to minimize task uncertainty and therefore develop coping strategies (Camerer & Weber, 1992; Camerer & Weber, 1992; Smithson, 2009; Weber & Camerer, 1987). Typical examples for strategies to cope with uncertainty reported in literature are information search, suppression of ambiguous information or information reduction (Lipshitz & Strauss, 1997). During information search decision makers search for relevant information to discriminate between alternatives and finally to confirm their decision (Betsch, Funke, & Plessner, 2011). If there is much ambiguous information, decision makers often use the suppression strategy, viz. they ignore or distort undesirable information to reduce subjective uncertainty in the way that their preferences and beliefs align with their decision (Lipshitz & Strauss, 1997). Information reduction is a similar concept to suppression, but decision makers rather reduce the amount of task information and focus only on relevant information. Haider and Frensch (1999) state in the information-reduction hypothesis that task-relevant information is distinguished from task-irrelevant information with practice. Consequently, mainly task-relevant information is processed, resulting in better task performance and reduced uncertainty as, for instance, the number of alternatives is reduced. Another way to make decisions under risk or uncertainty is addressed in the so called expected utility theory (Neumann & Morgenstern, 1947; Neumann & Morgenstern, 1947). In this theory decision makers compare the subjective expected utility values of the risky or uncertain prospects to make their choice. However, these strategies are not always applicable in the real world as access to information is usually limited. Therefore, new theories of so called “bounded rationality” were developed, which emphasize the inability to consider all relevant information during decision making as a reason for uncertainty, lack of knowledge and costly information (Simon, 2000). Consequently, humans often make irrational decisions due to the limited information. They use cognitive heuristics – rules of thumb – to simplify the decision making process like the earlier mentioned TTB heuristic (Gigerenzer & Gaissmaier, 2011). Another cognitive heuristic is the satisficing rule which describes a decision strategy whereby the search for an alternative stops as soon as a satisfactory result is found, but not necessarily the optimal result (Simon, 2000). This fast and frugal characteristic of heuristics enables people to make economic decisions that, however, may lead to erroneous judgments (Kahneman & Tversky, 1973). There are also other characteristics of human thinking that lead to incorrect decisions, for example, the tendency to search for a causal relationship even if there is none. In contrast, some animals like rats choose a better strategy in these random decision situations as they do not tend to search for causality and are able to respond more randomly (Wolford, Miller, & Gazzaniga, 2000). Nevertheless, humans have the advantage to be able to learn which decision strategy might be optimal in non-random decision situations. Rieskamp and Otto (2006) suggest a theory about strategy selection learning (SSL) that is based on reinforcement learning, i.e. learning from the consequences of the decisions. Within SSL it is assumed that individuals select a decision strategy by considering their expectations that base on past experiences and improve the strategy selection through feedback over time. Thereby the learning effects, i.e. the behavioral change due to the learned decision-outcome expectations, are stronger at the beginning than at the end. The importance of learning in the decision making process is also shown in a general model of decision making provided by the framework of Rangel et al. (2008). This framework assumes different processes involved in the decision making process (Figure 1.4).

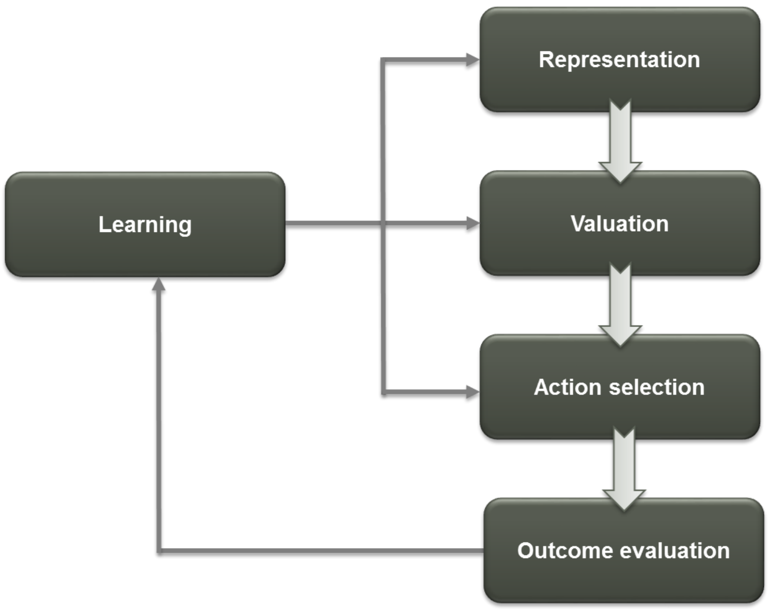


Figure 1.4: Framework for the 5 processes during decision making. Adapted from “A framework for studying the neurobiology of value-based decision making”, by Rangel et al. (2008, p. 546). For more details see main text.

At first, the decision maker creates a representation of the decision problem which involves internal as well as external states. Examples for an internal state might be fatigue level and for an external state threat level, for example, the difficulty of a task. This representation works as the basis for the second process, the valuation of different options. During the valuation process, the different decision options are weighted with regard to the likelihood to evoke a desirable state as a kind of preference formation. Then, the different options are compared with each other in the third process leading to action selection. Finally, the decision maker chooses the preferred option. After making the decision, the outcome of the decision is evaluated and finally used as feedback in learning processes to improve future decisions. Learning updates the other processes, i.e. the representation, the valuation as well as the action selection, except for outcome evaluation. These five processes presented in the framework, however, are not rigid with regard to the sequence and give only a rough overview (Rangel et al., 2008). The framework by Rangel et al. (2008) clearly shows that learning is an essential process during decision making that is closely linked to the representation of the decision problem. The aspect of uncertainty is not explicitly assumed in the presented framework, but it is integrated in the valuation process as the likelihood of the different decision options is taken into account. Thus, probabilities create uncertainty in the decision process as the outcome is not predetermined.

In literature different tasks are suggested to study decision making under uncertainty in laboratory settings that are based on a probability structure like lotteries (Arieli, Ben-Ami, & Rubinstein, 2011) or gambling tasks, for instance, the IOWA Gambling task (IGT; Bechara, Damasio, Tranel, & Damasio, 2005). In the IGT participants are instructed to maximize their profit during the selection of cards with monetary gain and losses. They have to decide between four decks of cards whereby two decks are riskier and have shortcomings in the long run. However, participants are not aware of that difference (Buelow & Suhr, 2009). The probabilities of losses and gains concerning the four decks can be varied to increase or decrease objective uncertainty. Nevertheless, the use of gambling tasks like the IGT in research has also disadvantages. For example, pathological gamblers show performance impairments in the IGT (Buelow & Suhr, 2009) and thus, gamblers' personalities might bias the results. Therefore, a task without gambling characteristics is chosen for the current experiments. Another bias might arise from the probability structure. Humans have difficulties to estimate probabilities accurately as numbers are abstract and often biased depending on memory performance as well as heuristics (Gigerenzer & Gaissmaier, 2011; Gigerenzer & Goldstein, 1996). For instance, high probabilities are often underestimated und low probabilities are often overestimated, because of regression to the mean (Beuer-Krüssel & Krumpal, 2009; Fischhoff & Beyth, 1975). In general, not only the probability structure, but also the induced uncertainty might evoke bias, as individual differences in perceiving uncertainty exist. For instance, people prone to sensation-seeking might perceive uncertainty as a motivating factor (Zuckerman, 1979) whereas other people more commonly associate uncertainty with anxiety as an inhibitory factor (Bammer & Smithson, 2008). In addition, the questionnaire-based study by Shuper et al. (2004) showed that also cultural differences exist with regard to the way how people deal with uncertainty. The results of the study indicated that Canadian students were more uncertainty oriented, i.e. they seek for information during uncertain situations, in comparison to Japanese students. In conclusion, studying probabilities seems to be biased in different ways. Nevertheless, the usage of probabilities to vary uncertainty systematically is the common method up till now (Smithson, 2009).

Evidence from neuroscience studies also suggests that uncertainty affects decision making. Functional magnetic resonance imaging (fMRI) experiments show that the activation of the prefrontal cortex is related to learning processes in uncertain situations (Fellows & Farah, 2007; Paulus et al., 2001; Platt & Huettel, 2008). In contrast, other studies report a larger neural circuit with the involvement of the orbitofrontal cortex and the amygdala that respond to the degree of uncertainty (Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005). Furthermore, other authors like Paulus et al. (2001) discuss a network of prefrontal, parietal and cingulate cortex during decision making under uncertainty. Overall, there seem to be many different brain regions involved in decision making under uncertainty, presumably indicating that uncertainty evokes more complex cognitive processing. The aim of the current work, however, is to assess uncertainty on a rather behavioral basis with the main focus on eye movement behavior. The following chapter explains how eye movements might provide an additional path to explore the central nervous system during the investigation of decision making under uncertainty.

1.3 Eye Movements, Attention and Learning

As already mentioned eye movements may be an important indicator to elucidate search behavior and several strategies respectively. Eye movements are motor skills used to rotate the eyeball in all directions consciously or unconsciously by the oculomotor system which is part of the central nervous system. The three most studied eye movement patterns are saccades, fixations and smooth pursuit movements. Saccades are rapid eye movements with a velocity between 30 and 500°/s lasting 30-80 ms (Holmqvist, Nyström, & Andersson, 2011, p. 23). They can be voluntarily and reflexively used to move the fovea to another position in the visual environment. The fovea is a small area of the visual field not larger than 2° in which humans have full visual acuity and are able to process the perceived information most efficiently (Holmqvist et al., 2011, p. 21). However, there is a field of view surrounding the fovea called perceptual span in which information is also processed even if the acuity is reduced. The size of the perceptual span depends on the nature of the task and can be at least 24°, for instance, during driving (Holmqvist et al., 2011, p. 381). During fixations, the eyes do virtually not move. However, this phase is characterized by micro-movements consisting of tremor, drift and microsaccades lasting 200-300 ms. Thus, fixations are often defined as a timespan in which the eyes remain relatively still. A series of both, fixations and saccades, is called scanpath. Smooth pursuits are used to track a moving target with the eyes driven by parts of the brain that are not activated during saccades (Duchowski, 2007, p. 46; Holmqvist et al., 2011, pp. 22–23). Some authors suggest that only during fixations new information is processed (Gilchrist, 2011; Rayner, 2009). Nevertheless, gaze shifts seem also to facilitate stimulus processing and influence the development of the mental representation (Irwin, 2004). Further, it “[…] seems reasonable to assume that fixation location corresponds to the spatial locus of cognitive processing and fixation or gaze duration corresponds to the spatial locus of cognitive processing of the material located at fixation.” (Irwin, 2004, p. 106). In other words, eye movements and attentional processes seem to have an impact on which stimuli are cognitively processed in what manner. This idea was earlier expressed in the aforementioned eye-mind hypothesis by Just and Carpenter (1980), stating a strong coupling between eye movements and attention (cf. Theeuwes, Belopolsky, & Olivers, 2009). Hence, a stimulus is cognitively processed for the same amount of time the person has fixated on it. This hypothesis, however, has its limits, for example, when considering memory retrieval processes as already shown by Anderson et al. (2004). The authors assume that retrieval processes are not necessarily reflected by eye movements. Moreover, Irwin (2004) highlights four main issues regarding the eye-mind hypothesis: First, the attentional focus of cognitive processing within the complete field of view may be wider than the fixation location as information is also processed in the visual periphery (cf. Fox, Merwin, Marsh, McConkie, & Kramer, 1996). Second, the focus of cognitive processing is often separated from the fixation location. Third, salient stimuli that pop out can guide the eye without cognitive control and without active search (bottom-up driven). Fourth, cognitive processing may happen during saccadic eye movements and during fixations. Further, Hyönä (2010) mentions that even if a learner attends to the relevant stimulus for a while, there is no necessity that the learner adequately comprehends the stimulus. In contrast, if the learner adequately comprehends the stimulus, the stimulus had to be attended to initially (top-down driven). Thus, attentional processes are at least partly reflected by eye movements and essential for information processing and further cognitive processing, for instance, learning.

The relation between the allocation of attentional resources and cognitive processing is already described by Wickens and Hollands (2000) in their model of human information processing. Their framework gives an overview of the processes which are involved in information processing and underlines the relevance of sensory and perceptual processes as the basis for cognitive processes which are relevant in the current thesis. There are a series of stages during information processing: sensory processing, perception, cognition, memory, response selection and execution. Perceptual processing is driven by bottom-up controlled sensory input as well as by top-down controlled input from long term memory including past experiences (Wickens & Hollands, 2000). Shifts of attention are either bottom-up driven (stimulus-driven), for example, if a salient stimulus captures attention automatically, or volitional, top-down driven (goal-driven) as knowledge about the task is activated (Buschman & Miller, 2007). The activated information is temporarily stored in working memory which is involved in all cognitive processing that is conscious and resource limited. In addition, a feedback loop is used for the evaluation of the goal achievement and suggests that there is no defined starting point for the flow of information. Finally, the model also takes into account that attention has to be allocated as resources are limited. These relations have to be considered when developing the experimental design of the thesis.

In order to investigate sensory and perceptual processes of visual stimuli, the visual search task (VST) was developed. This standard paradigm of visual search in basic research involves the detection of a target object amongst distractors (Eckstein, 2011; Huang & Pashler, 2005; Treisman & Gelade, 1980; Wolfe, 1994). Most commonly, set size and target-distractor discriminability are varied and the effect of this variation on reaction times is studied. The higher the target-distractor similarity, the longer the reaction time. However, the standard paradigm of visual search considers rather attention allocation than eye movements as attention shifts precede gaze shifts (Kristjánsson, 2011). In the following the term visual search behavior also implies eye movements.

Eye movements seem to be strongly coupled with attention as neuronal mechanism of attention and eye movements overlap to a great extent. Further, the evidence that attention is shifted before eye movements follow, reinforces the coupling between eye movements and attentional processes (Kristjánsson, 2011). Attentional processes and thus, also eye movements, are essential for cognitive processing already shown by top-down processes which rely on prior knowledge and guide attention voluntarily. Some authors summarized that eye movement metrics are related to cognitive processes reported in literature about evaluating the usability of technical systems (Ehmke & Wilson, 2007; Poole & Ball, 2006). They mainly mention that eye movement metrics, like fixation duration and number of fixations, indicate the efficiency of visual search and the importance or saliency of presented elements (Goldberg & Kotval, 1999). Other authors argue directly that eye movements "[…] provide an excellent on-line indication of the cognitive processes underlying visual search and reading." (Liversedge & Findlay, 2000, p. 6).

Besides of eye movements even blinks seem to be related to information processing (Wascher, Heppner, Möckel, Kobald, & Getzmann, 2015). Blinks describe the unconscious closure of the eye for a short moment and lead to a temporary loss of visual information. They are inhibited more often during the presentation of relevant information and thereby reflect the amount of information processing (Fogarty & Stern, 1989). Thus, it seems to be obvious that eye movement and blink behavior does not reflect only sensory but also cognitive processing. Especially top-down processes which are driven by experience and expectations should be reflected by eye movement behavior. Experience and expectations, however, are also strongly related to mental models.

Eye movement tracking technology has improved considerably in the last 10 years. Therefore, many studies investigated eye movements in relation to attention and learning processes. Eye movements are automatic, fast and frequent (Irwin, 2004; Spivey & Dale, 2011). These characteristics make eye movements an ideal dependent variable because they imply objectivity, might respond quickly to experimental manipulations and provide sufficient data for statistical analysis. Eye movements are not randomly distributed, but rather systematic and driven by automatic bottom-up and volitional top-down processes (Theeuwes, 2010). Additionally, eye movements seem to differentiate between experts and novices as experts usually possess a more accurate mental model than novices. In their study, Charness et al. (2001) report different eye movement patterns during chess for expert and intermediate players. Intermediate chess players showed a higher fixation frequency and smaller amplitudes of saccades, i.e. smaller angular distances of the saccades, than experts. Also Tai et al.

(2006) found fewer eye fixations and fewer gaze shifts between the displayed elements for experts than for novices during a science assessment amongst students studying different disciplines of science. Finally, developing expertise also depends on information processing abilities and working-memory capacity as described by Jipp (2016) which might be reflected by eye-movements.

Lai et al. (2013) emphasize in their review that eye movements enable to gain insights into the relationship of effects of learning and other cognitive activities. Learning processes are involved in the development of mental models and often measured by using methods like questionnaires, think-aloud protocols, diaries or interviews that are subjective and require a certain amount of consciousness (Lai et al., 2013). However, these methods seem to interfere with cognitive processing (Ericsson & Simon, 1980; Schooler, Ohlsson, & Brooks, 1993) and additionally do not allow to study stimulus driven processes people are not aware of. Thus, studying eye movements as a more objective measurement might give new insights into the development of mental representations.

1.4 Further Evidence: Eye Movements and Mental Representations

Eye movements could inform about cognitive states of the human in different field of application, for example, in decision making, problem solving, skill and knowledge acquisition described in the following.

Decision making:

Orquin and Mueller Loose (2013) published a review about eye movements in the context of decision making. In their review, they described that only stimuli that are fixated or inside the perceptual span of the nearest fixation can be considered within the decision-making process. Further, there is a robust gaze bias towards the finally chosen object indicated by fixation duration or/and fixation frequency (Glaholt & Reingold, 2009; Glaholt, Wu, & Reingold, 2010; Hegarty, Mayer, & Green, 1992; Shimojo, Simion, Shimojo, & Scheier, 2003; Tsai, Hou, Lai, Liu, & Yang, 2012) and might be used to predict selection, for example, of preferred faces (Glaholt, Wu, & Reingold, 2009). Thus, eye movements are often seen as a useful measurement to study decision making, as fixations can reveal the decision beforehand (Glaholt et al., 2010).

Problem solving:

In the field of problem solving researchers make use of eye tracking as it allows to get insights into the development of the mental model that also encloses the solution knowledge of the problem. Studies, for instance, already showed that eye movements and attentional processes can discriminate between successful and unsuccessful problem solvers. Successful problem solvers, for instance, focus more on relevant cues than unsuccessful problem solvers (Grant & Spivey, 2003; Tsai et al., 2012). Further, Ellis

(2012) investigated solving anagrams while eye movements were monitored. At the beginning of the experiment participants focused on distractors and target objects for the same duration of time. However, this pattern changed over time. Viewing times on the target object increased while simultaneously decreasing viewing times of the distractors.

Skill and knowledge acquisition:

In skill acquisition research studies showed that eye movements can discriminate between novices and experts, for instance, experts attend earlier and longer on task-relevant information comparable with the findings in the context of problem solving (Charness et al., 2001; Hyönä, 2010; Jarodzka, Scheiter, Gerjets, & van Gog, 2010; Tai et al., 2006). A greater visual focus lies on the relevant stimulus as shown by Jacob and Hochstein (2009), who report higher fixation frequency and longer fixation duration of detected card pairs than undetected card pairs in the Identity Search Task. This task requires to find two identical card pairs amongst twelve cards. They assume a three-stage model of the perceptual recognition process while searching for card pairs. This model seems to be highly relevant for the current thesis with regard to the interaction between learning processes and eye movements. In the first stage fixations seem to be randomly allocated whereas in a second stage, an implicit recognition of the card pairs may guide eye movements to the target location. In the last stage, the implicit knowledge develops into explicit knowledge which leads directly to an attention allocation on the target stimulus and accelerates reaction times.

This model seems to fit well to the three stages of the mental model acquisition reported by Schumacher and Czerwinski (2014) which was inspired by memory theories. Schumacher and Czerwinski (2014) also assume a three stage process, however, in contrast to Jacob and Hochstein (2009), the emphasis of their research lies on the acquisition of expert knowledge in the context of physical systems. The first pretheoretic stage is based on the fact that earlier experiences, as similarities to prior events, are used to understand the functioning of a physical system. The second experiential stage includes the development of some understanding of causal relationships and first abstractions that arises from recognition of similar characteristics across instances. In the final expert stage abstractions across different system representations take place even if dissimilarities exist superficially. In this stage users are able to recognize systemic patterns and retrieve prior knowledge about the system. Finally, knowledge can be easily transferred. Thus, experts developed a more abstract mental model including less irrelevant information than novices.

A further theoretical approach is provided by Ackerman (1988), who also suggests three phases of skill acquisition and complements the aforementioned models. The author proposes that the three skill-acquisition phases are linked with three ability classes: general intelligence, perceptual speed and psychomotor abilities which are a basis of individual differences in performance. However, this theory ignores individual experiences and internal processes during skill learning (Langan-Fox, Armstrong, Balvin, & Anglim, 2002). The first phase of Ackermann's skill acquisition theory is mainly associated with general abilities and demands cognitive abilities such as knowledge retrieval or reasoning. The influence of this intellectual ability on performance diminishes with practice. In the next phase, perceptual speed becomes more important for performance when rules for performance are adopted by the learners and attentional demands are reduced. Finally, in the third phase, processes become more autonomous and almost free of attentional demands. Thus, performance is more reliant on psychomotor skills (Ackerman, 1990). This last phase might be comparable with the proceduralization process, as declarative knowledge is no longer required and processes are transferred into automated actions (Anderson, 1982; Pomerol & Brézillon, 2003; Pomerol & Brézillon, 2003). In the following, the reviewed literature about cognitive processes, behavioral aspects and eye movements are brought together.

1.5 Overview and Research Objectives

The three aforementioned theoretical concepts of perceptual recognition (Jacob & Hochstein, 2009), mental model acquisition (Schumacher & Czerwinski, 2014) and skill acquisition (Ackerman, 1988) all assume a three-staged process and show a comparable structure from an novice to a more experienced state. These models might be strongly related to each other, as perceptual recognition processes are part of the mental model acquisition which in turn is also part of the skill acquisition. In other words, a mental model can only be developed if information is perceived and if the information stored in the mental model is finally used to prepare and execute actions. Thus, it seems reasonable to suggest that the stages of the models run in parallel and influence each other. If this is the case, fixations should reflect the state of learning as well as the state of the mental representation (see Figure 1.5). The combination of the three models would lead to the assumption that at the initiation of the mental model eye fixations are more randomly distributed because information has to be accumulated. With increasing expertise and a more abstract mental model, eye movements should become more focused on relevant stimuli information, because knowledge about the learning environment is still available. Studies already showed that the acquisition of mental representations is reflected by eye movements especially if participants look at "nothing". The "looking at nothing" phenomenon describes people's focus on spatial locations to retrieve information from memory which were present at these locations, even if the information is no longer available there. In other words, they are looking at nothing to facilitate memory retrieval (Scholz, Helversen, & Rieskamp, 2015). It seems to be obvious that the acquired knowledge should also influence visual search behavior.

Bowden (1997) already described that a mental representation is created first and then the stored information is used for searching additional relevant information in the environment as well as in memory. In the current study, an initial mental representation is developed by instructions informing about the task procedure and training trials before the actual experiment. Further, the validity and accuracy of the mental model is assessed via questionnaire. However, Gentner (2001) already stated that it is not sufficient to ask people directly about their mental model as they are often unable to verbalize their knowledge. Thus, additional behavioral data is needed to ensure that a mental representation has actually developed. In general, performance should increase over the time course of the task, as a more accurate mental representation develops simultaneously. In parallel reaction times are expected to speed up as actions become more autonomous with increasing knowledge as shown by Jacob and Hochstein (2009).

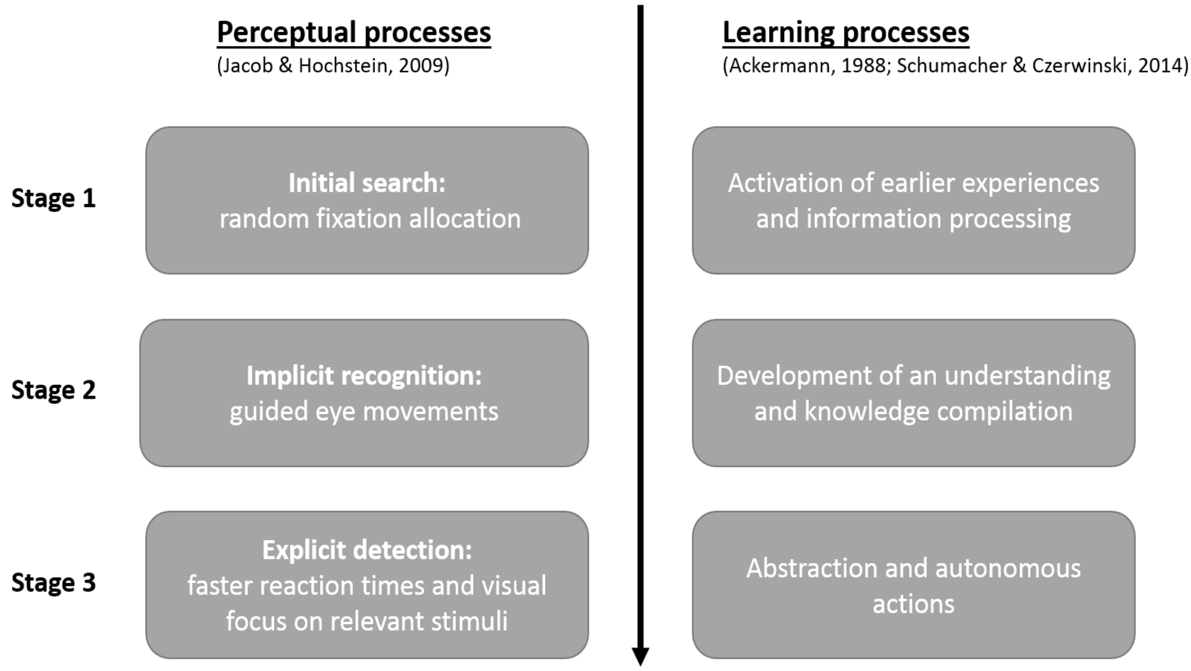


Figure 1.5: The three-stage models in comparison. Depicting the main statements of the proposed model by Jacob and Hochstein (2009) including perceptual processes and eye movements and the two learning processes concerning skill acquisition by Ackermann (1988) and mental model acquisition by Schumacher and Czerwinski (2014).

This thesis aims to study eye movements during the learning of a probability concept in an experimental setting. As mentioned earlier, probabilities are a common method to manipulate uncertainty and the results of the probability learning can be inquired explicitly. The research focus on representations of uncertain concepts is relevant for the investigation of mental models even though not much attention has been paid to uncertainty so far. In decision making, uncertainty is often investigated in the context of risky decisions (Glöckner & Herbold, 2011; Venkatraman, Payne, & Huettel, 2014). However, gains and losses of the used gambling tasks in these studies influence the participants' behavior and are accompanied by other task characteristics. Thus, this paradigm is not appropriate for the current thesis.

In visual search research, attention was given to saliency rather than uncertainty, even if task uncertainty is always present. Participants have to search for a salient stimulus, however, little consideration is given to the spatial uncertainty of the stimulus in the search paradigm, i.e. the location of the stimulus appearance is uncertain. Furthermore, it is still unclear how task uncertainty affects the mental model acquisition of, for example, technical systems. In this thesis, behavioral data and reported probabilities retrieved from memory are expected to provide insights into mental representations. Eye movement patterns are supposed to change in parallel to the learning process and the mental model acquisition reflecting cognitive processing. During the learning process, feedback as well as motivational processes are essential (Yi & Davis, 2003). In order to investigate the theoretical question empirically, an attempt was made to develop a new paradigm for a consistent implementation: Occluded Visual Spatial Search Task (OVSST; see Chapter 2.2 for details). This research paradigm is successively expanded and further developed over the course of the thesis. All experimental manipulations aim to manipulate the learning process and thus, the development of the mental representation. In the following an overview of the research objectives of the series of experiments is provided.

1.5.1 Experiment I: Predicting the Appearance of Geometric Objects

Humans seem to have an individual mental model from the external reality they are familiar with (Johnson-Laird, 1983; Jones, Ross, Lynam, Perez, & Leitch, 2011). Thus, it is necessary that participants have to learn new associations of rather unfamiliar and abstract constructs when investigating the acquisition of new mental models. The OVSST should meet that criteria. It was expected that participants learn the underlying concept of the task, that is explained in the next chapter, and thus, were able to improve task performance, viz the number of correct predictions. Thereby, eye movements should reflect the ongoing cognitive processes. The associated experiment reported in Chapter 3 tries to answer the following question.

Research question 1: Do eye movements allow insights into the acquisition of mental models under uncertainty during the performance of the newly developed OVSST?

1.5.2 Experiment II: Degradation of Search Targets

In this experiment we aimed to make visual search more difficult (Kumada, 1999; van Zoest & Donk, 2004) which is expected to interfere with the acquisition of an accurate mental model. Therefore, target objects were degraded (e.g., Sternberg, 1969) viz. background texture was changed with the aim to reduce target perceptibility. The effect of the disruptive background on visual search behavior and learning processes was tested in Chapter 4 to answer the following question.

Research question 2: Does the degradation of target stimuli interfere with the development of accurate mental models?

1.5.3 Experiment III: Dynamic Relearning

The findings from the previous experiments suggested that eye movements indeed inform about the state of learning and the degree of uncertainty during the development of mental models. However, learners usually not only have to learn a new context, they also have to relearn initial associations they are familiar with. Thus, they have to deal with problems like functional fixedness which is a cognitive bias preventing people to think outside the usual action strategies (Adamson, 1952; Knoblich, Ohlsson, & Raney, 2001). Existing mental models have to be adapted to the new task characteristics and might impair the learning process. In order to test this in an experimental setting, the OVSST was designed as relearning experiment with a learning and a relearning phase. In Chapter 5 it was analyzed if previous findings with regard to the eye movement patterns can be transferred to the relearning process by answering the following question.

Research question 3: Do eye movement patterns indicate different phases during relearning?

1.5.3 Experiment IV: Separate Tasks

OVSST as a new paradigm to gain insights into the acquisition of mental models seems to work well with regard to the findings in the previous experiments. However, so far it is not clear whether participants developed the mental model based on the first part of the OVSST, the goal-driven prediction task, as intended, or based on the second part, the stimulus driven reaction task. Therefore, Experiment 4 evaluates underlying processes during the performance of OVSST by testing both tasks separately. Finally, the following question should be clarified.

Research question 4: Is the developed mental model mainly based on the prediction task or on the reaction task?

1.5.4 Experiment V: Learning Different Probabilities

In all of the previous experiments the degree of objective uncertainty remained constant. In Experiment V the probability distribution was manipulated to investigate the effect of objective uncertainty on visual search behavior. Participants had to learn a probability concept with a higher and a lower probability distribution to answer the following question.

Research question 5: Do participants show more visual search behavior when dealing with a higher degree of uncertainty according to the assumption that information search is used to cope with uncertainty?

2 General Method

In the following I present all methodological details that apply for every experiment if not specified explicitly. That includes a description of the tested sample, the main experimental task, pre- and post-tests to control for cofounding variables, the design and procedure of the experiment, the used eye tracking method and a validation study of two different eye-trackers and finally, the data analysis. In a separate *Method* section of each experiment the method will be specified.

2.1 Participants

A total of 92 right-handed participants, between 18 and 35 years old, took part in the five experiments. All participants obtained informed consent and were naïve about the study's purpose before participating at the experiment. No participants took part in more than one experiment. Furthermore, all participants had normal vision without eyeglasses or contact lenses. At the end of the experiment they received either course credit or were paid for their participation (10€/10 C$ per hour).

2.2 Occluded Visual Spatial Search Task (OVSST)

The OVSST was developed for the purpose of this work. It is inspired by the tunnel effect that describes the visual perception when objects are hidden in a tunnel and finally return with different velocities (Burke, 1952; Flombaum & Scholl, 2006)[1]. The aim of the task was to investigate how people build up a new mental representation of a spatial probability concept over time without prior knowledge to avoid familiarity biases (Ellis, 2012; Starling, 2012). Further, the task should fulfill two requirements: 1) there should be enough capacity in the visual working memory to be sure that the presented stimuli are perceived, 2) stimuli have to be salient to gain attention. Then, allocated information can be processed, stored in long-term memory and retrieved from memory, which are the prerequisites for the development of a mental representation.

In the OVSST the presented stimuli are designed as dark grey (RGB: 128, 128, 128) geometric objects with a diameter of 2 cm. A black square ("room" or "tunnel") is constantly presented on the screen and is of the dimensions 20x20 cm positioned in the middle of the screen with three exits and a bottom entrance each 2 x 0.5 cm. Participants are instructed to perform two tasks: a prediction and a reaction task. First, participants observe one of three distinct dark gray objects (triangle, square or circle) moving from the bottom entrance into the quadratic room with the three exits (Fig. 2.1). After the object disappeared, participants have to predict as accurately as possible with the arrow keys (left,

[1] This concept of tunnel effect has to be distinguished from the tunnel effect which describes limitations in the visual field. The latter meaning is less relevant in the context of the current research paradigm.

top, right) on the keyboard at which of the exits (left, top, right) these objects will reappear (Instruction: "Versuche so gut wie möglich mit den Pfeiltasten (links, oben, rechts) vorherzusagen an welchem Ausgang das Objekt erscheint"). Participants were asked to put the index finger on the left key, the middle finger on the top key and the ring finger on the right key. Each object was associated with a higher probability of 74 % to one of the exits and with a lower probability of 11% to each of the other two exits (cf. Miller, 1998). In 4% of the trials the object reappeared at the entrance. In order to perform the prediction task as accurate as possible, participants have to learn the probability concept. Finally, the object reappears at one of the exits. The response was designed as a go/no-go task, thus in 50 percent of the trials the object reappeared with a changed color intensity (light gray, RGB: 223, 223, 223). In these trials participants had to press the appropriate arrow key as quickly as possible, for example, the left arrow key if the object reappeared at the left exit. If there was no color change, participants were instructed just to observe the situation (Instruction: "Sobald das Objekt aus einem der Ausgänge wieder austritt, solltest du schnellstmöglich die entsprechende Pfeiltaste drücken, aber nur wenn sich die Farbe des Objekts geändert hat. Ansonsten wartest du ab bis der nächste Durchgang beginnt."). This task was employed to force participants to process stimulus information and not only the physical emergence of the stimuli. In general, participants had to search visually for information at the three exits in order to receive performance feedback. This feedback is integrated in the mental representation of the participants so that they learn the probability structure of the OVSST and are able to improve the performance.

As mentioned above the OVSST contains another special characteristic: In 4% of the cases, objects reappear at the bottom entrance as a rare occurrence. The aim of this rare occurrence was to analyze how people deal with a special unexpected and uncertain situation. Auditory feedback indicates erroneous task performance in the prediction task, i.e. if the arrow key is pressed too late ("zu schnell") or too early ("zu langsam"), and in the reaction task if participants respond incorrectly ("falsch reagiert"). At the end of the task, participants are asked to complete a paper-pencil questionnaire concerning the participants' conscious representation of the probability concept – the Concept Awareness Questionnaire (see Appendix A).

The described procedure of the task was the basic version of the OVSST used in experiment I (see Fig. 2.1 for a schematic description). In further experiments stimuli, probabilities and procedures of the OVSST are manipulated in accordance to the research questions described in the introduction. The motivation for the manipulations is described in detail in the respective chapters. Task uncertainty is caused by the instructions, the probability structure of the OVSST as well as the rare occurrence of the stimuli at the bottom entrance. It includes the objective and manipulable uncertainty created by the task as well as the subjective uncertainty perceived by the participants. Over the course of time performing the OVSST, task uncertainty should be more and more dominated by objective uncertainty as the familiarity with the task characteristics increases.

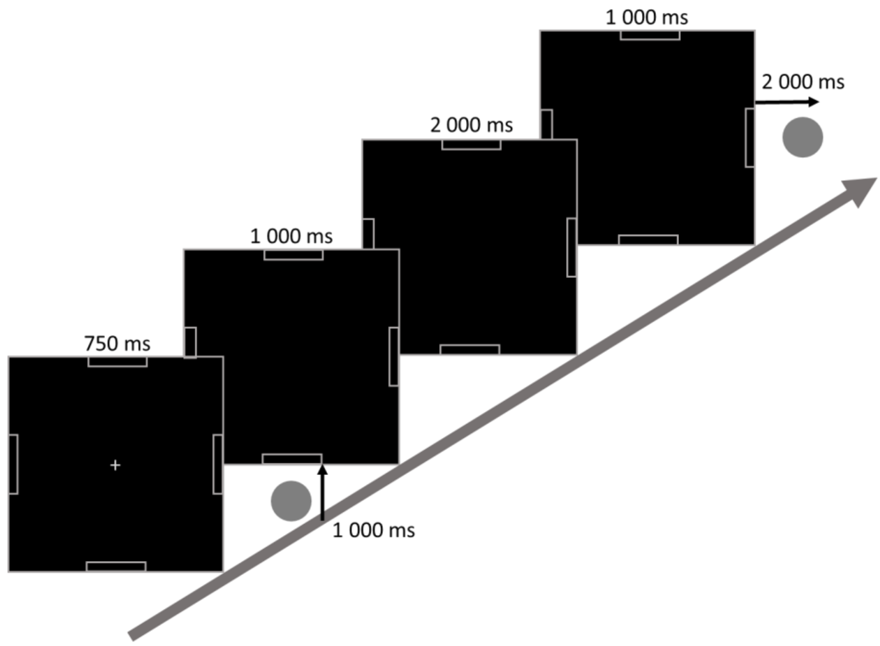


Figure 2.1: Schematic description of one trial of the OVSST: First, participants fixate a fixation cross in the middle of the room that appears for 750 ms. Then, one of the three objects (triangle, square, and circle) appears at the bottom entrance of the room for 1 000 ms and moves into the room. After 1 000 ms the object has disappeared into the room and participants are instructed to predict at which exit the object will reappear with the left, right and upper arrow key of the keyboard within 2 000 ms. Finally, the object reappears at one of the exits. The object moves out of the room to the exit position for 2 000 ms and pauses there 1 000 ms. If the object reappears with a changed color intensity, participants have to press the appropriate arrow key again. In total, a trial lasts 7 750 ms.

2.3 Pre- and post-tests

In our study, changes in attention capacity as well as motivation were additionally measured to check for the homogeneity of the sample and a potential interaction between these variables and the results of the experiments. Participants performed a computer version of the D2 Test of attention by Brickenkamp (1994) before and after the OVSST to control for systematic changes in attention capacity (e.g., due to fatigue) as a confounding variable. This test measures accuracy and speed during the discrimination of similar visual stimuli. The test consists of 14 lines with 47 symbols each. These symbols consist of the letters d or p with 0 or up to 4 dashes. Participants were instructed to mark all ds with two dashes and to ignore all other symbols. Each line shows 21 times the letter d with two dashes amongst 26 distractor symbols and is displayed for 20 seconds. (Brickenkamp & Zillmer, 1998).

Different parameters can be analyzed such as quantitative performance (quantity/speed), qualitative performance (errors/accuracy) and total error count. Bates and Lemay (2004) conclude that the D2 is an internally consistent and valid measurement of attention.

The Questionnaire on Current Motivation (QCM) developed by Rheinberg et al. (2001) was used to control for motivational aspects on behavioral performance. This questionnaire measures four motivational factors with a total of 18 items: anxiety, probability of success, interest and challenge. Rheinberg et al. (2001) report an internal consistency between Cronbach's alpha $\alpha = .66$ and $\alpha = .90$ for their scale which is deemed to be sufficient. In our studies, the items were reformulated into past tense because participants finished the questionnaire after the experiment. Thus, the items refer to a task which lies in the past. Participants answered the items on a seven-point scale. Overall results showed that participants perceived the OVSST as challenging, but they had stronger hope for success than fear to failure, presumably due to the laboratory setting.

2.4 Design and Procedure

Participants were tested in a single session that took approximately one and a half hour. All used test methods and instructions were in German language except in Experiment IV. After they were informed about the procedure of the study and signed the declaration of consent, they performed the D2 test of attention. Then, the eye tracking experiment with the OVSST started. Participants were seated approx. 70 cm in front of a computer screen. Their head was positioned on a chinrest to minimize head movements. Furthermore, lighting conditions were kept constant. The stimuli were presented on a white background of a 27-inch monitor (Acer HN274H B) with a screen resolution of 1920x1080 (Fig.2.2). Each participant completed a practice block (18 trials) of the OVSST with three geometric objects that differed from the experimental ones. Participants were instructed and became familiar with the procedure of the task during the performance of the practice block. After this first practice block, a control session followed. In the control session participants were instructed to improve task performance, but they were not informed about any probability structure. This control session functioned as a control condition called 100% condition: each object, the same as in the previous sessions, was associated with one of the exits to 100%. Thus, every object reappeared only at its predefined exit. Participants performed 2 blocks with 21 trials each of this 100% condition. Finally, they performed 324 trials of the actual OVSST with the 74-11-11 probability structure divided into 4 blocks (81 trials each block). After each block, a fixed pause of 2 minutes occurred. Before every block the eye-tracker was calibrated with a 9-point calibration and the actual gaze deviation was measured. Participants with a deviation more than 2° were excluded from the experiment. After every block, participants received feedback about the percentage of correct answers depicted on the display to encourage an improvement of their performance. At the end of the study, participants completed the Concept Awareness Questionnaire of the OVSST, the QCM and a second time the D2 test of attention.

A power analysis was conducted using the software package G*Power 3 to determine an appropriate sample size for the following experiments (Faul, Erdfelder, Lang, & Buchner, 2007). The analysis for a repeated measures ANOVA with four measurements and within factors, showed that a design with total sample size of 20 participants has a Power of 73% with a critical F=2.766. This seems to be sufficient on the one hand and also practicable with regard to financial resources and human resources on the other hand. Thus, it was aimed to test around 20 participants to have a sufficient statistical power.

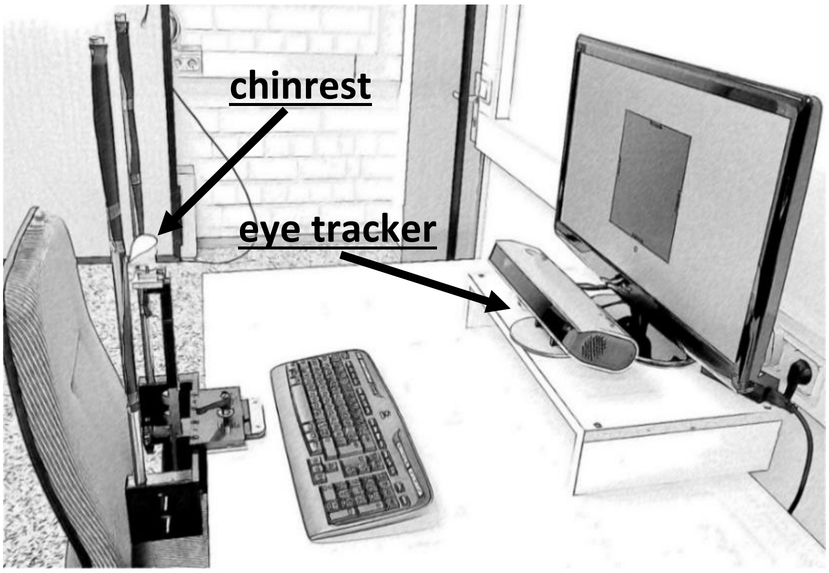


Figure 2.2: Experimental setup: Participants were seated in front of a screen and an eye tracker. The keyboard was used as input device. The chinrest was used to avoid head movements.

2.5 Eye Tracking

At the beginning of the experimental series, two different eye trackers were available: the SMI Red 500 (SensoMotoric Instruments, Teltow, Germany) and the EyeLink 1000 (SR Research, Ontario, Canada). The EyeLink 1000 allows for an average gaze position accuracy of 0.25°-0.5° and a sampling rate of 1000 Hz binocularly when the head is supported whereas the SMI Red provides a gaze position accuracy of 0.4° and a sampling rate of 500 Hz. Remote eye tracker like the SMI Red 500 are especially useful for the current studies because of different reasons. First, remote eye trackers use a contact-free measuring technique enabling a natural handling without restrictions. Second, head movements are allowed to a certain extent. Third, the time resolution is high – a sampling rate of 500 Hz measures the gaze position every 2 ms. However, the high time resolution causes big data sets which might make data handling more difficult, especially checking for plausibility and erroneous traces. The advantages of the EyeLink 1000 are the even higher time resolution as well as the higher possible accuracy. In order to find out which of the eye trackers is most suitable for the current studies, a validation study was set up to evaluate the precision (spatial variation between individual samples) and accuracy (spatial variation from the actual gaze point to the one measured by the eye tracker) of the recordings with both eye trackers. The same objects as in the OVSST were tested in this study to check for their usefulness at the same time. The experiment and the results are briefly described in Appendix B. Results of the validation study showed that data recorded by the SMI Red eye tracker was more precise than data of the EyeLink 1000. Furthermore, accuracy of the data did not differ between the recordings of both eye trackers. In addition, the calibration of the SMI Red eye tracker was less time-consuming and the remote version without head support was more practicable for the current experiments. Thus, we decided to use the SMI Red 500 remote eye tracker to track the eye movement binocularly in all of our studies. The system tracks the gaze position by using the corneal reflection technique: The cameras of the eye tracker detect the cornea reflection of the infra-red light and the gaze position is then calculated from the position of the corneal reflection relative to the pupil center (Duchowski, 2007, 54 ff.). Problems usually occur when detecting pupils of participants with dark eyes. Further, gaze position accuracy is lower for participants with contact lenses or glasses (cf. Nyström, Andersson, Holmqvist, & van de Weijer, 2013). For this reason, we tested only participants with normal vision who did not wear contact lenses or glasses. Moreover, participants were informed not to apply make-up to their eyes to prevent miscalculations of the eye tracking system.

2.7 Data Analysis

2.7.1 Data Preparation

Raw data from eye tracking recordings of the OVSST were aggregated per trial (e.g., number of fixations, fixation duration, etc.) for every participant. Before analyzing data, we cleaned the data set: All trials with missing reactions and/or missing judgments and/or trials with an insufficient amount of eye movement data points (less than 65%) were excluded from data analysis. Trials with missing reactions and/or judgments were not analyzed separately, because the error rate was too low and thus the resulting data set too small for appropriate analyses. Further, fixations within 100 pixels around the middle of the screen were excluded from data analysis to prevent inaccuracy as the allocation of the fixations to the areas of interest (AOI) could be difficult (Fig. 2.3).

2.7.2 Independent and Dependent Variables

The independent within-subject factors were *block* (1-4) and *judgment* (correct, incorrect). *Block* was employed to investigate the development of the dependent variables (see below) over time. *Judgment* was introduced to compare visual search behavior, reaction times as well as judgment times for correct and incorrect judgments.

The standard eye movement variables, already mentioned in Chapter 1.3, were used as dependent variables: fixation frequency, fixation duration, and number of gaze shifts. Fixations were detected with the saccade-detection algorithm of SR Research (Tatler, 2007) including a minimum saccade duration of 4 ms, a minimum fixation duration of 50 ms and a minimum velocity of 30°s-1. Fixation frequency was defined as the accumulation of fixations whereas fixation duration was defined as the mean time period of all fixations. The display was divided into four equal AOIs (Fig. 2.3). Gaze shifts between these AOIs were accumulated and stood for the variable number of gaze shifts.

An additional dependent eye movement variable was the number of blinks. Blinks were detected with the Event Detector for High Speed Event Detection provided by the iView software (SensoMotoric Instruments, Teltow, Germany). Blinks were recognized as an important indicator for cognitive processing. In general, blinks were usually inhibited during the performance of a task, for example during the decision process (Boehm-Davis, Gray, & Schoelles, 2016). However, the number of blinks depended also on the task characteristics. If the task was highly cognitively demanding, the number of blinks was expected to be lower. Also, the number of blinks depended on the state of the task performer, for example, if a person was fatigue (higher blink rate expected) or paid attention (lower blink rate expected; Stern, Boyer, & Schroeder, 1994; Wascher et al., 2015).

Further dependent behavioral variables were: judgment time, reaction time, task performance, subjective probability concept and judgment type. Judgment time was described as the time interval from the beginning of the prediction until a decision was indicated by the participants viz. one of the arrow keys is pressed. In contrast, reaction time was defined as the time interval from the reappearance of the target object at one of the exits until the appropriate key press. Task performance was reflected by the number of correct predictions. The subjective probability concept of the participants was assessed by the Concept Awareness Questionnaire and described the object-exit associations participants were able to retrieve.

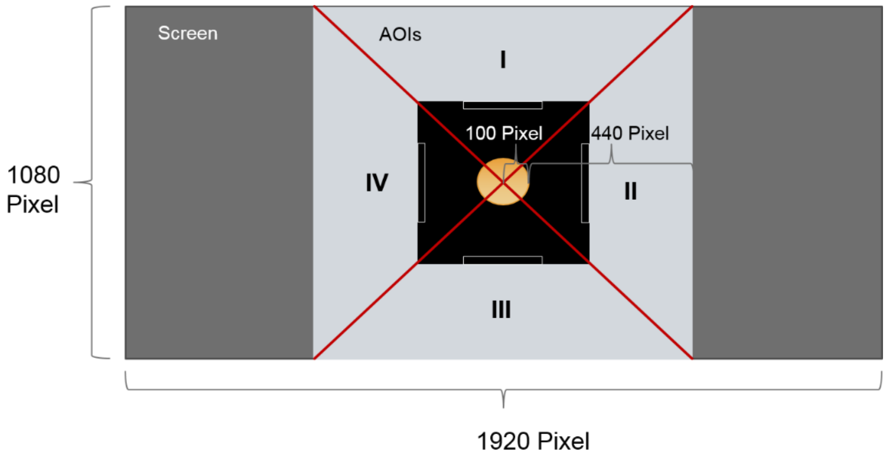


Figure 2.3: The four AOIs of the display: The red line marks the separation of the AOIs. Only eye movement data in the light gray area are analyzed in order to have equal sized AOIs.

2.7.3 Recurrence Quantification Analysis

Besides the aforementioned analyses, it was tested if recurrence quantification analyses might be an appropriate tool to gain deeper insights into the temporal development of eye movement behavior. The recurrence quantification analysis (RQA) was defined as "a method of nonlinear data analysis which quantifies the number and duration of recurrences of a dynamical system presented by its state space trajectory" (Marwan, 2017; Marwan, Romano, Thiel, & Kurths, 2007). Recurrence plots visualized at which point in time participants look at the same location as before by marking it with a black dot. Each fixation was recurrent with itself forming a diagonal line. The recurrence plot was mirrored around this diagonal line (Fig. 2.4). There were four important measures. First, *recurrence* (RR) was the percentage of refixations and thus showed how often participants refixated a location. Second, *determinism* (DET) described the refixations which form diagonal lines and thus indicated if a specific sequence of fixations was repeated. Third, *laminarity* (LAM) considered the percentage of refixations which form vertical lines. This measurement indicated if a distinct area was repeatedly refixated. Fourth, *center of recurrent mass* (CORM) was a measurement for the temporal distribution of recurrences. Low values indicated that refixations were made in short time intervals and high values indicated that refixations were made after relatively long time intervals (Anderson, Bischof, Laidlaw, Risko, & Kingstone, 2013). For example, the comparison of two recurrence plots might suggest that participants were more focused on distinct locations at the end of the experiment than in a trial at the beginning of the experiment as vertical lines were longer for the former trial. However, in total recurrences within a trial were quite rare in the current experiments. This might be due to the phenomenon that people tended to avoid to look at the same location which they attended shortly before called "inhibition of return" (IOR, Klein, 2000; Posner, 1980). Further, the short sequence of the trial did not allow much visual search behavior. The target object was the only relevant information participants were searching for and disappears during the prediction task. Thus, the visual search was mainly limited to the end of the trial when the target object reappears.

In a further step, we tested if cross recurrence plots are more conclusive than recurrence plots for the current experiments. They depicted the interrelations between two state space trajectories which described at which point in time the same location was attended while comparing two scanpaths. For example, a picture was scanned for 10s at an earlier and at a later point in time. This eye movement activity resulted in two scanpaths for the same field of observation which can be compared. The associated cross recurrence analysis (CRQA) was used to analyze the interrelations mathematically and was based on the same measurements as the recurrence quantification analysis reported before (Anderson et al., 2013). The idea was to directly compare two scanpaths with each other instead of comparing two separated plots as before. However, in the current experiment the length of the trials was different due to different prediction and reaction times and thus had to be standardized to make them comparable. Another problem was the large amount of trials that were recorded which made the cross-recurrence analysis rather complex and difficult to manage. Further, it was only reasonable to make a within analysis and not between subjects as the presentation of the target objects was randomized in every block. Finally, it was difficult to interpret the results in an appropriate manner. In conclusion, the RQA and CRQA did not seem to be an appropriate tool to analyze data of the current experiments as these analyses did not fit well to the experimental design and did not allow to draw conclusions about behavioral data.

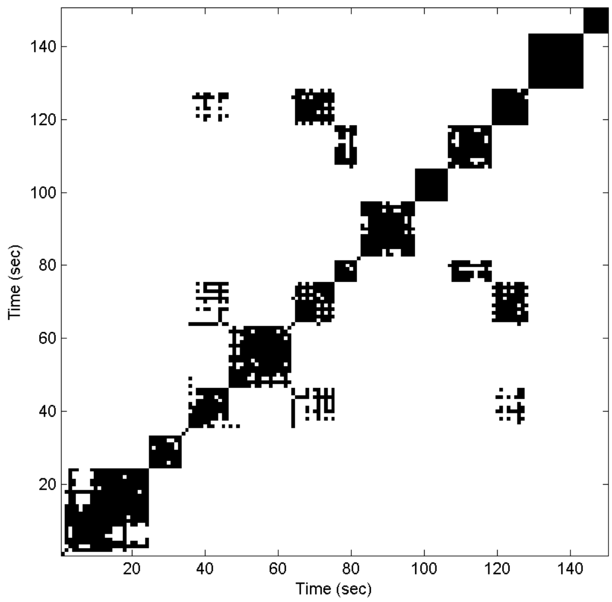


Figure 2.4: Example of a recurrence plot. For more details see text above.

2.7.4 Statistical Methods

Generally, a two-way repeated-measures ANOVA with the within-subject factors *block* (1-4) and *judgment* (correct, incorrect) was calculated with SPSS Statistics 23.0 (IBM, Armonk, USA) showing the main effects as well as the interaction and are extended depending on the research question. Means and standard deviations for correctly and/or incorrectly predicted trials of marginal or partly significant effects across blocks were listed in the Appendix.

A correlation analysis controlled motivation and attention effects on performance as it was assumed that increasing task performance was mainly dependent on the mental model development. In order to gain deeper insights into the effects of individual performance and motivation on the mental model development and visual search behavior, participants were split in two groups by hierarchical cluster-analysis. Planned *t*-tests were used to compare these groups and to detect significant group differences. The probability level for statistical significance was always set to α=.05. Before using planned *t*-tests, variables were tested for normal distribution by using the Shapiro Wilk test (p>.05) because this test is appropriate for small sample sizes. Whenever necessary, violations of sphericity were corrected by using the Greenhouse-Geisser Epsilon or using the Huynh-Feldt ε if Greenhouse-Geisser Epsilon is larger than .750. Unless otherwise specified, error bars in the graphics showed confidence intervals according to Morey (2008) that were calculated with a SPSS Syntax provided by O'Brien and Cousineau (2014). This was an appropriate method to calculate confidence intervals for within-subject designs. Further, Morey's approach was a corrected version of the earlier methods by Cousineau (2005) and Loftus and Masson (1994).

3 Experiment I – Predicting the Appearance of Geometric Objects

3.1 Introduction

The aim of Experiment I was to test if eye movement patterns inform about the acquisition of mental models under uncertainty while performing the OVSST. In the following, relevant literature on decision making, visual attention, eye movements, learning, feedback mechanism, mental models and uncertainty was highlighted specifically in the scope of Experiment I.

Decision making: The process of decision making can be separated into 4 stages as illustrated by Lee et al. (2017, p. 209): (1) the acquisition and integration of information, (2) interpretation and assessment of information, (3) planning and choice of one action, (4) monitoring and correction of the chosen action. According to the authors attentional resources as well as metacognition are necessary in all of these stages, especially within the decision-making cycle. Attention provides the starting point for every decision and can also influence the decision, for example, by the attention span and the level of attention (Lee et al., 2017, 209 ff.).

Eye movement and attention: Eye movements might provide insights into the visual attention allocation as both processes, eye movements and attention, share the same neural resources with attention as mentioned earlier (Kristjánsson, 2011). The operationalization of attention by eye movement parameters might also enlighten the influence of visual attention on decision making under uncertainty. For instance, Brunyé and Gardony (2017) investigated eye movements during different levels of perceptual uncertainty by using a perceptual decision making task. Results showed that eye movements reflect decision uncertainty as well as attention demands in a robust manner. For example, fixations were fewer and longer during conditions with uncertainty. Just and Carpenter (1976) found that longer fixation durations indicated difficulties in encoding information or a higher engagement during the performance of simple cognitive tasks. Also applied studies in HCI already showed that eye movement parameters are associated with attention processes as mentioned earlier (Ehmke & Wilson, 2007; Poole & Ball, 2006). For example, Goldberg and Kotval (1999) reported higher numbers of overall fixations and longer scanpath length while interacting with poor interfaces indicating less efficient search. The authors concluded that eye movements are an important indicator for the quality of the interface and thus can be useful for improving interfaces. In general, there seems to be an interaction between attention processes and the technical system that can be influenced.

Learning and feedback: Eye movements might not only inform about attention processes during decision making but also about learning processes which occur if the decision making is repeated. Learning can thereby include information about the location (Theeuwes et al., 2009) which is highly relevant for the performance of the OVSST. In general, learning results in faster and improved task performance due to more automatic and more experienced actions. Feedback mechanisms during learning are comparable with the aforementioned fourth stage of monitoring and correction within the decision making cycle mentioned by Lee et al. (2017). The comparison of the initial objective with the current situation is essential for an accurate learning process and thus influences behavior. Learning is also dependent on the frequency of use. As Thorndike states in the law of exercise and the law of effect (Thorndike, 1927), stimulus-response associations are reinforced if they are often used and produce satisfying effects. Accordingly, these associations are weakened if they are not used and associated with dissatisfying results (Olson & Hergenhahn, 2013).

Mental model development under uncertainty: The information which is processed during learning is finally stored in long-term memory and provides the basis for mental models. For example, if users interact with a technical system for a long period, they might have access to a detailed representation of the system and thus can deal with different situations. In turn, if users interact with the system for the first time, they probably have only little knowledge about the cause-effect relationship. At first, they have to develop a mental representation or mental model of the system which serves as the basis for the interaction with the system. The established mental model of the complex system is only a reduced representation of the reality. Altogether, the complexity of the system and the reduced representation lead to a certain degree of uncertainty users have to deal with. Uncertainty might also interfere with the mental model development. On the one hand, it depends on the user's personality how they deal with the uncertainty. Uncertainty can be used as a motivating or as an inhibiting factor (Smithson, 2009). Every person perceives and reacts to uncertainty in a different way. On the other hand, earlier mentioned coping strategies, including, for example, information search and suppression, have an influence on handling uncertainty (Lipshitz & Strauss, 1997).

However, optimal search strategies may only be a prerequisite to enhance the accuracy of mental models. At least two additional requirements have to be fulfilled for the development of mental models. First, users have to pay attention to the relevant stimuli to encode visual information (Mulligan, 1998) as already indicated by attention processes in the decision making cycle (Lee et al., 2017) at the beginning of the chapter, viz. eye movements are focused on the relevant stimulus information. Second, sufficient capacity has to be provided in the visual working memory to store the relevant information (e.g., Brady, Konkle, & Alvarez, 2011).

3.1.1 Research Question and Hypotheses

In Experiment I, we addressed the acquisition of mental representations under uncertainty[2]. For this attempt, the OVSST was employed and eye movements assessed. As mentioned in the introduction, uncertainty was understood in the sense of task uncertainty and defined as a lack of knowledge about the cause-effect relationship (Thompson, 1967) which encompasses both, the objective uncertainty evoked by the task characteristics and the perceived subjective uncertainty by the participants. The objective uncertainty is initiated by the probability structure of the OVSST. The probabilities were estimated by the participants at the end of the experiment via questionnaire.

Participants have to learn object-exit associations and outcome feedback to develop an accurate mental representation of the OVSST. It was assumed that this mental representation is then used to perform the prediction task of the OVSST as accurately as possible and to anticipate the correct exit for the reaction task. The acquisition process might be three-staged as suggested by Schumacher and Czerwinski (2014). As already mentioned earlier they divided the acquisition of the mental representation into three stages: 1) a reference to earlier experience, 2) a first understanding of the causal relationship and 3) an abstract and accurate representation (Schumacher & Czerwinski, 2014). Because none of the participants was familiar with the OVSST, one would assume that the pre-theoretic stage, in reference to Schumacher and Czerwinski (2014), does not focus on the retrieval of similar examples in memory, but rather on the accumulation of new information in this task. Hence, all participants start with the same degree of uncertainty and develop a completely new mental representation of the task. The acquisition of the mental representations already starts with informing the participants about the task via instructions and performing the training trials.

The acquisition of mental models as a process of information processing and storage is also combined with learning mechanisms. As already mentioned in the introduction, Rieskamp and Otto (2006) found stronger learning effects at the beginning of the learning process than at the end. They reported several alternative reinforcement learning theories and also, for example, the imagination model. This model describes the ability of people to imagine different options and their outcomes comparable with their mental representations. It predicts that learning processes are accelerated at the beginning of the task due to the reinforcement of also non-preferred strategies. This model might also be applicable to the OVSST as task characteristics are comparable. The OVSST contains different options and outcomes which are stored in mental representations of the task.

[2] Parts of Experiment I are already published (Renker & Rinkenauer, 2016).

In Experiment I, we set out to investigate if eye movements allow insights into the acquisition of mental models under uncertainty during the performance of OVSST. It was expected that task performance increases over blocks viz. the amount of accurate predictions increases as indicator of a learning process/ learning effect. This increase is expected to be highest at the beginning of the experiment when information is mainly accumulated (cf. Rieskamp & Otto, 2006). Furthermore, judgment times as well as reaction times are expected to decrease over blocks due to increasing practice and increasing task performance (cf. Jacob & Hochstein, 2009). In parallel, it is hypothesized that eye movement activity decreases over blocks reflecting the state of learning and the degree of uncertainty (cf. Lipshitz & Strauss, 1997). Reactions times as well as eye movements are dependent on the correctness of the judgment. Thus, it is expected that reaction times are faster during correctly than incorrectly predicted trials and eye movement patterns are less extensive (lower values) during correctly than incorrectly predicted trials. Finally, it was expected that the number of blinks is lower during incorrectly predicted trials due to a stronger need for attentional resources which are required to perceive the target object.

3.2 Method

3.2.1 Participants

In total, 20 participants (11 female) participated in the experiment at the Leibniz Research Centre for Working Environments and Human Factors (IfADo). As mentioned in the introduction, participants had to finish the Concept Awareness Questionnaire at the end of the experiment to gain insights into the participants' mental representations of the probability structure. On the basis of this test, two of the participants were excluded from data analysis, because they were not able to develop the expected representations. The mean age of the remaining 18 participants was 25 years (*SD*=3 years). 15 of the remaining participants were students from the Technical University Dortmund. All of them were right-handed and had normal vision without glasses or contact lenses.

3.2.2 Procedure

In Experiment I, we applied the procedure as described in the General Method section (Chapter 2) to compare objective data (eye movement data and behavioral data) with subjective data (data retrieval, i.e. subjective probability concept). Motivation and attention resources were used as controlled variables measured by the D2 test of attention (Brickenkamp, 1994) and the QCM (Rheinberg et al., 2001) to check for confounding effects on task performance. The D2 test was run before and after the experiment to determine the influence of the OVSST on the test results and the QCM was completed at the end of the experiment. Participants performed a practice block of the OVSST and two blocks of the 100% condition with geometric objects that differed from the experimental ones. In this condition, there is no uncertainty provided by the probabilities due to the constant allocation of the target objects to the exits. After the training and the 100% condition, the original OVSST with the 74-11-11 probability concept was completed. Finally, participants filled in the Concept Awareness Questionnaire to gain insights into their subjective probability concept.

3.2.3 Data Analysis

In the 100% condition, participants predicted 60.8% of the trials (26 out of 42 trials) correctly. This condition was used to control for outliers in the sample concerning the comprehension of the task after the control sessions. The box plot in figure 3.1 showed no outliers regarding the task performance of the current sample in the 100% condition. Nevertheless, there seemed to be a high interindividual variability indicated by the high range of the dependent variable.

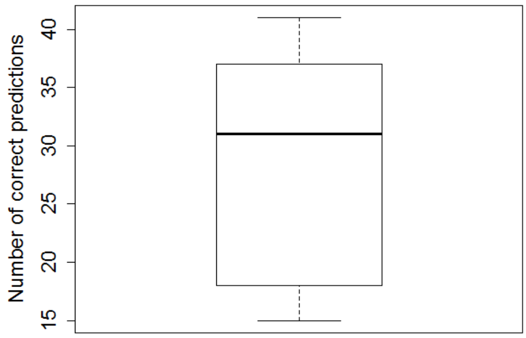


Figure 3.1: Box plot for the number of correct predictions in the 100% condition of Experiment I.

Data of the experimental condition were adjusted as follows: Heat maps were used to check for drifts in eye movement data set and thus, to ensure that data recording was accurate. 1.3% of the trials were excluded from data analysis due to missing responses to the judgment task. Another 0.3% of the trials were excluded, because less than 65% of the eye movement data in the trial were not available due to tracking errors. As mentioned in the introduction a two-way repeated measures ANOVA with the within-subject factors *block* (1-4) and *judgment* (correct, incorrect) was performed to analyze the development over time to evaluate the learning process. Dependent variables were eye movement parameters: fixation frequency, fixation duration, number of gaze shifts and number of blinks. Further dependent variables were judgment time, reaction time and task performance. Task performance was defined as the number of correct predictions. Judgment time was defined as the time interval from the beginning of the prediction - as soon as the object disappeared - until the key press indicating the prediction. Reaction time was defined as the time interval from the reappearance of the target object at one of the exit until the appropriate key press.

For the analysis of the participants' subjective probability concept, data of the Concept Awareness Questionnaire were compared with the objective probability concept (74-11-11) indicating the accuracy of the developed mental representation. Response behavior, i.e. the prediction of the exits, was also cumulated for each object-exit association and compared with the subjective probability concept to determine the relation between behavioral and mental processes. Further, the influence of the rare occurrence (trials with object appearance at the entrance) on response behavior was tested by comparing response behavior before and after the rare occurrence. The impact of the confounding variables on task performance, the number of correct predictions, was measured by correlation analysis and repeated measures ANOVA.

Additional analyses were run due to the novelty of the OVSST in Experiment I. Eye movement data within the trials were analyzed as well. Two time intervals were determined: First, the judgment-interval between the appearance of the object at the bottom entrance and the key press whereby participants predict the exit and second, the reaction-interval after predicting the exit until the object stays in this final position at the exit (Fig. 3.2). Paired *t*-tests were computed to identify differences between eye movement patterns in both intervals, also including the dwell time: time participants spend in the areas of interest (AOI). AOIs were divided into AOI_{target} where the target object finally reappears, $AOI_{predict}$ which exit-zone participants have predicted and unassigned AOIs, called AOI_{other}, for the within trial analysis. In addition, the last fixation within the first interval was compared with the participant's prediction.

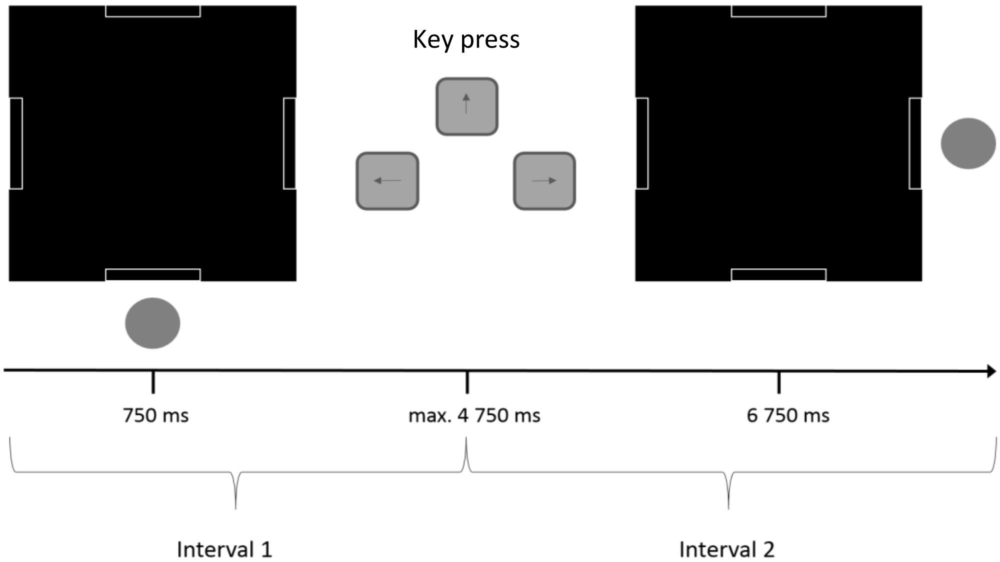


Figure 3.2: Trial analysis separated into two intervals: (1) From object appearance at the bottom entrance until key press and (2) from key press until final position at the exit.

Blinks were separately analyzed within the trial because an effect directly after the prediction and reaction task would be suggested based on the earlier discussed literature (see General Method for details). The blink rate for the time course of the trial (blinks for each 100ms) was investigated to have a more detailed analysis.

Hierarchical cluster analysis of the task performance was used to split participants into two groups (high and low performers). Group differences with regard to eye movement patterns were assessed by using independent *t*-tests. Before using *t*-tests, variables were tested for normal distribution by using the Shapiro Wilk test. This test was used due to the small sample size (n=18).

Averaged predictions for correct and incorrect exits were fitted to two learning models, a power model and an exponential model. These models are used to estimate the rate of learning and to predict the maximal learning performance by the asymptote of the learning curve. The power model and the exponential model were chosen for the analysis because these two models are mainly discussed in literature (e.g., Heathcote, Brown, & Mewhort, 2000; Newell & Rosenbloom, 2013; Speelman & Kirsner, 2008). In literature, the exponential function dominates the power function for large sample sizes. However, Newell and Rosenbloom (2013) found that the power function describes learning data consistently better. The main differences of the two models is that the exponential function has a fixed based raising to a variable exponent, whereas power functions have a variable base raising to a fixed exponent. We tested both models to check which one fits best for the current data set.

3.3 Results

In the following, only significant (p<.05) results or trends (p<.10) were reported, except if the results were relevant for the aforementioned research questions.

Task performance:

In total 65.7% of all trials were correctly predicted. Descriptive statistics showed that the number of correct predictions increased across blocks, indicating the learning process (for details see Appendix C, Tab. 9.1). However, this increase was significant only from Block 1 to Block 2, $t(16)=4.16$, $p=.001$. Thus, Block 1 seems to be the most important one for the learning process and therefore a further analysis was run by splitting the block into 4 equal parts. Within the first 20 trials participants already predicted the likely exits in 63.7% of the cases. In the last part of Block 1 participants chose in 92.5% of the cases the likely exits.

Judgment time:

Results showed that judgment time decreased significantly across blocks, $F(3,51)=11.61$, $p<.001$, $\eta_p^2=0.285$, and was significantly shorter during correctly than incorrectly predicted trials, $F(1,17)=6.77$, $p<.019$, $\eta_p^2=0.406$, (Fig. 3.4A). The variable judgment time was analyzed in detail to gain deeper insights into the cause-effect relationship of the within-subject factor *judgment*. Therefore, judgment times were split into correct and incorrect judgments for likely and unlikely exits (see Fig. 3.3). Correct judgments for unlikely exits could not be analyzed because participants mainly predicted the likely exit and if they chose the unlikely exits, the probability is low that the judgment is correct. Thus, there were not enough data to integrate this combination in the analysis. For all other combinations, results showed that participants predicted the likely exits faster than the unlikely exits no matter if the prediction was correct or incorrect (likely exit: correct prediction *M*=.387, *SD*=.142; incorrect prediction *M*=.386, *SD*=.143; unlikely exits: incorrect prediction *M*=.455, *SD*=.195). These effects were significant for likely correct predictions, $t(19)=2.57$, $p=.03$, and showed a trend for likely incorrect predictions, $t(10)=1.94$, $p=.08$, when comparing it with judgment times for unlikely incorrect predictions.

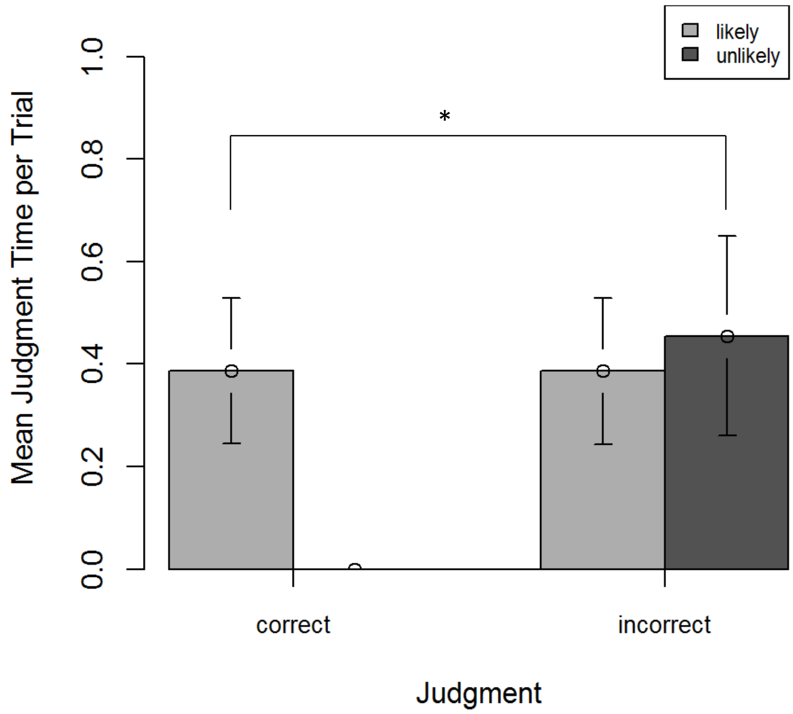


Figure 3.3: Judgment times in Experiment I: Correct and incorrect judgments for likely or unlikely exits. The bar for correct judgments and unlikely exits was not depicted as this combination was rare (n=3) and thus, not valid for a comparison. Error bars depict the standard deviation.

Reaction time:

We observed a trend of *block*, $F(3,51)=2.54$, $p=.067$, $\eta_p^2=0.130$, indicating that reaction times decreased across blocks, which was underlined by descriptive statistics (see Appendix C, Tab. 9.2 for details). In addition, we found a main effect of *judgment*, $F(1,17)=80.45$, $p<.001$, $\eta_p^2=0.826$, indicating shorter reaction times during correctly than incorrectly predicted trials (Fig. 3.4B).

Fixation frequency:

The main effect of *block* was significant, $F(3,51)=5.32$, $p=.012$, $\eta_p^2=0.238$, which suggested a decrease of fixation frequency across blocks. We also found a main effect of *judgment*, $F(1,17)=23.69$, $p<.001$, $\eta_p^2=0.582$, indicating a lower fixation frequency for correctly predicted than for incorrectly predicted trials (Fig. 3.4C).

Fixation duration:

Analysis of *block* revealed a significant main effect, $F(3,51)=4.00$, $p=.045$, $\eta_p^2=0.190$, suggesting a significant decrease of fixation duration across blocks (Fig. 3.4D). However, we observed no significant main effect of *judgment*, $F(1,17)=2.12$, $p=.164$, $\eta_p^2=0.111$, in contrast to fixation frequency and the number of gaze shifts which both showed similar patterns of results.

Number of gaze shifts:

The main effect of *block* was significant, $F(3,51)=11.16$, $p<.001$, $\eta_p^2=0.396$, indicating that the number of gaze shifts decreased across blocks. There was also a main effect of *judgment*, $F(1,17)=37.86$, $p<.001$, $\eta_p^2=0.690$, indicating that the number of gaze shifts was smaller for correctly predicted than for incorrectly predicted trials (Fig. 3.4E).

Number of blinks:

A main effect of *block* was found, $F(3,48)=4.820$, $p=.018$, $\eta_p^2=0.232$, suggesting a significant increase of the number of blinks across blocks in contrast to all other eye movement parameters (Fig. 3.4F). In addition to the main effect, results showed a trend for a hybrid interaction between the factors *block* and *judgment* of the dependent variable number of blinks, $F(3,48)=2.784$, $p=.058$, $\eta_p^2=0.148$, (Fig. 3.3F; for details see Appendix C, Tab. 9.1). The hybrid interaction was indicated by a higher number of blinks for correctly predicted than for incorrectly predicted trials in Block 1 and Block 2, however, this pattern changed and was reserved in Block 3 and Block 4.

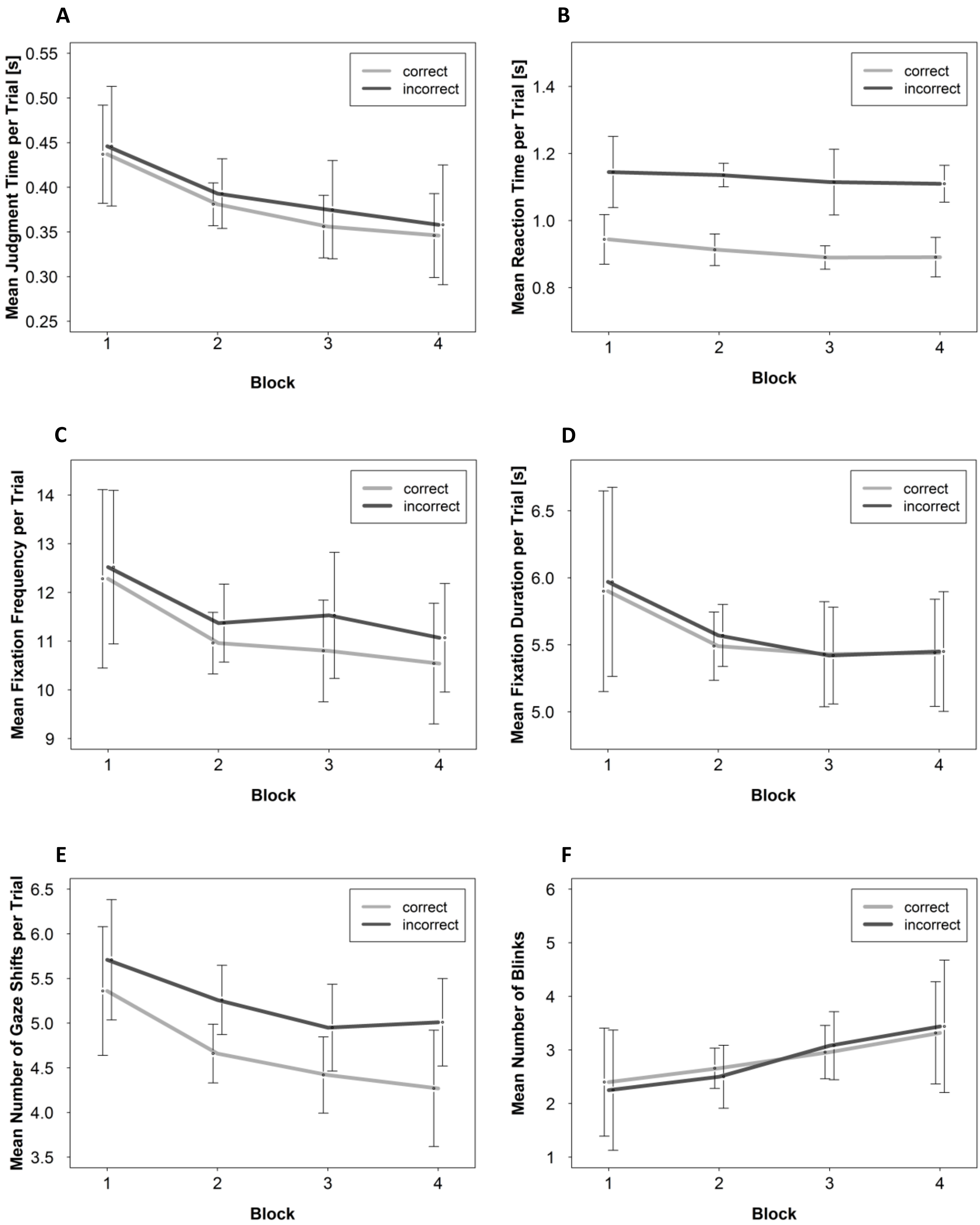


Figure 3.4: Times course of the variables in Experiment I: Judgment time (A), reaction time (B), fixation frequency (C), fixation duration (D), number of gaze shifts (E)and number of blinks (F) across blocks for correctly and incorrectly predicted trials as function of *block* and *judgment*.

Analysis of the subjective probability concept and response behavior:

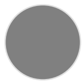
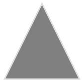
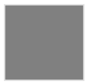
In general, the Concept Awareness Questionnaire showed that estimated probabilities of the participant's subjective probability concept were close to the given probabilities as shown in table 3.1. Participants estimated the probability relation of the object to the exits on average with 15.5% for the unlikely exits and 67% for the more likely exit. In comparison with the given probability concept of 74% for the likely exit and 11% for each of the two unlikely exits, the results of descriptive statistics show that participants underestimated higher probabilities and overestimated lower probabilities. However, planned *t*-test showed only a trend for an underestimation ($p<.06$) and only three of the unlikely object-exit associations were significantly overestimated, $t(16)=2.82$, $p=.006$ (circle-left), $t(16)=2.21$, $p=.021$ (square-top), $t(16)=2.78$, $p=.007$ (square-right). All other unlikely object-exit associations were not significantly overestimated, however, results suggested a trend for an overestimation ($p<.06$).

In contrast to the estimation of the probabilities, behavioral data rather showed a reversed pattern. Table 3.1 shows that likely exits were more often predicted by the participants than the object actually reappeared at this position. However, interindividual variability of the predictions were large, ranging from 72.14% to 98.66% for the likely exits ($M=90.66$; $SD=7.41$) and from 1.61% to 27.86% for the unlikely exits ($M=9.34$; $SD=7.41$).

Figure 3.5 shows a more detailed visualization of the predictions for every of the three objects per exit and block. This figure provides an overview of the development in the course of time. Participants obviously learned the associations of the objects to the exits already within the first block and intensified their behavior over time according to the probability concept. Thus, the likely exit was predicted more often whereas the unlikely exit was less often predicted with regard to the objective probability.

In order to check if the rare occurrence (the object reappears at the bottom entrance) had an influence on the participants' usage of the developed response strategy and thus, on the learning process additional analyses were run. The response during the rare occurrence was compared with the next trial after the rare occurrence presenting the same object. The results show that in 82% of the cases the response strategies did not change after the rare occurrence and that the response behavior still matched with the optimal response strategy participants used most of the time.

Table 3.1: Memory representation of the probability concept and behavioral probabilities in Experiment I

Object & Exit		Subjective Probability Concept	Performed Predictions
	left	16 % (6.4)	7 % (5.9)
	top	**71 % (11.6)**	**89 % (7.8)**
	right	13 % (6.0)	4 % (3.4)
	left	16 % (10.9)	5 % (6.0)
	top	14 % (6.0)	8 % (3.5)
	right	**70 % (14.9)**	**87 % (8.9)**
	**left**	**66 % (17.1)**	**87 % (7.8)**
	top	18 % (13.0)	9 % (4.5)
	right	16 % (8.7)	4 % (4.8)

Note. The object-exit associations that inhere a higher probability are shown in bold. Values in brackets show the standard deviation.

<u>*Analysis of the control variables:*</u>

As already mentioned, the D2 test was used to assess attention before and after the experiment. Participants selected significantly more targets of the D2 test after the experiment than before, $t(16)=6.44$, $p<.001$. In the pre-test they detected 71% (SD=10.31%) of the cases on average and in the post-test 79% (SD=11.43%) of the cases on average, presumably due to a learning effect. The error rate did not change from the pre-test to the post-test, $t(16)=0.22$, $p=.829$. Thus, attention did not seem to decrease from the beginning to the end of the OVSST.

The analysis of interest in the task showed a significant negative correlation (r=-.469, p=.025) between the subscale *interest* of the QCM and the number of correct predictions, by using the Spearman's rank correlation coefficient. Thus, to be interested in the task seemed to be negatively related to task performance.

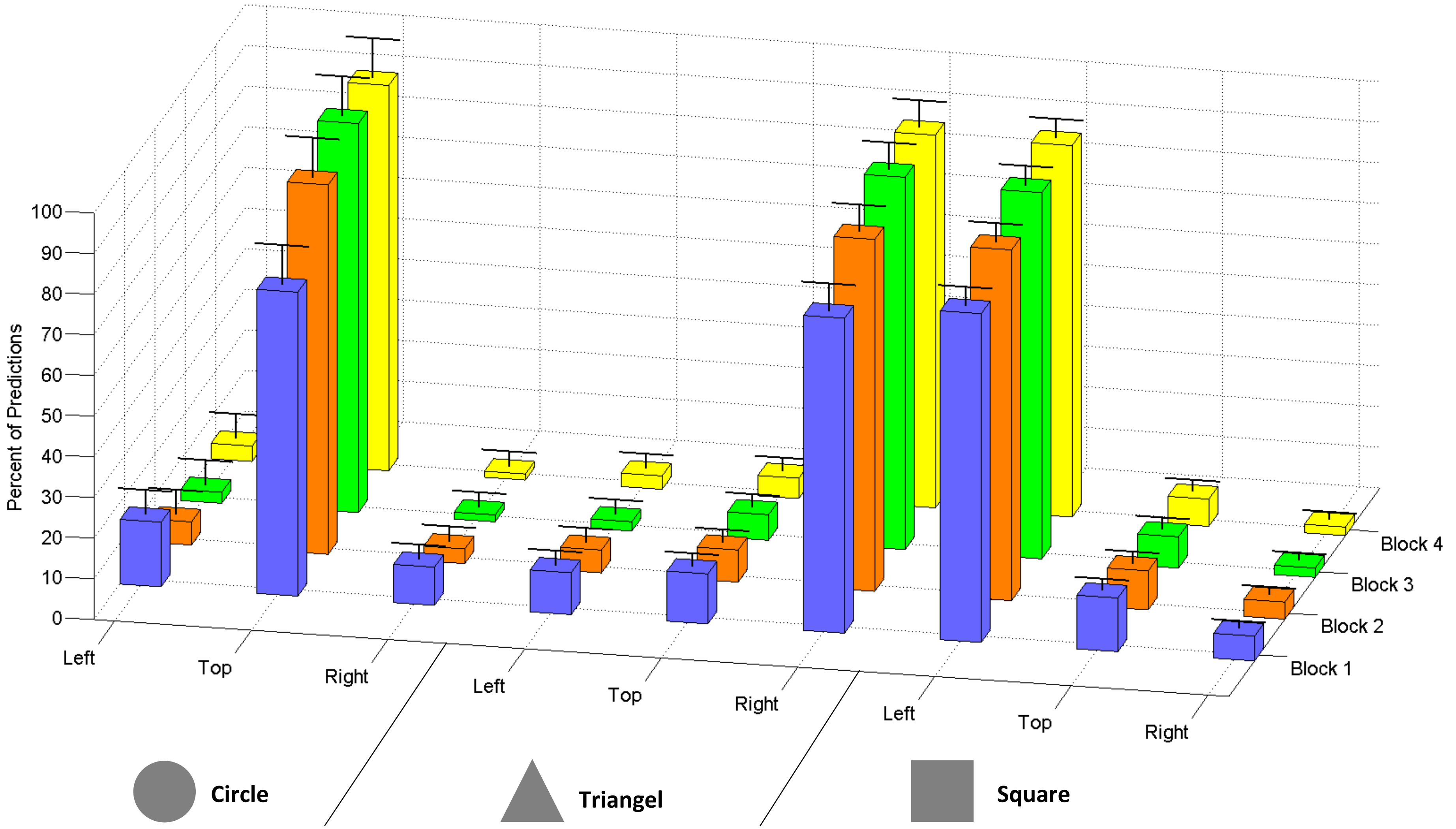


Figure 3.5: Response frequencies in Experiment I: Response frequencies of the three objects (circle, triangle, square) per exit (left, top, right) and block (1-4)

<u>*Within trial analysis:*</u>

Table 3.2 shows the main results of the within trial analysis. There were significant differences between visual search activity in the first judgment-interval and the second reaction-interval. Participants' gaze was mainly located in the AOI_{bottom} during the judgment-interval until the key press, presumably due to the initial presence of the stimuli in this AOI. In the rare case that participants moved their eyes into another AOI within the judgment-interval, they preferred the AOI_{target} instead of the AOI_{other}, $t(17)=2.552$, $p=.021$. Appropriately, last fixations in the judgment-interval were mainly in the AOI_{bottom}, namely 78.5%. However, 17.8% of last fixations in the judgment-interval were in the $AOI_{predict}$. This AOI is the predicted target AOI and thus indicated with the appropriate key press at the end of the judgment-interval in the context of the prediction task. Thus, if participants showed such a visual search behavior already in the judgment interval, then they most likely anticipated or prepared their later prediction. In this case, eye movements indicated the final choice of the target exit before actually predicting the exit. Generally, participants showed extensive visual search behavior only in the reaction-interval with significant longer dwell time spent in AOI_{left}, AOI_{top} and AOI_{right}. Accordingly, the gaze shifted more often between the AOIs, presumably due to the fact that participants had to detect the target object to perform the reaction task.

Table 3.2: Trial analysis in Experiment I: Statistics between interval 1 and interval 2

	judgment-interval	reaction-interval	paired *t*-tests
Dwell time AOI_{bottom}	$M=1.913$, $SD=.182$	$M=.251$, $SD=.143$	$t(17)=36.49$, p<.000
Dwell time AOI_{left}	$M=.040$, $SD=.047$	$M=.803$, $SD=.245$	$t(17)=13.76$, p<.000
Dwell time AOI_{top}	$M=.071$, $SD=.063$	$M=1.311$, $SD=.223$	$t(17)=23.60$, p<.000
Dwell time AOI_{right}	$M=.038$, $SD=.042$	$M=.819$, $SD=.187$	$t(17)=17.73$, p<.000
Number of Gaze Shifts	$M=.499$, $SD=.384$	$M=2.723$, $SD=1.342$	$t(17)=7.10$, p<.000
Fixation Frequency AOI_{target}	$M=.330$, $SD=.270$	$M=3.441$, $SD=1.133$	$t(17)=12.99$, p<.000
Fixation Frequency $AOI_{predict}$	$M=.354$, $SD=.358$	$M=3.298$, $SD=1.171$	$t(17)=12.05$, p<.000
Fixation Frequency AOI_{other}	$M=.220$, $SD=.137$	$M=1.353$, $SD=.927$	$t(17)=5.41$, p<.000

Analysis of the blinks:

The number of blinks were analyzed in detail by means of frequency measures and planned *t*-tests of events or actions and peak values visualized in figure 3.6. This figure shows a higher blink rate per second after the three main events: (1) initial appearance of the object, (2) prediction of the exit and (3) detection of a possible change in the color intensity. The number of blinks of all participants was accumulated and split into intervals of 100ms length. When the object shape appeared the first time the blink rate increased from 90 blinks/s (interval: 700-800ms) to 430 blinks/s (interval: 1100-1200ms). The increase was highest and most rapid after performing the prediction task increasing from 120 blinks/s (interval: 2600-2700ms) to 640 blinks/s (interval: 2900-3000ms) without a long time delay. There was more variation when the object reappeared and participants had to react. While the object moved out of the room the blink rate first decreased down to 30 blinks/s (interval: 5000-5100ms) and then increased with a time delay slowly to a peak of 400 blinks/s (interval: 5900-6000ms). The time delay presumably occurred due to the fact that the object became gradually visible. It was impossible to perceive the object immediately and to react and thus, there was a larger scattering.

Cluster analysis:

A hierarchical cluster analysis was performed to group participants into performance clusters and thus to investigate the effect of individual performances on eye movement pattern as described in the introduction. Clusters were formed on the basis of the number of correct predictions, i.e. low performers (n=6) and high performers (n=12). Low performers had a significantly higher fixation frequency, $t(16)=3.39$, $p=.004$, and showed significantly more gaze shifts, $t(16)=3.06$, $p=.025$, than high performers (see Tab. 3.3). In contrast, fixation duration did not differentiate between the clusters, $t(16)=0.35$, $p=.732$.

Table 3.3. Descriptive statistics of the Performance Cluster in Experiment I

	Cluster	Mean	Standard Deviation
Number of Gaze Shifts	Low performance	6.53	2.17
	High performance	3.75	0.71
Fixation Frequency	Low performance	14.89	5.03
	High performance	9.27	2.12

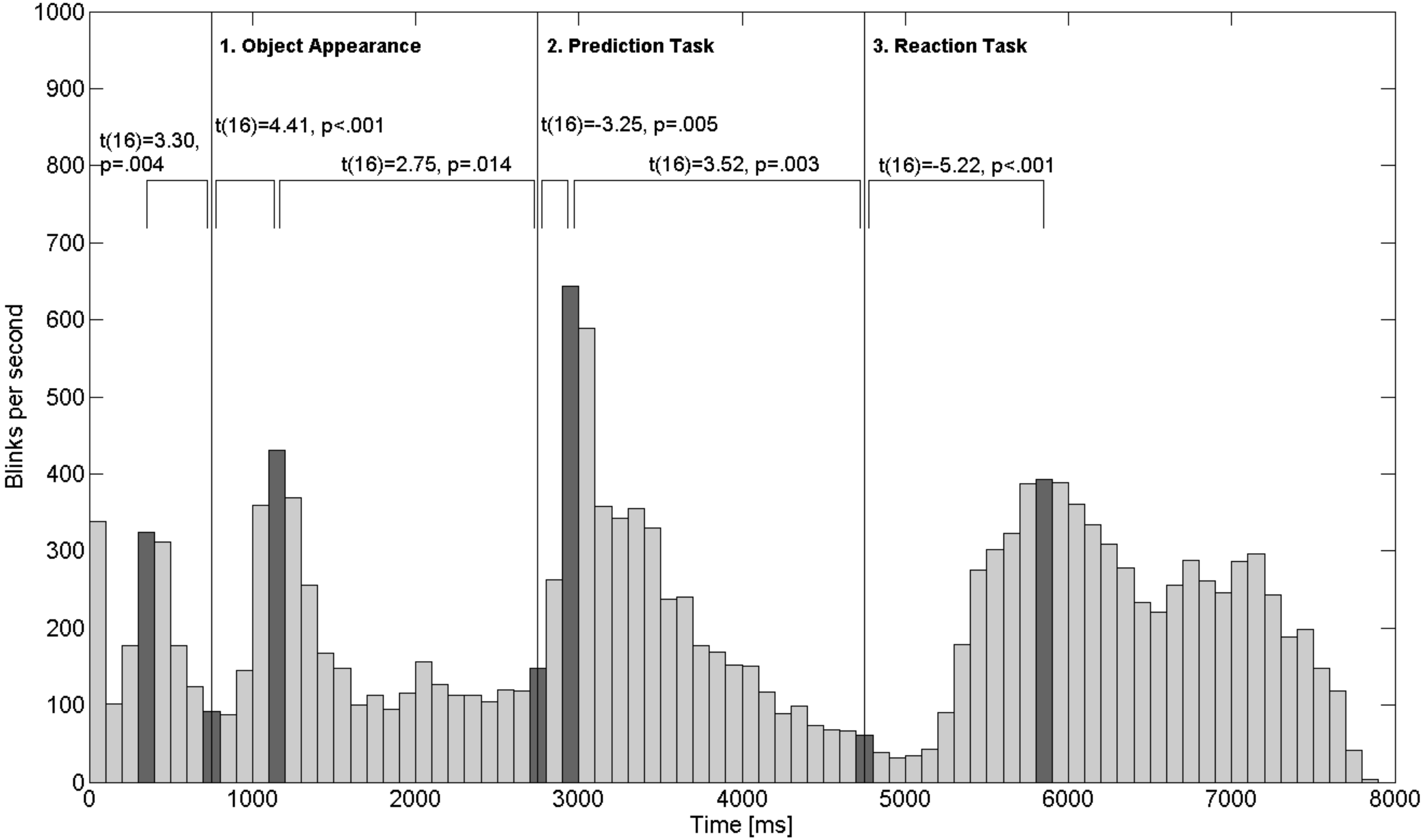


Figure 3.6: Blinks in Experiment I: The number of blinks per second accumulated over all participants during the time course of the trial for correctly and I incorrectly predicted trials. One bar displays all blinks within 100ms. Dark gray bars highlight the values for the *t*-test (event or action and peak values). The first vertical grey line marks the beginning of the object appearance (1), the second line marks the beginning of the prediction task (2) and the third line marks the beginning of the reaction task (3) from left to right.

Learning models:

Learning models were fitted to get a future prognosis with regard to the response strategy (e.g. TTB) of the participants. Data of the fitted learning models showed that the power model revealed to be a better fit in total (R^2=.92) than the exponential model (R^2=.89), presumably due to the fact that several individuals were analyzed (Speelman & Kirsner, 2008, p. 18). The asymptote of the power function approaches 99.26% for the likely exits and 0.45% for the unlikely exits. Thus, the model fit suggests that the likely exit, the high probability, is obviously represented as nearly a 100% probability.

3.4 Discussion

The purpose of Experiment I was to explore eye movement patterns during the acquisition of mental representations under uncertainty. Participants improved their performance by learning the underlying probability concept of the OVSST. This learning process was reflected by a decrease of judgment times, reaction times and eye movement parameters (fixation frequency, fixation duration, number of gaze shifts). The development across blocks was strongest at the beginning of the experiment, presumably due to the fact that the main learning processes occurred in this time period indicating a strong coupling between learning, behavioral data and eye movement data. Additionally, the correct predictions were associated with faster judgment times and reaction times as well as fewer fixations and gaze shifts. Finally, participants were able to develop a quite accurate mental representation of the probability concept and seemed to use both strategies described in literature to cope with the given uncertain situation, i.e. information search as well as suppression of ambiguous information (Lipshitz & Strauss, 1997).

Decision making:

It was expected that participants improve their decision-making behavior by learning the object-exit associations of the OVSST. Indeed, means suggested that task performance increased across all blocks, however, the increase was only significant from Block 1 to block 2. Thus, learning effects seemed to be strongest from Block 1 to Block 2 corresponding to the findings of Rieskamp and Otto (2006) who found that learning processes concerning the strategy selection for a decision were accelerated at the beginning of the task sequence. A detailed analysis of Block 1 showed that the main tendencies of the probability structure were already learned in the first 20 trials of the experiment. Thus, participants internalized the probability structure quickly, presumably due to the fact that the chosen object-exit associations were easy to memorize.

The analysis of the response behavior showed that object-exit associations with higher probabilities were preferred and object-exit associations with lower probabilities were neglected. Further, the reappearance of the target object at the bottom entrance, the rare occurrence (4%), did not influence the response behavior. Thus, participants ignored ambiguous information and maintained their initial response strategy consistent with the coping strategy mentioned by Lipshitz and Strauss (1997).

The decision strategy used by the participants corresponded with the optimal response strategy Take The Best (TTB) to maximize the number of correct predictions (Dougherty et al., 2008), also known as probability maximizing. As described by TTB, participants chose the recognized alternative initially and with increasing expertise they chose the alternative with the best cue value to perform the task (see Introduction for more details). The usage of TTB was strongly corroborated by the fitting results of the power model showing the tendency that participants would hardly never predict exits with low probabilities and almost always exits with high probabilities in further blocks. Thus, participants seemed to ignore lower probabilities and to redefine the given uncertainty as certainty in their mental model.

The response behavior was consistent with those of Edwards (1961) studying probability learning. The author also found that participants used a probability maximizing strategy, i.e. they predicted more often specific events than they actual appeared. These findings of probability maximizing were contradictory to the robust phenomenon of probability matching which is well described in literature (e.g. Fantino & Esfandiari, 2002). This phenomenon states that humans usually try to match their choice to the given probabilities which eventually results into a less optimal strategy. Yu and Huang (2014) studied visual search behavior under conditions of uncertainty during matching and maximizing strategies in decision making. Participants had to identify random-dot stimulus patches whereby in one condition an equal probability and in a second condition a biased probability was tested. They found similar to the current findings an overmatching of the biased probabilities and a quick internalization of the probability concept. Further, they assumed that humans choose matching strategies to adapt to a changing world. Thus, the absence of an overall probability matching phenomenon in the current study might be due to the laboratory setting and the steadiness of the experimental setting.

Effects on cognitive processing:

The analysis of judgment times and reactions times showed a parallel development across blocks to the task performance. Participants performed the prediction task and reactions task considerably slower at the beginning of the task than at the end, probably due to the learning effects. Additionally, the decrease of judgment times over time might indicate some kind of proceduralization in decision making (Anderson, 1982; Pomerol & Brézillon, 2003; Yi & Davis, 2003), meaning that rules of the decision strategy embedded in the mental model were automatically retrieved from memory without any further cognitive processing. Such a rule could be, for example, to predict the right exit whenever the triangle appears.

Assessing the within-subject factor *judgment* clearly showed that reaction time was faster for correctly predicted trials, presumably due to a beneficiary preparation for the reaction task. Also, prediction time was faster during correct predictions, especially when the optimal decision strategy (TTB) was used, i.e. when the likely exit was chosen. The reason for this might be some kind of strategy switch cost (Lemaire & Lecacheur, 2010). As mentioned before, participants might have automated the learned response patterns for the optimal strategy, also called proceduralization, so that a deviation from this strategy required the inhibition of automatic processes leading to longer judgment times.

Visual search behavior:

Besides learning effects across blocks, it was expected that eye movement parameters reflect these processes. Indeed, eye movement patterns showed a function of *block* as well as task performance, judgment time and reaction time. They informed about the state of learning and the degree of task uncertainty by a decrease of visual search activity, i.e. fixation frequency, fixation duration and the number of gaze shifts decreased across blocks. The number of blinks showed a reversed pattern and increased across blocks indicating less attention effort across blocks (e.g. Wascher et al., 2015). In general, visual search for relevant information seemed to be more extensive at the beginning of the trial than at the end, presumably due to a high initial subjective uncertainty and the novelty of the information. Rayner (2009) stated that new information is only processed during fixations. Thus, information search seemed to help to cope with the initial subjective uncertainty evoked by the OVSST.

The analysis of *judgment* clearly showed that fixation frequency as well as the number of gaze shifts were lower for correctly predicted trials, presumably due to a beneficiary preparation for the reaction task. Conversely, this means that incorrect predictions might require more visual search activity to perceive the object at the missed exit to perform the reaction task. It was expected that the number of blinks is lower during incorrectly than correctly predicted trials as more attentional resources were required to perceive the target object and thus to perform the reaction task. This was true at the beginning of the experiment (see Fig. 3.4F) and in accordance with previous findings that a lower blink rate is related to higher visual attention and information processing (e.g., Wascher et al., 2015). However, reasons for a reversed pattern at the end of the experiment remain unclear.

The further analysis of blinks showed that blinks were suppressed before the execution of the prediction and reaction task suggesting the acquisition of higher attentional resources to perform the task. After the completion of the tasks, when cognitive demands are lower, number of blinks increased. Experiments with different visual choice response tasks already showed that blinks were suppressed during cognitive processing and in expectation of new information (Fogarty & Stern, 1989; Wascher et al., 2015). These results were also true for the OVSST of the current experiment. Thus, the blink rate seemed to provide a relevant indicator for attentional effort.

Mental model development:

Behavioral data as well as eye movement data clearly showed a learning process as expected. In addition, data of the Concept Awareness Questionnaire indicated that participants were able to develop a mental model that was close to the given probabilities of the OVSST. High probabilities were usually underestimated and low probabilities were overestimated presumably due to a tendency to the mean (Beuer-Krüssel & Krumpal, 2009). The reported results of the subjective probability concept in the current experiment showed similar tendencies. However, the objective response behavior showed rather a reversed pattern as participants rather predicted mainly higher probabilities and neglected lower probabilities. Participants might be able to reflect the overmatching of their response behavior and to adapt their conscious understanding of the probability concept according to the received feedback.

In addition, the high probability of 74% given by the OVSST might facilitate the accurate acquisition of the mental model as it points clearly into one direction. Probabilities and their associated meanings were already discussed by Jungermann et al. (2010, p. 161). They presented possible associations between numeric terms and the appropriate verbal terms within 25%-steps ranging from 0% associated with “never” to 100% associated with “always”. They assumed that 75% is associated with the word “often” and this might additionally work as an anchor for the subjective probability estimation. In general, it should be taken into account that a large interindividual variability in estimating the subjective probabilities occurred, also found in the earlier mentioned study by Yu and Huang (2014).

Taken together the findings of eye movement data and behavioral data over time suggest that the mental model of the OVSST was mainly developed in the first block. Thus, after the first block high performers might already be in the expert stage and were able to recognize patterns and to retrieve former knowledge for the development of an optimal decision strategy under uncertainty according to the ideas of mental model acquisition suggested by Schumacher and Czerwinski (2014).

Group-specific performance:

Not all participants internalized the optimal decision strategy in the same manner and performed equally well. A cluster analysis was used to group participants into high and low performers. Interestingly, eye movement patterns in the current experiment allowed to differentiate between high and low performers whereby high performers showed less visual search activity compared to low performers. This was in accordance with the investigation of expert and novice chess player by Charness et al. (2001). In their study, high performers made less fixations and fixated more on relevant information than low performers. An explanation for these findings might be that more fixations indicate inefficient visual search behavior, also stated by Goldberg and Kotval (1999). These participants might have difficulties to cope with the objective uncertainty using coping strategies like information search and suppression of ambiguous information (Lipshitz & Strauss, 1997) in an appropriate manner.

Task analysis:

Due to the novel character of the OVSST, we investigated mental processes within trials and found that visual search behavior was mainly shown after the prediction task of the OVSST, thus reinforcing the need of the reaction task. It seemed that attention was paid to the relevant features in the respective phases of the OVSST (cf. Kaakinen, Hyönä, & Keenan, 2002). Within trial analysis also showed that participants preferred AOI_{target} in case they switched their gaze to another AOI before they made their prediction, presumably due to an early anticipation of the later response. If the last fixation was in another AOI than AOI_{bottom}, then the $AOI_{predict}$ was preferred according to the decision making literature assuming that the chosen alternative is lastly fixated (Orquin & Mueller Loose, 2013).

Influence of confounding effects:

Attention as a possible confounding variable did not seem to affect learning results. However, interest in the OVSST could not be excluded as a possible confounding variable, but did not affect the experimental conditions in a systematic way and thus, might be interpreted as noise. Higher interest in the task led to less correctly predicted trials, possibly due to a more complex strategy and the phenomenon of probability matching (Fantino & Esfandiari, 2002). Participants who were more interested might have tried to develop a better strategy adapted to the probability concept considering also the lower probabilities whereas less interested participants always chose the likely exit. However, the latter strategy was more effective. Laude et al. (2012) reported similar results comparing choices between a more likely (75 %) and less likely (50 %) option for food of hungry motivated pigeons and less hungry and thus, less motivated pigeons. The highly motivated pigeons chose the suboptimal option more often presumably due to a greater impulsivity whereas less motivated pigeons preferred the optimal 75 % reinforcement more often. In a study about human gambling behavior Molet et al. (2012) also emphasized a tendency for suboptimal strategies for participants with higher gambling motivation, viz. more self-reported gambling activities.

Further research:

The developed mental model in Experiment I was unexpectedly accurate. It might be that an increase in task difficulty influences visual search behavior and further the learning process. In order to increase task difficulty and to impair visual search, the stimulus quality could be manipulated by including a distracting environment. The issue will be addressed in Experiment II.

3.5 Conclusion

The OVSST served as a valid method to study the development of mental models in an uncertain setting. Thereby, it was possible to investigate location uncertainty, discriminability of target objects, distraction and initial knowledge. The results of the study suggested that eye movement patterns provided information about the state of learning and the degree of the task uncertainty during the mental model acquisition. Furthermore, both coping strategies reported by Lipshitz and Strauss (1997), information search and ignorance of irrelevant information, seemed to matter for dealing with task uncertainty of the OVSST in this experiment. Finally, the estimation of the probability concept by the participants was better than expected, presumably due to a low level of task difficulty.

4 Experiment II – Degradation of Search Targets

4.1 Introduction

As became visible from Experiment I, participants were able to build up a mental representation of the OVSST already in the first block, and showed a good task performance in general. The learning process was reflected by the visual search behavior which became more focused across blocks with increasing experience. One important component for the development of mental models is the perception of the situation and its following processing steps. Therefore, the aim of Experiment II was to assess the influence of perceptual processes on the development of mental models by increasing visual search difficulty.

In the following, relevant literature about visual search is summarized to clarify possible effects of visual search difficulty on eye movements and further on cognitive processing as an extension of the aforementioned literature. In the typical paradigm of visual search, as mentioned earlier, a target stimulus is displayed amongst distractor stimuli whereby the target stimulus is not always present. Participants were asked to respond as quickly and as accurately as possible when they perceive either the presence or absence of the target stimulus (Müller & Krummenacher, 2006). In these experiments, the analysis of reaction times revealed two different visual search modes, namely serial and parallel visual search. These search modes are integrated in the feature integration theory (Treisman & Gelade, 1980) and consists of two stages. In the first, pre-attentive stage the focus is on parallel visual processing of basic features such as color, orientation and size (feature mapping) and occurs automatically. In the second stage of focused attention, all observed features are combined in order to ensure holistic perception requiring attention. If the target object differs from the distractors in a fundamental feature, the object is automatically and quickly detected without conscious awareness (pop-out effect) whereas the detection of targets with several different features requires a serial scan of the distractors (Treisman & Gelade, 1980).

The different visual search modes affect search efficiency, as expressed in search time and accuracy. Search efficiency decreases with increasing target-distractor similarity and decreasing distractor similarity (Wang, Cavanagh, & Green, 1994) and thus, should be reflected by eye movements. Search efficiency is influenced by task difficulty which can be manipulated by the discriminability of target and distractor stimulus. Target-distractor similarity affects performance directly (e.g., response times) and indirectly (learning behavior; Ahissar & Hochstein, 1997).

Another well-investigated phenomenon in visual search is the inhibition of return (IOR) effect which might be an additional factor influencing visual in the current context. IOR refers to a bias not to redirect attention to objects or locations shortly after they have been attended to (Klein, 2000; Posner, 1980). Thus, if a target object amongst similar looking distractors was ignored at one position, it takes much longer to find this target object now as in visual search novel locations were attended first before returning to familiar locations. In sum, the presence of distractors influences visual search in a negative way. However, not only distractors affect visual search, also the degradation of the target stimuli viz. if the shape of the stimuli is blurred. Sternberg (1967) investigated information processing by degrading stimuli. Results showed that visual stimulus degradation as an additive factor, leads to longer encoding time and thus longer reaction times.

Further studies, for example in the context of IOR research, suggest a strong overlap between the oculomotor system, attentional processes and visual working memory (Hoffman & Subramaniam, 1995; Theeuwes et al., 2009). Eye movements and attention processes are tightly connected: Eye movements are typically directed to the location where attention is allocated and attention is allocated at locations that are possible saccade targets. However, there is an inability to shift overt attention to one location and simultaneously move the eyes to another location (Hoffman & Subramaniam, 1995). In a nutshell, attention is the basic prerequisite for cognitive processing and thus, affects memory performance (Marois & Ivanoff, 2005). Therefore, it is suggested that cognitive processes involved in learning the object-exit association of the OVSST are also influenced by attentional processes and thus by visual search.

In summary, the reported findings in literature emphasize the strong coupling between visual search and cognitive processing. Thus, visual search influenced by the degradation of the stimuli is expected to affect the mental model development and lead to the following research questions.

4.1.1 Research Question and Hypotheses

The main research question of this experiment is whether the degradation of target stimuli interfere with the development of accurate mental models. As discussed, distractors might hinder visual search and thus, influence cognitive processing in the way that the accurate mental model acquisition is hampered due to the strong coupling of visual search and cognitive processing. To test this assumption, the target stimuli of the OVSST were degraded by inserting a distracting white-gray pattern in the background. This procedure is comparable with the stimulus degradation in the first experiment by Maisto and Baumeister (1975) who used a checkerboard pattern to degrade the stimulus.

First, it is expected that visual search is increased due to the distracting effect of the unstructured background. Thus, it is assumed that eye movement parameters show higher values in comparison to the first experiment. In literature, it was already reported that participants move their eyes more often with increasing target-distractor similarity (Eckstein, 1998). Second, it is expected that reaction times are longer than in the Experiment I (Maisto & Baumeister, 1975; Sternberg, 1969), especially during trials with incorrect predictions because here the pop-out effect is attenuated due to the distracting background. Third, it is suggested that more resources are required for the visual search due to degradation of the stimuli. Therefore, it is assumed that the development of the mental model takes longer so that the mental representation of the OVSST is less accurate at the end of Experiment II than Experiment I, viz. the estimation of the probability concept made by the participants should differ more from the presented probability concept than in Experiment I.

4.2 Method

4.2.1 Participants

A total of 23 participants (11 female) were tested individually at IfADo. Two of the participants did not affirm to see a relation of the target objects to the exits and were excluded from data analysis. Another four persons were excluded from data analysis due to a low data quality of the eye movement recordings. Finally, 17 participants (mean age=24 years, *SD*=3 years) entered the analysis. They were all students except for one. All of them fulfilled the conditions of participation in the experiment: normal vision and right-handedness.

4.2.2 Procedure

The procedure of the Experiment was similar to Experiment I (see General Method for a detailed description). The experiment was divided into a practice block (18 trials), the 100% condition (42 trials) and the experimental session (324 trials), whereby a 74-11-11 probability structure of the OVSST was used. The rare occurrence of 4% (reappearance of the target object at the bottom entrance) was maintained in order to compare results between the experiments. Participants were instructed to predict as accurately as possible and to react as quickly as possible. There was only one modification in comparison to Experiment I: The background of the screen was a gray-white pattern (Fig. 4.1) instead of white. The design of the unstructured background should impair visual search. Thus, a pattern with gray colored objects was chosen to increase target-distractor similarity. On the one hand, the shape of the objects for the unstructured background should not equal the target objects to prevent biases. On the other hand, the shape of the distractor objects should be similar (although not identical) to the target objects to decrease search efficiency (Eckstein, 1998; Wang et al., 1994). For that reason, an asymmetry was created by breaking the pattern apart and rearranging it to decrease distractor similarity. The distractors and target overlapped hereby causing an additional stimulus degradation (Sternberg, 1967).

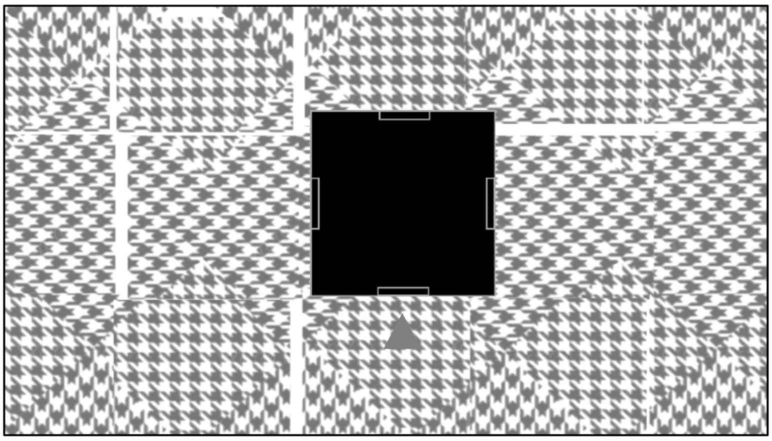


Figure 4.1: Unstructured background of the screen: The gray triangle starts to move into the room.

4.2.3 Data Analysis

In the 100% condition of Experiment II, participants predicted 70.9% of the trials (28 out of 42 trials) correctly. In analogy to Experiment I, the 100% condition was used to check for outliers of the sample after the practice block. There were no outliers for the task performance of the current sample in the 100% condition. Nevertheless, the box plot in figure 4.2 indicated a high interindividual variability.

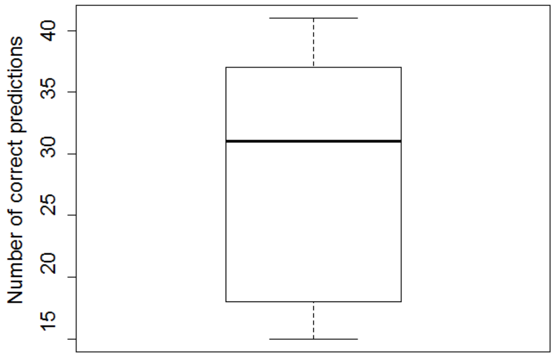


Figure 4.2: Box plot for the number of correct predictions in the 100% training condition of Experiment II.

In the experimental condition, 1.7% of the trials with missing predictions and another 0.9% of the trials with insufficient eye movement data points were excluded from data analysis. A two-way repeated measures ANOVA with the within-subject factors *block* (1-4) and *judgment* (correct, incorrect) and the between-subject variable *background* (white, white-gray patterned) was employed. The latter factor refers to the comparison between Experiment I and II. As in Experiment I, dependent variables were fixation frequency, fixation duration, the number of gaze shifts in terms of eye movement parameters and judgment time and reaction time in terms of behavioral data. Task performance, viz. the number of correct predictions was also used as dependent variable.

In order to compare data of Experiment II with data of Experiment I, mixed ANOVAs were performed. Variables were checked for normal distribution by using the Shapiro Wilk test before running planned t-tests. Shapiro Wilk test was chosen due to the small sample size (n=17). In the case of missing normal distribution, the non-parametric Mann-Whitney U test requiring no normal distribution, was used to check for the effects. Correlation analysis and repeated measures ANOVA were used to study the impact of the confounding variables on task performance.

4.3 Results

In the following, only significant (p<.05) results or trends (p<.10) were reported, except if the results were relevant for the aforementioned research questions.

Task performance:

Overall, 68.5% of the trials in Experiment II were correctly predicted. The number of correct predictions increased across blocks as shown by descriptive statistics (see Appendix C, Tab. 9.3 for details). However, only the increase from Block 1 to Block 2, $t(16)=2.89$, $p=.011$, and from Block 2 to Block 3, $t(16)=3.24$, $p=.005$, was significant.

Judgment time:

The main effect of *block* was significant, $F(3,48)=3.05$, $p=.038$, $\eta_p^2=0.160$, indicating a significant decrease of judgment time across blocks. In addition, we found a main effect of *judgment*, $F(1,16)=8.35$, $p=.011$, $\eta_p^2=0.343$, in the way that judgment times were longer in incorrectly predicted trials (Fig. 4.4A).

Analogous to Experiment I, a more detailed analysis was run for the variable judgment time to check for the reliability of the results. Judgment time was split into correct and incorrect judgments for likely and unlikely exits (see Fig. 4.3). Results for predicting the unlikely exit correctly were not depicted as the likelihood for this combination was low i.e. only one participant made correct predictions by predicting the unlikely exit which could be integrated in the analysis. Results of descriptive statistics showed that participants predicted the likely exit faster than the unlikely (likely exit: correct prediction $M=.442$, $SD=.150$; incorrect prediction $M=.443$, $SD=.150$; unlikely exits: incorrect prediction $M=.488$, $SD=.183$). However, this difference showed only a trend for the comparison of incorrect likely predictions and incorrect unlikely predictions, $t(6)=2.17$, $p=.073$, and not significant for the comparison of correct likely predictions and incorrect unlikely predictions, $t(6)=1.48$, $p=.190$.

Reaction time:

Analysis of *block* revealed a significant main effect, ($F(3,48)=3.07$, $p=.037$, $\eta_p^2=0.161$, indicating a significant decrease of reaction times across blocks. We also found a main effect of *judgment*, *F*(1,16)=8.35, *p*=.011, η_p^2=0.343, indicating shorter reaction times during correctly than incorrectly predicted trials (Fig. 4.4B).

<u>Fixation frequency:</u>

There was neither a main effect of *block*, *F*(3,48)=0.58, *p*=.568, η_p^2=0.035, nor a main effect of *judgment*, *F*(1,16)=0.25, *p*=.623, η_p^2=0.015, indicating no significant change (Fig. 4.4C).

<u>Fixation duration:</u>

We found no main effect of *block*, *F*(3,48)=1.441, *p*=.242, η_p^2=0.083, and no main effect of *judgment*, *F*(1,16)=0.32, *p*=.579, η_p^2=0.020, indicating no significant change (Fig. 4.4D).

<u>Number of gaze shifts:</u>

The main effect of *block* was significant, *F*(3,48)=4.06, *p*=.025, η_p^2=0.202, indicating that the number of gaze shifts decreased across blocks. We found also a main effect of *judgment*, *F*(1,16)=30.88, *p*<.001, η_p^2=0.659, indicating less gaze shifts for correctly predicted than for incorrectly predicted trials (Fig. 4.4E).

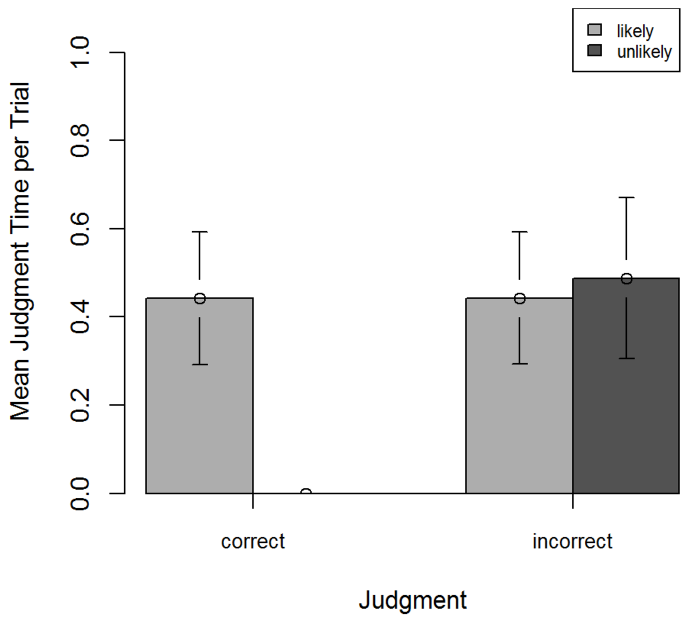


Figure 4.3: Judgment times in Experiment II: Correct and incorrected judgments for likely or unlikely exits. The bar for correct judgments and unlikely exits was not depicted as this combination was rare (n=1) and thus, was not valid for a comparison. Error bars show the standard deviation.

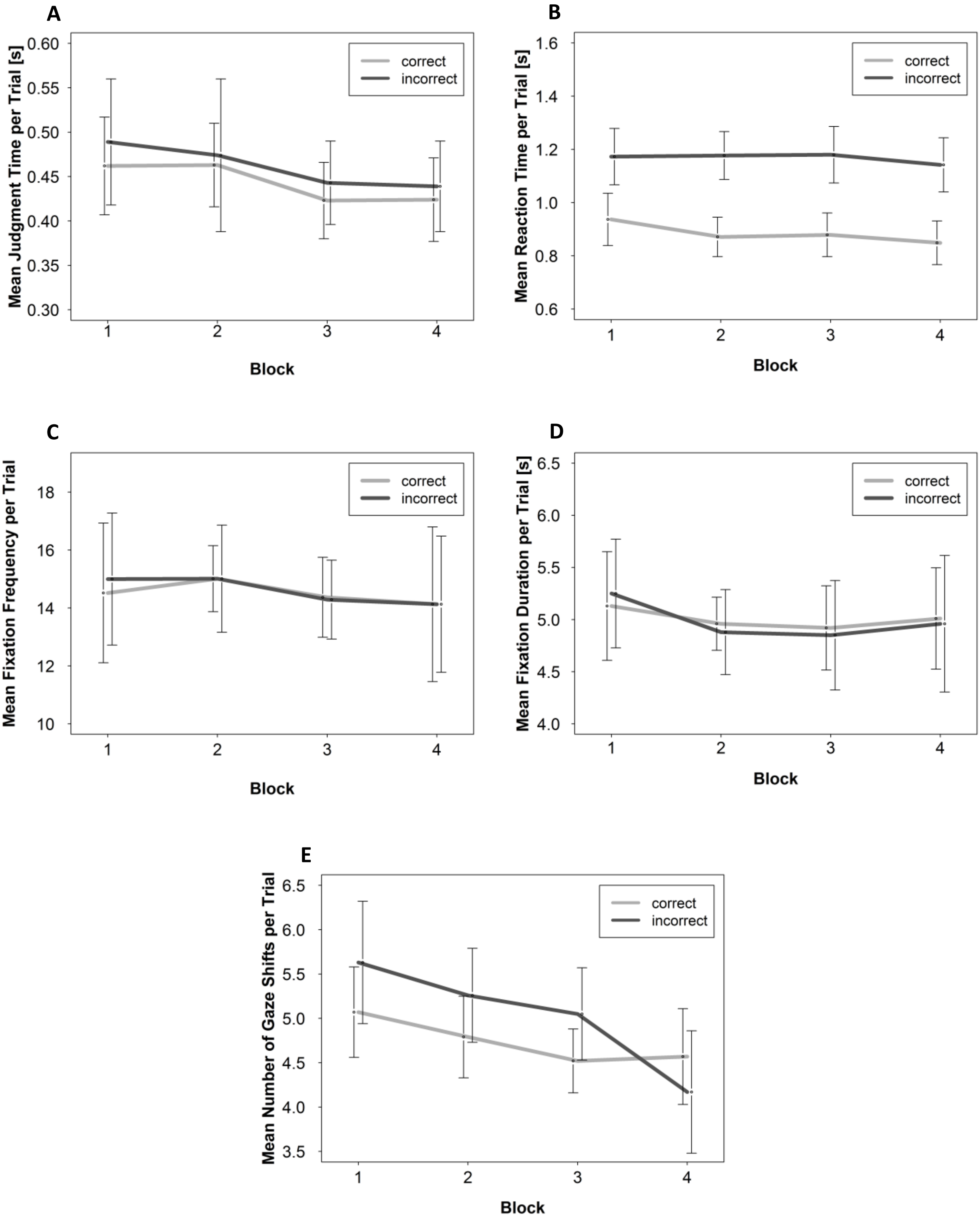


Figure 4.4: Results of the variables in Experiment II: Judgment time (A), reaction time (B), fixation frequency (C), fixation duration (D) and the number of gaze shifts (E) for correctly predicted and incorrectly predicted trials as a function of *block* and *judgment.*

Comparison of eye movement patterns and behavioral data between Experiment I and II:

Results of the between-subject factor *background* showed that participants made significantly more fixations per trial in Experiment II with a white-gray patterned background than in Experiment I with a white background, $t(33)=2.56$, $p=.015$. All other eye movement parameters did not differ significantly between Experiment I and II and there were also no interactions between the factor *background* and *block* or *judgment* ($p>.22$). There was also no significant difference between reaction times and judgment times in Experiment I and Experiment II, neither for correctly nor for incorrectly predicted trials ($p>.22$). In addition, performance clusters (high performance: n=9, low performance: n=8) resulting from hierarchical cluster analysis showed, in contrast to Experiment I, no significant group differences with regard to fixation frequency ($p>.65$), fixation duration ($p>.23$) and the number of gaze shifts ($p>.19$).

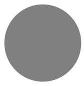
Analysis of the subjective probability concept and response behavior:

Behavioral data in Experiment II showed that participants mainly predicted the likely exit and were still able to make a realistic estimation of the objects' actual emergence at the exits indicated by the Concept Awareness Questionnaire (Tab. 4.2). The separation of the first experimental block into 4 equal parts showed that the associations between objects and exits were again already learned in the first 20 trials: in 69.4% of the cases the likely exit of the target object was predicted. Overall, interindividual variability ranged from 79.05% to 98.74% predicting the likely exit ($M=91.85$, $SD=6.087$) and 1.26% to 20.95% predicting the unlikely exits ($M=8.15$, $SD=6.087$).

Comparison of probability estimation and response behavior between Experiment I and Experiment II:

In order to compare the probability estimation and the response behavior in Experiment I and Experiment II, task performance and the amount of likely exit predictions were analyzed in a between-subject design. Further, the prediction error viz. the deviation of the estimated value from the target value (74% for higher probabilities and 11% for lower probabilities) for every target object was considered. As data in Experiment II was not normally distributed, the appropriate non-parametric Mann-Whitney tests were also calculated (Shapiro-Wilk: $p<.008$). However, all results were not significant ($p>.20$)

Table 4.1: Memory representation of the probability concept and behavioral probabilities in Experiment II

Object & Exit		Subjective Probability Concept	Performed Predictions
	left	12 % (6.7)	4 % (3.4)
	top	**78 % (9.6)**	**92 % (6.3)**
	right	10 % (5.2)	4 % (3.9)
	left	12 % (6.6)	4 % (4.3)
	top	11 % (7.4)	4 % (3.6)
	right	**77 % (10.0)**	**92 % (7.2)**
	**left**	**79 % (9.8)**	**92 % (6.8)**
	top	11 % (5.5)	4 % (4.0)
	right	10 % (6.5)	4 % (3.9)

Note. The object-exit associations that inhere a higher probability are shown in bold. Values in brackets show the standard deviation.

Analysis of Control Variables:

In analogy to Experiment I performance of the D2 test was better after the experiment than before, $t(16)=10.10$, $p<.001$. The participants detected on average 57% ($SD=10.65\%$) of the targets before running the experiment and 68% ($SD=12.43\%$) after the experiment. Furthermore, the error rate did not change from the pre-test to the post-test, $t(16)=0.58$, $p=.568$. Thus, attention did not seem to decrease from the beginning to the end of the OVSST.

Interest as a subscale of the QCM ($M=3.26$, $SD=1.16$) and its impact on task performance, defined as the number of correct predictions, was also tested. Spearman's rank correlation coefficient showed a trend for the correlation ($r=.357$, $p=.080$) between interest and task performance which might be a hint that interest in the task affects task performance positively.

4.4 Discussion

The aim of Experiment II was to investigate to what extent the stimulus degradation affects visual search behavior and the accuracy of learning a probability concept. Task difficulty seemed to be increased indicated by a higher fixation frequency in Experiment II than in Experiment I. However, reaction times did not differ between Experiment I and II. Finally, the performance of participants in Experiment II did not differ from Experiment I and participants estimated the probabilities of the object-exit associations as accurately as in Experiment I. Thus, the accuracy of the mental representation did also not seem to be affected by the stimulus degradation.

Mental model development:

Due to the increased task difficulty, as provided by the stimulus degradation in Experiment II, it was expected that task performance declined from Experiment I to Experiment II. However, the number of correct predictions did not differ between both experiments and also in Experiment II participants built up an accurate mental representation of the probability concept. Participants rather overestimated high probabilities than underestimating them as in Experiment I. In general, it seemed that the white-gray pattern used for the stimulus degradation rather supported learning instead of hampering cognitive processing. This assumption might be emphasized by a higher fixation frequency because Rayner (2009) assumed that new information is only processed during fixations. If participants were forced to make more fixations for identifying objects and to spend more attentional resources for encoding information, it might be that information processing and in a next step association learning was facilitated. This is in accordance with Kintsch (1988) who already highlighted the importance of attentional processes for an accurate mental model development in the context of text comprehension.

Effects on cognitive processing:

The increase in task difficulty was also expected to be accompanied by longer judgment times and reaction times, especially for incorrect predictions. This point was already addressed in the introduction of this chapter by outlining the state of the art about visual search. However, findings of Experiment II showed that judgment time was not influenced by the stimulus degradation. Results of the variable judgment time showed similar affects as in Experiment I, presumably due to the fact that the predictions were performed while the stimuli were not visible and thus, no further information was available through visual search. Interestingly, the findings of the detailed analysis of judgment times replicated the findings of Experiment I. Judgment times for unlikely predictions tended to be longer, fostering the assumption that a strategy change required more cognitive resources due to the inhibition of selecting the preferred strategy.

In addition, reaction times for correctly as well as for incorrectly predicted trials did also not differ significantly from Experiment I to Experiment II. One reason for this might be the short distance between the target object locations enabling to still perceive the target at the exit through peripheral vision which was identical in both experiments. In addition, the degraded target objects in the current experiment seemed to be still distinguishable from the white-gray patterned background due to the design of the pattern. One gray figure was created, duplicated several times and aligned next to each other with an offset. This pattern was broken apart in several pieces and rearranged so that it overlapped or drifted apart from each other so that similar features may be bound together within one group (feature integration theory; Treisman, 1998). Due to the fact that the target object inhered different features than the pattern of the background and moved out of the exit, the target object seemed to be still salient. This saliency in combination with peripheral vision and the learning effects probably lead to similar reaction times also for incorrect predictions as in Experiment I.

Furthermore, the results of the OVSST might not be comparable to the standard paradigm of visual search. For example, the four positions of the target objects at the exits were known in advance and the set size of the distractors was not manipulated in contrast to the typical design of visual search tasks (cf. Müller & Krummenacher, 2006). This consistency of the white-gray patterned background might be another reason why the stimulus degradation was not as distracting as expected. Solman and Smilek (2010) studied three different levels of consistency (random, repeated, intermediate) during visual search of a target letter and found that response times as well as fixation frequency increased with decreasing consistency. Additionally, in the study of Solman and Smilek search efficiency was improved during a repeated search condition in contrast to a random search condition. The first-mentioned condition might be comparable with the condition in the current experiment assuming that search efficiency was not highly impaired. The importance of consistency was also reported by Kristjánsson (2011) who found better task performance due to the consistent presentation of stimuli. Chun and Jiang (1999) published that visual search for a target was facilitated by the presence of the same distractors across all trials. The consistency of the background pattern is also accompanied with familiarity. Wang et al. (1994) already mentioned that the familiarity of the background reduces visual search and support parallel search. In sum, the current background design seemed not to influence reaction times due to the patterned background. An accurate mental model of the target locations and the object-exit associations seemed to be sufficient for the task performance of the OVSST because "[…] humans compute something close to an accurate posterior probability map and then use that map to determine the next fixation location efficiently." (Najemnik & Geisler, 2005, p. 390).

Visual search behavior:

It was expected that visual search activity was increased in Experiment II due to the white-gray patterned background. Results of analyzing eye movement data showed that participants made significantly more fixations in Experiment II than in Experiment I as expected. However, fixation duration and the number of gaze shifts did not change significantly from Experiment I to Experiment II. One reason for this might be that gaze shifts between the AOIs mainly took place after incorrect predictions according to a less beneficial action preparation and the redirection to the target object. As participant's task performance in the currently reported experiment was very good, preparation benefits could still be used and thus, participants did not have to switch often between the target AOIs and the incorrectly predicted AOIs. In literature, longer fixations were frequently related to deeper processing. For example, objects which did not fit in the context or words that are less frequent are usually longer fixated (Henderson, Weeks, & Hollingworth, 1999; Holmqvist et al., 2011; Rayner, 1998). Longer fixation duration was also associated with a higher degree of uncertainty (Brunyé & Gardony, 2017). Fixation duration did not increase possibly because the underlying task of the OVSST was not modulated from Experiment I to Experiment II and thus, demands on cognitive processing and the degree of uncertainty had not changed. Huang and Pashler (2005) already argued that the nature of the task is rather relevant for attentional processes than search efficiency, for example the complexity of the task. In contrast to fixation duration, fixation frequency might be increased in Experiment II perhaps due to its relevance for perceiving the target objects and distinguishing the object from the patterned background. Fixation frequency was already described as an indicator for search efficiency and task difficulty (cf. Goldberg & Kotval, 1999; Jacob & Karn, 2003). In conclusion, the stimulus degradation and thereby increased task difficulty presumably required more attentional processes for encoding of the target objects reflected by increased fixation frequency.

Influence of confounding effects:

Just like in Experiment I, attention as a possible confounding variable did not seem to affect learning results. In contrast to Experiment I, motivation rather influenced task performance positively than negatively. Possibly, motivated participants primarily focused on perceiving the object due to the unstructured background and not on developing a probability matching strategy as in Experiment I. Another reason might be that intended actions triggered by the instructions were followed more carefully influencing attentive processes and thus task performance (Bekkering & Neggers, 2002). Further, increasing task difficulty evoked by the unstructured background might lead to a higher motivation to increase task performance according to Locke (1968). However, the actual reasons remain unclear.

Further research:

Experiment II was designed to investigate how stimulus quality, as a kind of visual uncertainty, influenced the development of mental representations. In the next experiment, the influence of prior knowledge about the task on the development of the mental representation was studied. Therefore, the investigation of relearning might provide more specific information how fast an established mental representation can be adapted to a new situation, i.e. when the environment stays the same but the relations change.

4.5 Conclusion

The results of the second experiment showed that a distracted background which degrades stimuli did not generally impair cognitive processing and thus, memory performance. In the actual experiment, distraction led to an increased fixation frequency and rather facilitated learning the object-exit associations. In sum, the effect of stimulus degradation on task performance and thus the development of the mental model seemed be rather dependent on the design of the distractor and the nature of the task.

5 Experiment III – Dynamic Relearning

5.1 Introduction

Besides the stimulus design and the nature of the task, the mental model acquisition might also be affected by prior experiences according to Kim and Rehder (2011) who found that attention was guided by prior knowledge during category learning. Acquired knowledge about features of a specific category facilitated focusing on relevant information. The relevance of prior knowledge for mental model development was already mentioned in the main introduction and a brief theoretical motivation of Experiment III is provided in the following.

Experiment III was designed as a relearning experiment to investigate the learning process when prior knowledge has to be inhibited and to analyze the new development phase of the mental representation in a more detailed way. Relearning can be defined as the learning of new stimuli-outcome associations as the result of modifications in order to adapt to the change. Relearning might be more difficult than new learning as users had to give a second meaning to the actions which were learned before, probably leading to a larger ambiguity (Bouton, 2002). In task switching two different effects were already reported which might also interfere with relearning. Carry-over of prior relevant information as well as inhibition of prior irrelevant information might influence the stability of already learned object-exit associations which now have to be overcome (e.g., Koch, Gade, Schuch, & Philipp, 2010, for review). The inhibition of automatic responses to the earlier learned associations might also be related to a functional fixedness. The phenomenon of functional fixedness describes a cognitive bias preventing people to think outside the usual action strategies (Adamson, 1952; Knoblich et al., 2001). After a series of experiments on attention and perception, Shiffrin and Schneider (1977) postulated a dual process model of information processing including automatic and controlled processes, which was also relevant in later research (Chein & Schneider, 2012; Schneider & Chein, 2003). Controlled processes require attention and underlie the restrictions of short term memory whereas automatic processes do not require attention and could run in parallel. Furthermore, controlled processes can be turned into automatic processes via training, however, an inability of verbalizing the automatically processed actions might occur. This model reinforced the assumption that relearning is more difficult than learning because automatic processes has to be turned into controlled processes and thus, attention has to be guided. The change in control, however, should also affect eye movement behavior.

In the following research about conceptual changes, describing the adaptation of conflicting concepts, is shortly summed up as processes were comparable with those relevant during relearning.

Researchers particularly tried to answer the question at what point in time conceptual changes occur, which was also highly relevant, for example, in the context of teaching (Chi, 2008). Jones et al. (2015) investigated the role of attentional processes during conceptual change. The authors reported an effect of attention allocation on conceptual changes via cognitive engagement which was defined as the quality participants think about cognitive strategies. Attention seemed to evoke a high cognitive engagement which in turn increased the likelihood of conceptual change (Jones et al., 2015). It was argued that difficulties during conceptual changes mainly occurred due to a lack of awareness. In everyday life, mistakes in categorization (e.g., a whale is a fish) hardly occur. Therefore, the need for a recategorization is not present. However, conceptual changes occur if new information contradicts prior knowledge (Chi, 2008). Chi and Roscoe (2002), for instance, studied the incorrect belief of students about the heart circulation system as mentioned earlier in Chapter 1.1. The authors of the study assumed that conceptual shifts were difficult because misconceptions were often embedded in naïve and robust theories. Jones et al. (2015) underlined that strong prior knowledge led to a low probability of conceptual change. Additionally, Dole and Sinatra (1998) stressed motivational factors as well as dissatisfaction as crucial for conceptual changes. Overall, there seem to be many different factors influencing the relearning process. Therefore, the question arises how to gain insights into the process of relearning and its difficulties.

One possibility to approach this question was to consider research about eye movements and problem solving in conflicting situation. Knoblich et al. (2001) studied impasses during arithmetic matchstick problems. Participants were presented incorrect arithmetic statements including numbers and operators (plus and equal sign) with match sticks. Participants were asked to correct these arithmetic statements by relocating match sticks. Results of the study showed longer fixation times during impasses when problem solvers faced difficult problems. Furthermore, increased attention allocation to the operators revealed by eye movements, indicated the end of problem solving. Thus, eye movement patterns gave insights into problem solving. Knoblich et al. (2001) concluded that eye movement recordings are a powerful method to gain insights into problem solving, because attention allocation becomes trackable and shows the approach of the problem. Also other authors reported that eye movements reflect cognition and thus, strategic aspects during problem solving (e.g., Epelboim & Suppes, 1997; Grant & Spivey, 2003; Hegarty & Just, 1993). The findings of Knoblich et al. (2001) reinforced the representational change theory by Kaplan and Simon (1990) describing that initial mental representations that are not useful to solve the problem can be seen as a reason why problem solvers encounter impasses (cf. Jones et al., 2015). Further, this theory assumes that before solving the problem, useless initial mental representations have to be deactivated or inhibited. However, some of the inactive knowledge might also be relevant for the solution of the problem. To resolve the problem, initial mental representations have to be revised, for instance by constraint relaxation and chunk decomposition. The first idea describes "[…] the deactivation of some knowledge element that has acted as a constraint on the options initially considered […]" whereas the latter idea focusses on "[…] the separation of the components of a perceptual chunk […]" (Knoblich et al., 2001, p. 1001). In sum, during relearning well-established information has to be revised, attention has to be allocated to the now relevant information and finally, the mental representation has to be adapted. The research question and hypotheses, derived from the described literature reviews, were listed in the following.

5.1.1 Research Question and Hypotheses

The previous experiments investigated the learning of a probability concept, or in other words the development of a completely new mental model. In Experiment III the question arises how people relearn concepts and are able to adapt their existing mental model to changed circumstances. Thus, in this experiment it is investigated if eye movement parameters allow insights into relearning processes and the associated changes in subjective uncertainty. For this purpose, participants have to learn an initial probability concept of the OVSST as in the previous experiments. Then, they have to relearn the initial probability concept by integrating new object-exit associations for the same target objects as before without being informed about this change of concepts. This design focusses on higher-level cognitive processes that drive mental model development rather than on perceptual variations.

It is expected that relearning a concept is more difficult than learning a new concept because earlier learned associations have to be suppressed and attention has to be relocated. Thus, task performance, viz. the number of correct predictions should be larger during learning than relearning. It is also expected that all measured eye movement parameters (fixation duration, fixation frequency and the number of gaze shifts) indicate the relearning process in the way that participants show more visual search behavior, especially at the beginning of the relearning phase, to search for information and thereby to reduce uncertainty.

5.2 Method

5.2.1 Participants

25 students (10 female) were tested at IfADo. One participant had to cancel the test session due to dry eyes and the use of eye drops. Another participant did not finish the experiment, as calibration indicated that data quality of the eye movement recordings was too low. Further, four participants were excluded from data analysis because they were not able to develop the expected mental representation. Two of them filled in equal probabilities for all object-exit associations for one of the

probability concepts. Another two stated no understanding of any probability structure. All remaining participants affirmed that every object is associated to the exits with different relations via questionnaire. Finally, 19 participants with mean age 25 year (*SD*=4 years) entered the analysis. All of them were right-handed and had normal vision as required.

5.2.2 Procedure

The procedure remained the same as in Experiment 1 and 2. Participants had to predict the appearance of three distinct stimuli (circle, triangle, square) at three different exits (left, top, right) as accurately as possible presented on the white-gray patterned background used in Experiment II (Fig. 4.1). After the reappearance of the target object, participants were instructed to react on changes of the color intensity as quickly as possible. A practice block was run with 18 trials, followed by the 100% condition with an unambiguous allocation of the target objects to the exits (42 trials). In the experimental session, participants performed 324 trials (four blocks) of the OVSST while a first probability concept (concept I) with the following likely object-exit associations was presented: circle - top, triangle – right, square – left. The probability distribution for the target objects to the exits was the same as in the first two experiments with a 74% probability to reappear at the likely exit and 11% probability to reappear at each of the two unlikely exits. After four blocks, the concept changed (concept II) and another four blocks with 324 trials in total had to be performed with new object-exit associations. Concept II includes the following likely associations: circle-right, triangle-left, square-top. The probability distribution did not change, only the object-exit associations. Participants were not informed about any probability concept at the beginning of the experiment, nor was there any indication of the changed probability distribution after the first four blocks. The order of the two concepts was counterbalanced across participants by presenting all participants with odd numbers concept I at first and then concept II. Participants with even numbers had to initially learn concept II and afterwards concept I. A fixed pause of two minutes was included after every block in order to ensure that every participant had the same pause length. The pause was followed by a new calibration of the eye tracker. At the end of the experiment participants were informed about the two concepts by completing the Concept Awareness Questionnaire for both concepts. They also filled out the motivational questionnaire QCM. Additionally, participants performed the D2 test before and after the experiment to check for changes in attention as confounding variable.

5.2.3 Data Analysis

In the 100% condition of the current experiment, participants predicted 63.5% of the trials (26 out of 42 trials) correctly. The box plot diagram in figure 5.1 showed no outliers regarding the task performance of the current sample in the 100% condition. Nevertheless, there seemed to be a high interindividual variability indicated by the high range of the dependent variable.

Heat maps were used to check for drifts and thus, to ensure that data recording was accurate. 2.4% of the trials were excluded from data analysis due to missing predictions. Another 0.1% trials were excluded because of technical issues causing less than 65% valid eye movement data points within the trial. Data was analyzed as in the previous experiments: repeated measures ANOVAs with the within-subject factors *block* (1-4), *judgment* (correct, incorrect) and additionally *learning phase* (learning, relearning) were performed. Fixation frequency, fixation duration, number of gaze shifts, judgment time and reaction time as well as task performance, respectively the number of correct predictions, were used as dependent variables. Effects of the confounding variables on task performance were investigated by performing correlation analysis and repeated measures ANOVA.

In a first step, the presentation order of concept I and II, due to counterbalancing, was considered as a between-subject variable. Results showed no significant effect of this variable with regard to the aforementioned dependent variables (p>.26). Therefore, the data was collapsed over concept I and II. However, the analysis of the Concept Awareness Questionnaire revealed distinct group characteristics, which were analyzed in more detail in the corresponding paragraph.

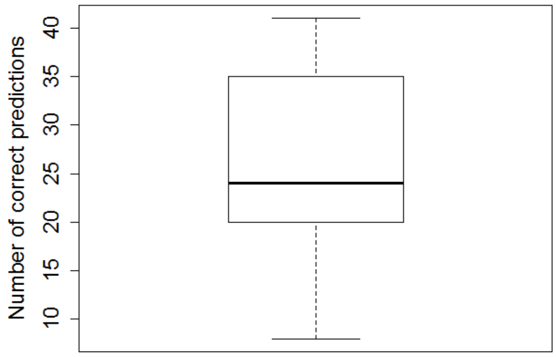


Figure 5.1: Box plot of the number of correct prediction in the 100% condition of Experiment III.

5.3 Results

In the following, only significant (p<.05) results or trends (p<.10) were reported, except if the results were relevant for the aforementioned research questions.

<u>*Task performance:*</u>

There was a significant main effect of *block*, $F(3,54)=4.35$, $p=.018$, $\eta_p^2=0.195$, indicating that the number of correct predictions increased across blocks during learning and relearning. We also observed that the number of correct predictions decreased significantly at the beginning of the relearning from Block 4 to Block 5, $t(18)=4.91$, $p<.001$. However, we did not observe a main effect of *learning phase*, $F(1,18)=1.72$, $p=.206$, $\eta_p^2=0.087$. Thus, correct predictions during learning (M=50.74; SD=6.26) and during relearning (M=48.79; SD=10.20) in contrast to our expectations (Fig. 5.3F).

Judgment time:

We found a main effect of *block*, $F(3,54)=6.11$, $p<.001$, $\eta_p^2=0.253$, indicating a significant decrease of judgment times across blocks. In addition, the main effect *judgment* was significant, $F(1,18)=8.97$, $p=.008$, $\eta_p^2=0.333$, indicating that participants judged faster during correctly than incorrectly predicted trials (Fig. 5.3A). Judgment times were analyzed in detail by splitting it into correct and incorrect judgments for likely and unlikely exits. Results showed that participants predicted the unlikely exit significantly slower than the likely exit. Judgment times for correctly judging likely exits differed significant from incorrectly judging unlikely exits, $t(11)=5.10$, $p<.001$, as well as incorrectly judging likely exits from incorrectly judging unlikely exits, $t(11)=4.84$, $p=.001$ (Fig. 5.2). The bar for correct predictions and unlikely exits was not depicted as this combination was only true for one participant due to its low probability and thus, too rare for a valid measurement. The tendencies are similar to those in Experiment I and Experiment II.

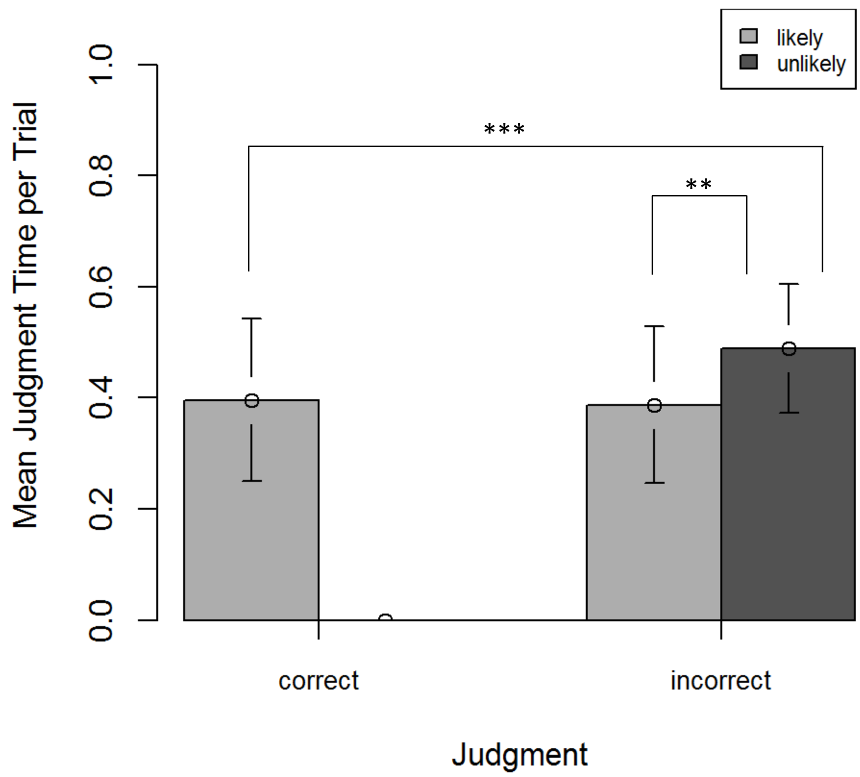


Figure 5.2: Judgment times in Experiment III: Correct and incorrected judgments for likely or unlikely exit. Judgment time for correct predictions and unlikely exits was not depicted as the occurrence of this combination was too rare (n=1). Error bars show the standard deviation.

<u>*Reaction time:*</u>

A main effect of *block* was found, $F(3,54)=3.05$, $p=.036$, $\eta_p^2=0.145$, indicating that reaction times decreased across blocks. We also found a main effect of *judgment*, $F(1,18)=148.46$, $p<.000$, $\eta_p^2=0.892$, indicating shorter reaction times during correctly than incorrectly predicted trials (Fig. 5.3B).

In addition, results concerning the reaction time showed an interaction of *block* and *judgment*, $F(3,54)=5.68$, $p=.002$, $\eta_p^2=0.240$, (Fig. 5.3B). Descriptive statistics showed that reaction times in correctly predicted trials decreased across blocks in the learning phase as well as in the relearning phase while reaction times in incorrectly predicted trials were almost constant across blocks in the learning phase as well as in the relearning phase (see Appendix C, Tab. 9.4 for details).

<u>*Fixation frequency:*</u>

Analysis of *block* revealed a significant main effect, $F(3,54)=3.95$, $p=.023$, $\eta_p^2=0.180$, indicating a significant decrease of fixation frequency across blocks. In addition, a main effect of *judgment* was observed, $F(1,18)=61.51$, $p<.001$, $\eta_p^2=0.774$, indicating fewer fixations during correctly than incorrectly predicted trials (Fig. 5.3C). There is also a significant increase of fixations from Block4 to Block5

<u>*Fixation duration:*</u>

There was neither a significant main effect of *block*, $F(3,54)=0.783$, $p=.466$, $\eta_p^2=0.042$, nor a significant main effect of *judgment,* $F(1,18)=0.09$, $p=.774$, $\eta_p^2=0.005$), indicating no significant change (Fig. 5.3D).

We only found a significant effect of *learning phase* for fixation duration, $F(1,18)=5.39$, $p=.032$, $\eta_p^2=0.230$ (Fig. 5.2E), indicating that fixation duration was significantly longer in the learning phase than during relearning.

<u>*Number of gaze shifts:*</u>

We observed a main effect *block*, $F(3,54)=4.28$, $p=.009$, $\eta_p^2=0.192$, indicating a significant decrease of the number of gaze shifts across blocks. We found also a main effect of *judgment*, $F(1,18)=151.75$, $p<.000$, $\eta_p^2=0.894$, indicating that participants showed less gaze shifts for correctly predicted than for incorrectly predicted trials (Fig. 5.3E).

In addition to the main effect, we found a significant interaction of *learning phase* and *judgment*, $F(1,18)=8.014$, $p=.011$, $\eta_p^2=0.308$, (Fig. 5.3E), i.e. the number of gaze shifts in incorrectly predicted trials was almost similar in the learning phase ($M=4.96$, $SD=1.44$) and the relearning phase ($M=4.93$, $SD=1.58$) whereas the number of gaze shifts in correctly predicted trials decreased from $M=4.20$ ($SD=1.61$) in the learning phase to $M=3.98$ ($SD=1.78$) in the relearning phase.

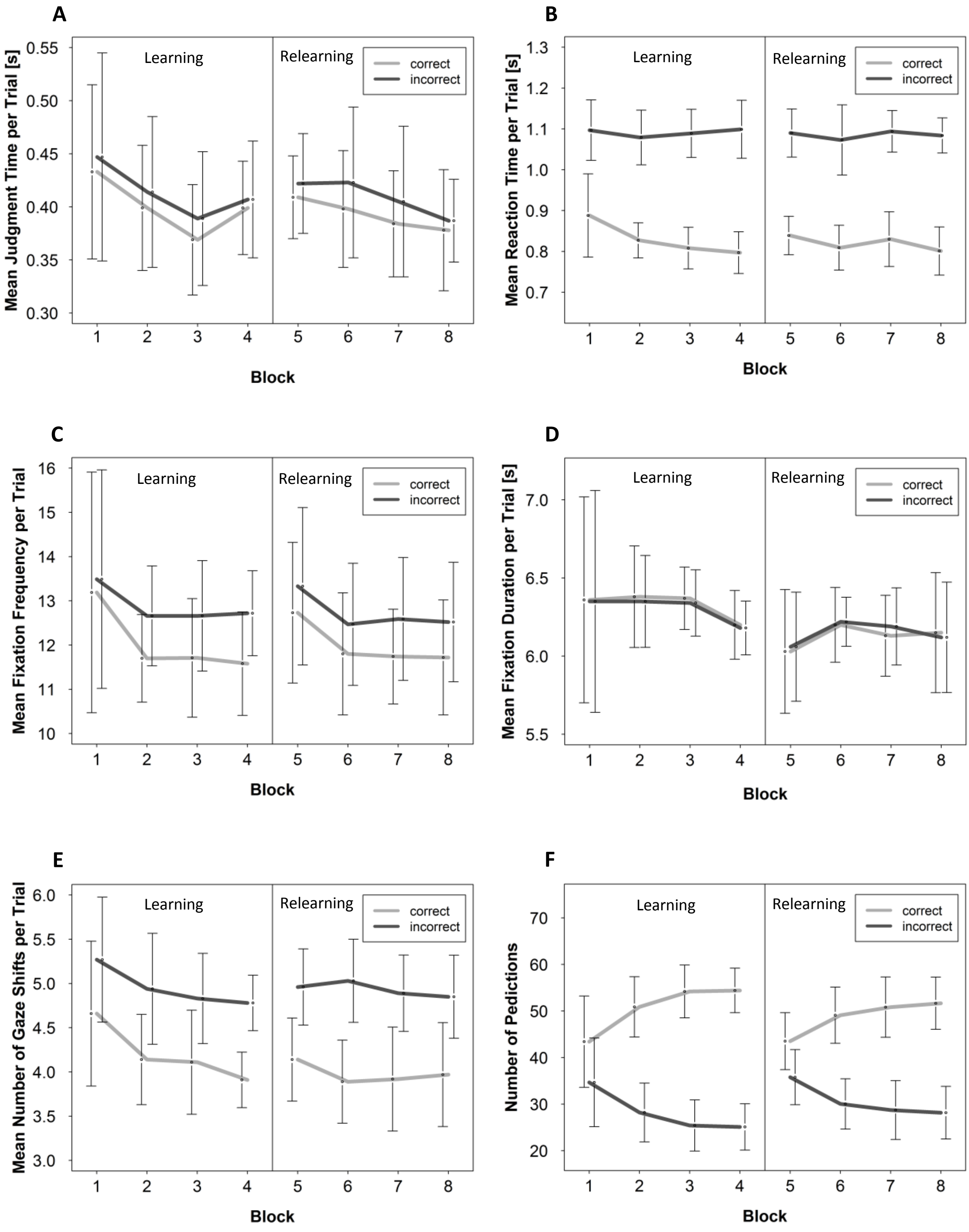


Figure 5.3: Results of the variables in Experiment III: Judgment time (A), reaction time (B), fixation frequency (C), fixation duration (D), number of gaze shifts (E) and number of predictions (F) during learning and relearning for correctly and incorrectly predicted trials as function of *block*, *judgment* and *learning phase*.

Analysis of the subjective probability concept and response behavior:

In the experimental condition, participants predicted on average 63.1% of the cases correctly during learning and 56.8% during relearning. As shown in table 5.1 and 5.2 participants were able to estimate the probabilities for concept I, which was also used in Experiment I and Experiment II, almost accurately even if they had to learn or relearn the concept. However, the probabilities of concept II were considerably less accurately estimated. Therefore, both concepts were analyzed in more detail in the following.

The prediction error of concept I was not significant regarding likely exits estimations of concept I, i.e. the estimation of likely exits did not differ significantly from the given value of 74% ($p>.14$). In contrast, the subjective probability concept participants developed of concept II showed significant differences between the likely exit estimations and the given probability of 74% with regard to the circle, $t(18)=2.73$, $p=.014$, and the triangle, $t(18)=2.89$, $p=.010$. Results showed only a trend for the prediction error of the square, $t(18)=1.99$, $p=.062$. Further, there was a trend that estimations of the likely exits differed from each other. There was a trend that the square-top exit association differed from the triangle-left exit association, $t(18)=1.45$, $p=.082$. In sum, the shape of the objects with regard to the exit position might influence learning.

Behavioral data might provide an explanation for differences between learning and relearning of the concepts: During learning of concept I and concept II, participants predicted mainly the likely exits. During relearning, participants who had to relearn concept I predicted mainly the likely exits in 85% of the cases. However, participants who had to relearn concept II did not show such a strong tendency for the likely exits shown in table 5.1 and 5.2. They predicted the likely exits in 57% of the cases. Inference statistics showed no significant difference between likely exits prediction during relearning concept I and II, $t(9)=1.22$, $p=.130$, but Levene's test indicated unequal variances ($F=5.92$, $p=.026$). Thus, data seems to be inconsistent with regard to the relearning phase.

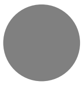
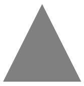
Behavioral data also indicated that the previous learned probability concept affects relearning, especially relearning of concept II, in the way that earlier learned associations have to be suppressed. Planned *t*-test showed a trend for differences between learning and relearning concept II with regard to the task performance for the following likely associations: circle and top exit, $t(9)=1.81$, $p=.055$), square and left exit, $t(9)=1.678$, $p=.065$, but not for triangle and right exit, $t(17)=1.25$, $p=.120$. Further, variances were not equal for the circle-top association ($F=7.997$, $p=.012$) and the square-left association ($F=9.681$, $p=.006$). The same statistics for the according object-outcome associations of concept I showed no significance ($p>.15$) and thus equal variances.

Analysis of Control Variables:

Results of analyzing the D2 test seem to be robust for the OVSST due to similar task performance as in Experiment I and Experiment II indicating that OVSST is not affected by changes in attentional performance. Data of one participant was excluded from data analysis due to a misunderstanding of the task. Overall, participants again showed better task performance after the experiment than before, $t(17)=7.74$, $p<.001$. They performed on average 62% (SD=13.26%) of the cases before running the experiment and 70% (SD=11.27%) of the cases after the experiment. The error rate stayed at the same level, $t(17)=0.50$, $p=.625$.

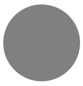
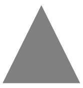
Interest measured by the appropriate subscale of the QCM (M=3.09, SD=1.16) showed no effect on task performance, viz. the number of correct predictions. Spearman's rank correlation coefficient showed a non-significant unilateral correlation (r=.185, p=.225) between interest and task performance.

Table 5.1: **Probability concept I** in Experiment III - Memory representation and behavioral probabilities during learning and relearning the concept

Object & Exit		Subjective Probability Concept	Performed Predictions during learning	Performed Predictions during relearning
	left	14 % (6.3)	7 % (3.9)	6 % (8.1)
	**top**	**72 % (10.9)**	**86 % (6.6)**	**84 % (13.9)**
	right	14 % (5.8)	7 % (6.1)	9 % (6.8)
	left	18 % (15.1)	6 % (4.1)	9 % (7.7)
	top	14 % 4.9)	7 % (7.3)	3 % (4.2)
	right	**68 % (18.6)**	**87 % (9.4)**	**88 % (10.7)**
	left	**67 % (20.8)**	**82 % (12.0)**	**84 % (14.3)**
	top	18 % (16.2)	11 % (8.1)	14 % (12.6)
	right	15 % (14.6)	7 % (5.6)	2 % (2.5)

Note. The object-exit associations that inhere a higher probability are shown in bold. Values in brackets show the standard deviation.

Table 5.2: **Probability concept II** in Experiment III - Memory representation and behavioral probabilities during learning and relearning the concept

Object & Exit		Subjective Probability Concept	Performed Predictions during learning	Performed Predictions during relearning
	left	21 % (18.0)	8 % (8.8)	9 % (9.2)
	top	21 % (17.9)	8 % (7.2)	21 % (18.9)
	right	**58 % (25.1)**	**84 % (15.7)**	**70 % (26.9)**
	left	**57 % (26.1)**	**85 % (15.4)**	**79 % (21.6)**
	top	13 % (5.9)	7 % (7.5)	6 % (6.3)
	right	30 % (25.8)	8 % (8.3)	15 % (17.1)
	left	18 % (14.2)	6 % (6.1)	19 % (21.0)
	**top**	**66 % (18.3)**	**87 % (12.5)**	**73 % (26.6)**
	right	16 % (10.2)	7 % (6.6)	8 % (8.5)

Note. The object-exit associations that inhere a higher probability are shown in bold. Values in brackets show the standard deviation.

5.4 Discussion

In addition to the learning process, in Experiment III the relearning of concepts was examined. Task performance of the OVSST did not differ significantly between both phases. However, behavioral data indicated group differences of relearning concept I and II resulting in a less accurate subjective probability of concept II. Eye movement parameters did not reflect this difference. However, they seemed to inform about the conceptual change process during relearning and entailed specific characteristics. More specifically, fixation frequency signalized the beginning of relearning immediately, but reached a rather stable level after one block. The number of gaze shifts showed a similar pattern, however, the specification was less noticeable. Fixation duration decreased in parallel to the learning curve and thus reflected the general learning progress. Further, fixation duration responded to the relearning phase with a time delay and thus might indicate the delay in the adaptation of the mental representation.

Mental model development:

Unexpectedly, task performance during learning was not significantly better than during relearning the object-exit associations. However, unequal variances and prediction errors during relearning concept II showed a trend that the relearning phase was more difficult with regard to the suppression of prior knowledge, especially for some of the participants. Despite of a non-significant main effect of group, there seemed to be group differences at least during relearning concept II which indicated that results have to be interpreted with caution. Reasons for that unexpected group difference might be that probability concept I was more extinction resistant than concept II and thus there might be a bias elicited by the different object-exit associations. Another reason might be the small sample size and the partitioning of participants into two groups (concept I and concept II) resulting in lower power of the data. Other reasons for the overall missing difference between learning and relearning might base on the performed task. The OVSST seemed to be a simple task as shown by the steep learning curve at the beginning of the task indicating early learning effects. Furthermore, a lack of awareness that might impede conceptual changes as reported in the introduction seemed not to be relevant in the current case as actions were not rare, but rather repeated in every trial (cf. Chi, 2008). Thus, attention was always allocated to the display. All these aspects facilitated relearning in the present context and might be the reason for the unexpectedly good task performance during the relearning phase.

Interestingly, unequal variances of correct predictions indicated greater interindividual variability during relearning of concept II. It seems that some participants are able to adapt to the changed situation better than others, presumably due to different levels of cognitive processing and delayed realization of the modification (cf. Jipp, 2016). The prediction error of the subjective probability concept might indicate that concept I was more dominant in the way that participants were not able to relearn concept II in an accurate manner. There seemed to be still some artifacts of concept I in estimating the unlikely probabilities regarding concept II, for example, the triangle of concept II was still associated to reappear to 30% at the right exit which was the likely exit of the triangle in concept I. Furthermore, behavioral data showed that initially learned associations were more often used during relearning of concept II whereas during relearning concept I almost no behavioral differences between both phases were shown. Therefore, it might be that the object-exit associations of concept I were easier to relearn and more resistant to extinction. In addition, it seemed that participants learned the square-top exit association better than other associations of concept II. Thus, other target objects should be used in further studies to avoid such biases. Furthermore, the movement of the target objects out of the exit might give directional information also supporting biases. Thus, it would additionally suppress biases if target objects fade in at the exits and do not move (cf. Itti, Koch, & Niebur, 1998; Treisman, 1985).

Effects on cognitive processing:

Judgment time and reaction time patterns in the learning phase replicated almost all findings in Experiment I and Experiment II and thus, supported the validity of the results, as they were faster for correct predictions and decreased across blocks. Interestingly, an interaction effect occurred between *block* and *judgment* for reaction times. Reaction times were constantly high for all eight blocks in incorrectly predicted trials but decreased in correctly predicted trials over all blocks independently of *learning* phase. An explanation might be again the missing anticipation of the correct exit during incorrect prediction and the general learning of the OVSST for the decreasing reaction times during correct predictions.

Visual search behavior:

Results of eye movement data analysis confirmed the assumption that participants showed more visual search behavior at the beginning of the relearning phase. Eye movement parameters indeed informed about difficulties to adapt to the changed concept, but they provided distinct information and did not differ between relearning concept I and concept II, indicating that they only reflected general characteristics of the task. The number of gaze shifts and especially fixation frequency might reflect disturbances in performing the OVSST due to the changed probability concept. Extensive visual search developed into a more focused search as reported in earlier studies (Ellis, 2012; Jacob & Hochstein, 2009) was also valid for the relearning process. Generally, the number of gaze shifts and fixation frequency might be lower for correctly predicted trials as the relevant exit was already anticipated. However, the number of gaze shifts for correctly predicted trials decreased from the beginning to the end of each learning phase whereas the number of gaze shifts for incorrectly predicted trials was rather constant in both phases, presumably due to the aforementioned missing anticipation which did not allow to improve visual search behavior. The decreasing number of gaze shifts during correctly predicted trials in both phases might be related to the increasing number of correct predictions and thus, to the increased accuracy of the mental representation which allowed to improve visual search and to focus more on relevant stimuli.

Further, results seemed to indicate that fixation duration was a possible indicator to differentiate between learning and relearning. Fixation duration was significantly longer in the learning than relearning phase. In literature, fixation duration was usually associated with difficulties in extracting information (Ehmke & Wilson, 2007; Jacob & Karn, 2003; Poole & Ball, 2006). In the current task, however, the way information had to be extracted did not change. It might be rather the familiarity with the task over time that led to an increased fixation duration in the relearning phase.

In contrast to fixation duration, the number of gaze shifts and fixation frequency, fixation duration responded to the changed concept with a time delay presumably due to the delayed awareness about the modification of the concept. The situation might be comparable with impasses studied by Knoblich et al. (2001) who reported longer fixation duration during impasses as already mentioned in the introduction of the chapter. At the beginning of the relearning phase participants had to be aware about the changed concept and thus, had to inhibit prior knowledge about the object-exit associations. During this process participants might also encounter a kind of impasse which might be the reason why fixation duration increases in the second block of the relearning phase.

Finally, the crucial question to be asked was what eye movements actually represent. Eye movements seemed to represent the learning process in a new environment and the search of relevant information to adapt to changed situations, but they did not represent the degree of accuracy of the mental representation. Difficulties in relearning concept II were not represented by eye movement parameters and thus, other processes like information processing seemed to be crucial for the relearning phase presumably because of the stable environment.

Influence of confounding variables:

The findings of the current experiment showed that attentional resources did not seem to affect learning results of OVSST. Thus, the findings of Experiment I and Experiment II could be replicated. In contrast to Experiment I and Experiment II, results showed no effect of motivation on task performance, neither in a positive nor in a negative way. Thus, the earlier mentioned findings by Dole and Sinatra (1998) could not be reinforced by the current experiment. The authors argued that motivational factors were relevant for conceptual changes. In sum, motivation did not seem to qualify a systematic influence on task performance.

Further research:

Another issue, also mentioned in the introduction, was the interaction between the acquisition of the mental representation and the two tasks of the OVSST. It was still unclear whether participants develop a mental representation of the OVSST based on the prediction task, the reaction task or a combination of both, as participants had to perform both tasks in Experiment I, II and III. Further, judgment times and reaction times in almost all experiments showed similar effects: they were slower for incorrect predictions and decreased across blocks. Thus, they did not indicate which of the corresponding task was essential for the mental model development.

5.5 Conclusion

Overall, results of the study showed that there was no appreciable difference between learning and relearning the probability concepts of the OVSST with regard to the task performance. However, eye movement patterns informed about the state of learning and the effect of relearning during the adaption to new features of the OVSST in different ways. Subjective data indicated a less accurate mental representation of concept II after the experiment and a bias regarding the shape of the objects. Therefore, stimuli of the OVSST should be modified in future studies.

6 Experiment IV - Separate Tasks

6.1 Introduction

In all previous experiments, participants were able to develop an accurate mental representation of the OVSST. However, some underlying processes of the OVSST were still unclear, especially whether the mental representation of the probability concept was developed due to the performance of the prediction or reaction task. Judgment times and reaction times were affected by learning in the previous experiments and thus, did not allow to conclude which of the two tasks was essential for the mental model development. Within the scope of the current experiment, we investigated to what extent the prediction and the reaction task affected the mental model development of the OVSST.

The OVSST might be related to aspects of the spatial cueing paradigm, introduced by Posner (1980) in the context of visual attention. In the spatial cueing paradigm, participants were asked to respond to a target which was presented after a cue. The cue indicated the most likely location for the target. Exogenous cues highlighted the target location and directly attracted attention, whereas endogenous cues pointed towards the target location and required the conscious guidance of attention. Finally, the cue type affected response times in the way that responses in trials with exogenous cue were usually faster than in trials with endogenous cues (Posner, 1980). Some aspects of the spatial cueing paradigm might be comparable with the OVSST. In the prediction task, the target shape also functioned as an endogenous cue for the target exit. In the reaction task, the changed color of the target might directly attract attention like the endogenous cue.

Endogenous and exogenous attentional shifts commonly typify top-down and bottom-up control (e.g., Kahneman & Tversky, 1973; Theeuwes, 2010). Due to the respective task characteristics, the prediction task could be assigned to top-down processes which were assumed to be goal-driven, slow, volitional and endogenously oriented. On the other hand, the reaction task might be a bottom-up process, characterized by being stimulus-driven, rapid, automatic and exogenously oriented (e.g., Connor, Egeth, & Yantis, 2004; Desimone & Duncan, 1995; Hauer & MacLeod, 2006; Kristjánsson, Mackeben, & Nakayama, 2001). Hauer and MacLeod (2006) added that cognitive control was involved in top-down processes whereas bottom-up processes required no cognitive control. Based on their study about attentional cueing of words, they concluded that endogenously cued attention, for example, a row of arrows pointing towards a word, led to more active learning and affected later memory processes in a beneficial way. This dichotomy was also visible when analyzing eye movements, namely goal-driven attention, should be reflected by slower eye movements than stimulus-driven attention (Engelkamp, 2006; van Zoest et al., 2004; van Zoest, Donk, & Theeuwes, 2004).

In the field of reasoning and decision making there was also clear evidence for a dual-process reported by Evans and Stanovich (2013) inspired by Kahneman (2011). They distinguished between more rapid and autonomous processes, called Type 1, and higher order reasoning processes relying highly on working memory, called Type 2. The dual process model assumed that two forms of cognitive processing were evoked by cognitive tasks. The reflective Type 2 process led to a mental simulation and contained explicit knowledge of the task whereas the intuitive Type 1 process was independent of cognitive abilities and contained implicit knowledge (Evans, 2003; Evans & Stanovich, 2013).

In the field of learning research there was also a dichotomy reported, namely implicit learning and explicit learning. Implicit learning involved the acquisition of knowledge with a lack of conscious awareness. Thus, learners were not able to verbally report the learned knowledge whereas explicit learning resulted in conscious knowledge (e.g., Ziori & Dienes, 2012). Regarding the two tasks of the OVSST it seemed to be plausible that the prediction task was based on explicit learning including higher-level cognition to make decisions. In contrast, the reaction task rather required implicit learning as participants only had to react on color changes of the stimuli, however, they might accelerate their reaction times by learning the probability concept and anticipating the correct exit (cf. Chun & Jiang, 1998; Jungé, Scholl, & Chun, 2007).

In research about category learning, faster and slower processes during decision making were already reported. For example, Chen et al. (2016) published that participants made faster decisions during category learning when the decisions were based on simple associations and slower decision when strategies had to be consciously selected. Thus, the prediction and the reaction task might be reflected by reaction times. Research about category learning also indicated effects of different learnings processes on the accuracy of the mental model development. Ziori and Dienes (2012) outlined that in studies with single tasks and salient features, comparable with the OVSST, implicit learning processes might lead to less accurate mental representations than explicit learning processes. Further, they concluded that explicit knowledge was based on accurate mental representations which were learned via practicing, i.e. the frequent repetition of an event. Thus, learning the probability concept of the OVSST accurately and allocating attention to the relevant information seemed to require the frequent repetition of target locations (cf. Kabata & Matsumoto, 2012; Ziori & Dienes, 2012).

In both tasks of the OVSST, the prediction and the reaction task, target locations were repeated and feedback was provided. In the prediction task, target objects reappeared at one of the exits. Thus, participants could compare their prediction with the actual appearance of the target object at the exits and conclude if the prediction was correct. Finally, they could adapt their decision-making behavior to the target probabilities. In contrast to the prediction task, participants only received error feedback, but no performance feedback in the reaction task. The following research questions regarding the prediction and the reaction task were devised based on the mentioned literature.

6.1.1 Research Question and Hypotheses

The aim of Experiment IV was to assess how the two tasks of the OVSST actually influence the development of the mental representation by studying the prediction task and the reaction task separately. According to the aforementioned research about implicit and explicit learning (Batterink, Reber, Neville, & Paller, 2015; Ziori & Dienes, 2012), it might be that both tasks influence the development of the mental representation. However, research about top-down and bottom-up processes would suggest that only the prediction task influences the development of the mental representation due to a deeper processing of the task (Connor et al., 2004; Desimone & Duncan, 1995; Hauer & MacLeod, 2006; Kristjánsson et al., 2001). Thus, it remains an open issue how the prediction and the reaction task influence the acquisition of the mental representation about the probability concept of the OVSST.

If the prediction task is related to goal-driven top-down processes and the reaction-task to rather stimulus-driven bottom-up processes, it is expected that the subjective probability concept is only learned accurately during the prediction task and not during the reaction task (Craik & Tulving, 1975). Therefore, probabilities estimated via the Concept Awareness Questionnaire are expected to be close to the actual probabilities for the concept presented in the prediction task, but not for the concept presented in the reaction task. As mentioned earlier, cognitive processing should be reflected by eye movements according to the eye-mind assumption. Thus, it is expected that eye movement patterns reflect the learning processes in the prediction task, i.e. by decreasing fixation frequency, fixation duration, number of gaze shifts, scanpath velocity and scanpath distance. In the reaction task, these eye movement parameters are expected to remain unchanged due to the missing learning. It is also expected that eye movements during the prediction task are slower than during the reaction task according to van Zoest et al. (2004) and Engelkamp (2006). They stated that fast eye movements are stimulus driven whereas slow eye movements are goal driven.

In contrast, if the prediction task and the reaction task are related to explicit and implicit learning processes, it is expected that reaction times in the reaction task decrease across blocks due to benefits of the unconsciously learned probability concept (e.g., Chen et al., 2016; Chun & Jiang, 1998; Chun & Jiang, 1999; Jungé et al., 2007; Ziori & Dienes, 2012). Eye movement parameters are also expected to reflect the learning process by decreasing values across blocks in the prediction as well as in the reaction task. However, due to the fact that participants cannot report the unconsciously learned probability concept in the reaction tasks, the estimated probabilities are expected to be accurate for the concept presented in the prediction task but not for the concept presented in the reaction task.

6.2 Method

6.2.1 Participants

A total of 26 people participated in the experiment at the University of British Columbia in the Brain and Attention Research Lab. The analysis of the Concept Awareness Questionnaire showed that nine participants could not comprehend any probability structure which is crucial for the analysis of the mental model development. Thus, they were excluded from data analysis, except for the cluster analysis. Additionally, the calibration values of two participants indicated low data quality of the eye movement recordings and data recordings of another two were erroneous. Another person had to be excluded due to technical issues. Data of the remaining 12 participants with mean age 26 year (*SD*=4 years) was included in the data analysis. All of the remaining participants were right-handed and had normal vision without visual aid. Further, they all affirmed that every object is associated to the exits with different relations by filling out the Concept Awareness Questionnaire.

6.2.2 Procedure

In Experiment III there was a bias with regard to the shape of the object associated to the exits. Thus, it was suggested to improve the OVSST by using different target objects and avoiding the movement out of the exits. These suggestions were realized in the current experiment. There was no longer an unstructured background, but the objects itself were degraded and presented on a gray background: Fuzzy Gabor figures (cf. Thornton & Gilden, 2007) with horizontal, vertical and diagonal lines were created (Fig. 6.1).

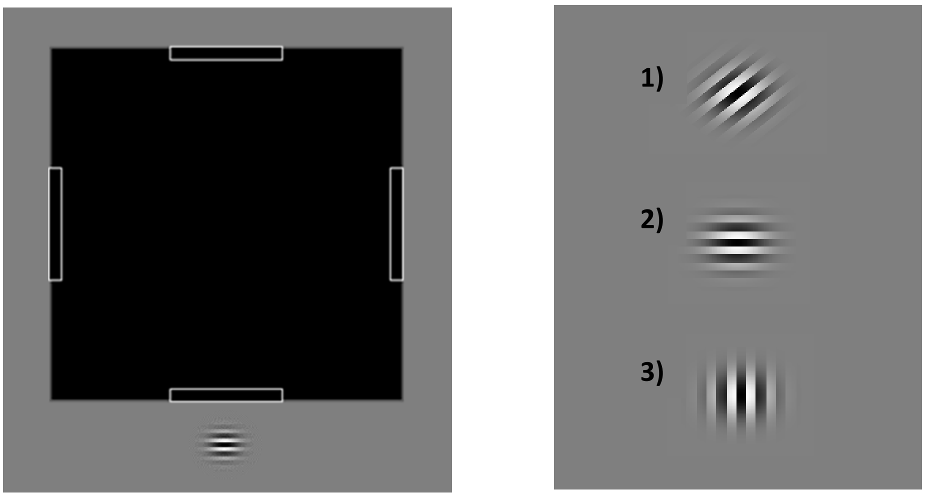


Figure 6.1: The OVSST with fuzzy Gabor figures in Experiment IV: Fuzzy Gabor figures moving from the bottom entrance into the black room and fading in at the exits (left). Fuzzy Gabor figures with diagonal (1), horizontal (2) and vertical stripes (3) (right).

The Gabor figures faded in within three seconds at the exits and thus, no longer moved out of the exits. As the change of color intensity would not be salient due to the gray background, participants were asked to react on the reappearance at the exits in case of a color change from black-white to red-green.

In addition, the OVSST was split into two separate parts: a prediction task and a reaction task. After a phase of training of the particular task, participants performed four blocks of the OVSST only predicting the exit of the Gabor figures as quickly and as accurately as possible. After another training phase for the second task, participants were instructed only to react on color changes of the Gabor figures for four blocks as quickly and as accurately as possible (Fig. 6.2). Participants had a fixed two minutes' pause after every block and a new calibration was run before starting the next block.

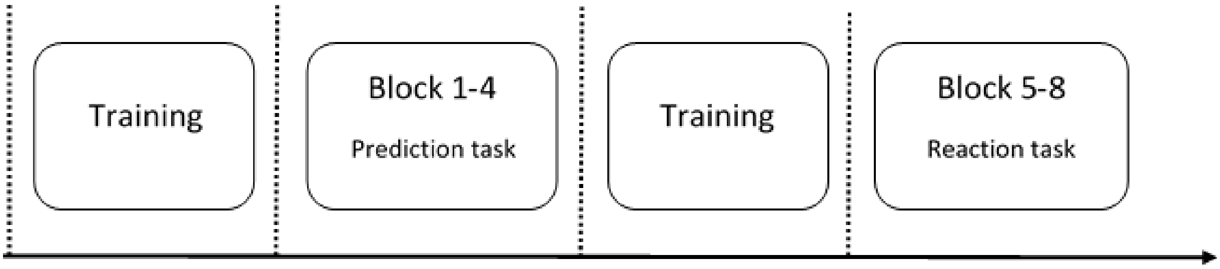


Figure 6.2: Temporal sequence of Experiment IV. Dotted lines indicate when participant received instructions.

The order of the two tasks was counterbalanced across participants by instructing all participants with odd numbers to perform the prediction task at first and then the reaction task. Participants with even numbers had to initially perform the reaction task and afterwards the prediction task. Both tasks were based on different object-exit associations. The prediction task comprised the following likely relations between Gabor figures and exits: vertical pattern – top exit, horizontal pattern – right exit and diagonal pattern – left exit. The reaction task comprised another relation: vertical pattern – right exit, horizontal pattern – left exit and diagonal pattern – top exit. In two thirds of the cases the color changed at the exits, viz. there were 54 go trials per block in the reaction task. The 100% condition was excluded to reduce the duration of the experiment and due to the fact that findings in the previous experiments were consistent. In Experiment I to III, only the prediction task allowed performance feedback, i.e. depicting the percentage of correct predictions at the end of every block, possibly influencing the learning process. Therefore, the performance feedback was no longer depicted in the current experiment due to the separation and comparison of the prediction and the reaction task. All test materials and instructions were translated into English language and checked by native speakers. Additionally, the validated English test version of the QCM was used to be able to compare data with previous studies of the experimental series. Before the experiment, participants had to perform an online version with 24 plates of the Ishihara Color Blindness Test (Ishihara, 1917) to ensure the ability to perceive color changes. After the experiment, participants completed the Concept Awareness Questionnaire for both tasks, the prediction and reaction task. The whole test session lasted up to three hours. Due to the technical equipment at the UBC, eye movements were recorded with the same eye tracker model, but with a lower sampling rate of 120 Hz instead of 500 Hz.

6.2.3 Data Analysis

Eye movement data was checked via heat maps for drifts which were corrected if necessary to ensure that data recordings were accurate. 4.2% of the trials of the prediction task were excluded from data analysis due to missing predictions and another 1.6% of the trials were excluded as less than 65% of the eye movement data points within the trial were valid. In the reaction task only go-trials were analyzed. Additionally, 2% of the trials in the reaction task were excluded from data analysis because of an insufficient number of valid eye movement data points. Data of the prediction and reaction task were separately analyzed. Two-way repeated measures ANOVAs with the within-subject factors *block* (1-4) and *judgment* (correct, incorrect) were run to analyze learning effects. In the reaction task only the factor block was investigated due to the missing prediction. Judgment time and reaction time were analyzed as dependent variable in the specific condition. Task performance, viz. the number of correct predictions was also used as dependent variable. Further, dependent variables were the earlier mentioned eye movement parameters: fixation frequency, fixation duration, number of gaze shifts and additionally, scanpath distance and velocity. To expand eye movement analyses, scanpath distance and gaze velocity were added to the former set of variables. Research already showed that longer scanpaths indicated less efficient searching (Ehmke & Wilson, 2007; Goldberg, Stimson, Lewenstein, Scott, & Wichansky, 2002). Gaze velocity used to measure cognitive arousal viz. the cognitive activation level (Holmqvist et al., 2011). These new measures might provide additional information, especially with regard to the level of arousal that might be relevant for distinguishing voluntary and automatic processes. The scanpath distance was calculated by adding up the length between all data points. In a next step, the scanpath length was divided by the trial duration and thus, the velocity (mm/s) is calculated. Saccadic velocity would be more informative as scanpath velocity depended largely on the scanpath distance and showed comparable results. However, the accurate detection of saccades in the current experiment might be problematic due to the lower sampling rate.

Planned *t*-tests were used to identify differences, for example, between the accuracy of the subjective probability concepts and between go and no-go trials in the reaction task. To identify differences in the visual search behavior between the reaction and the prediction task eye movement variables were used as within-subject variables of the repeated measures ANVOA. By means of hierarchical cluster analysis, participants were split into a high and low performer group. Data of participants who reported no understanding of the probability concept were also included in this analysis to increase sample size and the performance variance. Differences between these groups were investigated by using independent *t*-tests. Finally, confounding effects of interest and attention on task performance were analyzed by running correlation analysis and ANOVA.

6.3 Results

In the following, only significant (*p*<.05) results or trends (*p*<.10) were reported, except if the results were relevant for the aforementioned research questions.

6.3.1 Prediction Task

Task performance:

Generally, 54.3% of the trials were correctly predicted. There was a trend for a main effect of *block*, *F*(3,39)=2.74, *p*=.056, η_p^2=0.174, indicating that task performance increased across blocks (Fig. 6.4A).

Judgment time:

Analysis of *block* revealed a significant main effect, *F*(3,39)=4.20, *p*=.039, η_p^2=0.244, indicating a significant decrease of judgment times across blocks. However, there was no main effect of *judgment*, *F*(1,13)=1.21, *p*=.291, η_p^2=0.085 (Fig. 6.4B). Judgment time was analyzed in depth as in the previous experiment. Figure 6.3 shows the results of judgment times for correct and incorrect judgments and in each case for likely and unlikely exits. Data analysis showed no significant differences of judgment times for all pairings. Neither correct predictions for the likely exit (*M*=0.457, *SD*=0.168) and the unlikely exit (*M*=0.435, *SD*=0.216), nor incorrect predictions for the likely exit (*M*=0.478, *SD*=0.170) and the unlikely exit (*M*=0.519, *SD*=0.193) differed significantly (*p*>.17). However, at least on a descriptive level results seem to be similar to Experiment I, II and III. Interestingly, more participants predicted the unlikely exit correctly than in the previous experiments (*n*=10) so that this combination could be analyzed and depicted for the first time.

Fixation frequency:

We only observed a significant effect of *judgment*, *F*(1,13)=23.86, *p*<.001, η_p^2=0.647, indicating fewer fixations during correctly than incorrectly predicted trials. There was no significant effect of *block*, *F*(3,39)=1.89, *p*=.171, η_p^2=0.127, indicating no significant change (Fig. 6.4C).

Fixation duration:

There was neither a significant main effect of *block*, *F*(3,39)=1.32, *p*=.283, η_p^2=0.092, nor a significant main effect of *judgment, F*(1,13)=1.13, *p*=.308, η_p^2=0.080), indicating no significant change.

<u>*Number of gaze shifts:*</u>

We observed a main effect *block*, $F(3,39)=9.50$, $p<.001$, $\eta_p^2=0.422$, indicating that the number of gaze shifts decreased across blocks. We found also a main effect of *judgment*, $F(1,13)=57.26$, $p<.001$, $\eta_p^2=0.815$, indicating that participants showed less gaze shifts for correctly predicted than for incorrectly predicted trials (Fig. 6.4D).

<u>*Gaze velocity:*</u>

A main effect *block* was found, $F(3,39)=8.58$, $p=.008$, $\eta_p^2=0.398$, indicating a significant decrease of gaze velocity across blocks. There was only a trend for a main effect of *judgment*, $F(1,13)=4.29$, $p=.059$, $\eta_p^2=0.248$ (Fig. 6.4E).

<u>*Scanpath distance:*</u>

We observed a main effect *block*, $F(3,39)=9.12$, $p=.005$, $\eta_p^2=0.412$, indicating a significant decrease of scanpath distance across blocks. In addition, we observed a main effect *judgment*, $F(1,13)=6.69$, $p=.023$, $\eta_p^2=0.340$, indicating a shorter scanpath distance in correctly than incorrectly predicted trials (Fig. 6.4F).

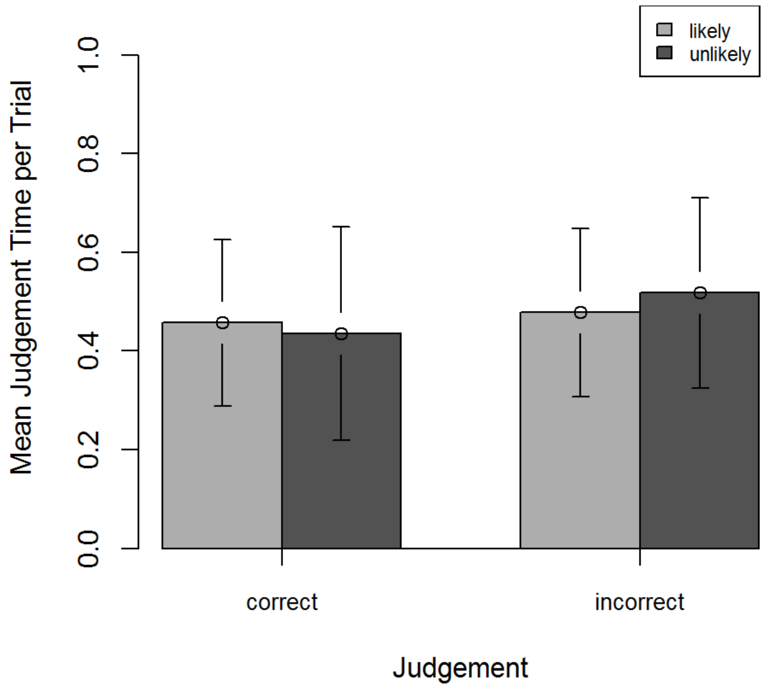


Figure 6.3: Judgment times in Experiment IV: Correct and incorrect judgments for likely or unlikely exits. Error bars depict the standard deviation.

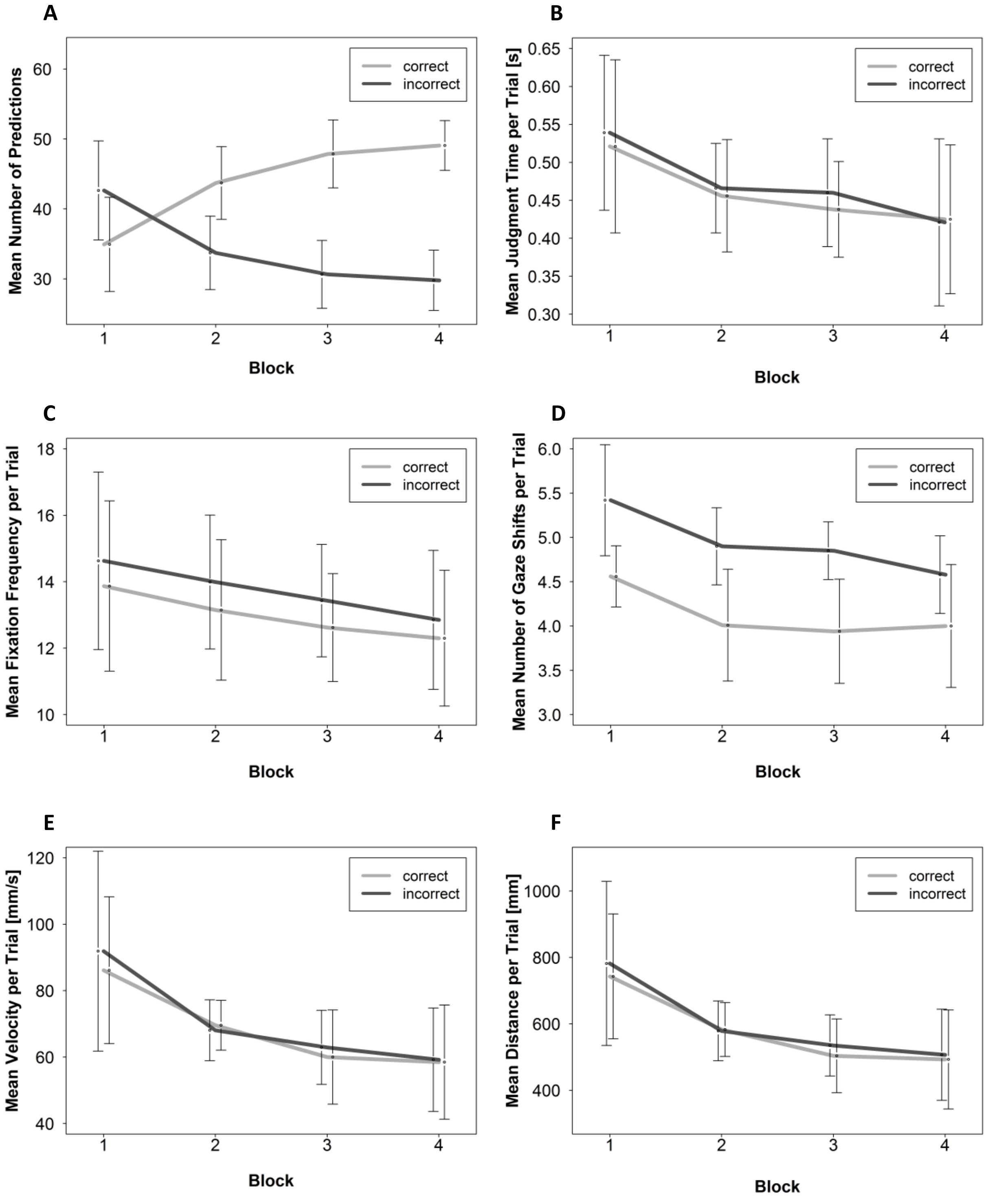


Figure 6.4: Results of the variables in the prediction task of Experiment IV. Number of predictions (A), judgment time (B), fixation frequency (C), number of gaze shifts (D), velocity (E) and scanpath distance (F) as a function of *block* and *judgment*.

Cluster Analysis:

Cluster analysis was used to group participants into low (n=9) and high performers (n=12) independent of the estimated subjective probability concept. Five participants of the overall 26 participants had to be excluded from the test due to technical issues. The grouped sample also included participants who reported no understanding of any object-exit association as they showed significantly less correct predictions per block (M_{Block}=23.00, SD_{Block}=3.19) than participants who reported an understanding of the underlying probability structure (M_{Block}=43.89, SD_{Block}=3.19; $t(19)$=4.758, p<.001). Finally, planned t-tests were used to test if differences in performances are reflected by eye movement parameters. As shown in table 6.1, high performers (M_{Block}=47.58, SD_{Block}=6.97), who made more correct predictions in the prediction task, shifted their gazes significantly more often between the AOIs than low performers (M_{Block}=22.72, SD_{Block}=2.04; $t(19)$=2.780, p=.012). However, other eye movement parameters, i.e. fixation frequency, fixation duration and scanpath distance indicated no significant differences between low and high performers (p>.26).

Table 6.1: Statistics of the Performance Cluster and Number of Gaze Shifts in Experiment IV

	Cluster	Sample size	Mean per Block	Standard Deviation
Number of Gaze Shifts	1	9	5.17	0.29
	2	12	3.83	0.40

Note. 1=low performers, 2=high performers

6.3.3 Reaction Task

Dependent variables:

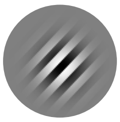
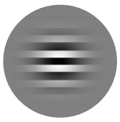
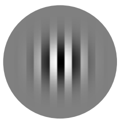
There was no main effect of *block* for reaction time, $F(3,39)$=2.49, p=.105, η_p^2=0.161. All eye movement variables showed no significant effects of *block* (p>.09). However, the comparison of go-trials and no-go trials in the reaction task showed significantly more gaze shifts in no-go trials than go-trials, $t(13)$=3.19, p=.007.

6.3.2 Comparison of the Prediction and the Reaction Task

Analysis of the subjective probability concept:

The subjective probability concept of the participants was measured after the completion of both tasks. Results showed that there was no significant effect of the order in which the tasks were performed (p>.09). However, the comparison of the subjective probability concepts showed that the likely estimations of the participants for all Gabor figures differed significantly between the prediction and reaction task: Gabor figures with diagonal lines, $t(13)=4.96$, $p<.001$, Gabor figures with horizontal lines, $t(13)=4.06$, $p<.001$, and Gabor figures with vertical lines, $t(13)=5.50$, $p<.001$, as shown in table 6.2. Results he Concept Awareness Questionnaire, which was completed after the two tasks, showed that likely exits in the prediction task were accurately recognized whereas in the reaction task participants estimated almost the same tendencies of the probability concept as in the prediction task. It seems that participants were not able to recall the probability concept of the reaction task and thus, transferred the concept presented during the prediction task also to the reaction task.

Table 6.2: Memory representation of the probability concept of the prediction and reaction task in Experiment IV

Gabor figures	Exit	Subjective Probability Concept - **Prediction**	Subjective Probability Concept - **Reaction**
	left	**70 % (21.9)**	55 % (25.9)
	top	15 % (12.3)	**25 % (17.4)**
	right	15 % (11.3)	20% (11.3)
	left	16 % (11.6)	**26 % (15.7)**
	top	20 % (16.7)	22% (12.5)
	right	**64 % (23.4)**	52% (25.7)
	left	15 % (11.3)	18 % (11.1)
	top	**70 % (21.3)**	60 % (23.8)
	right	15 % (10.3)	**22 % (15.8)**

Note. The object-exit associations with higher probabilities are shown in bold. Values in brackets show the standard deviation.

Eye movements:

For an additional comparison of eye movement parameters between the two tasks differences across blocks were analyzed. The analysis of the within-subject factor *task* showed no significant difference between fixation duration in the prediction and the reaction task. However, scanpath velocity, $F(3,78)=9.56$, $p=.001$, $\eta_p^2=0.269$ (Fig. 6.5A), scanpath distance, $F(3,78)=10.03$, $p<.001$, $\eta_p^2=0.278$ (Fig. 6.5B), and fixation frequency, $F(3,78)=3.11$, $p=.050$, $\eta_p^2=0.107$ (Fig. 6.5C), showed a significant interaction of the factor *block* and *task* in the way that eye movement parameters decreased across blocks during the prediction task but rather remained unchanged during the reaction task. Furthermore, results showed a significant interaction of the factor *block* and *task* with regard to the number of gaze shifts, $F(3,78)=10.03$, $p<.001$, $\eta_p^2=0.278$ (Fig. 6.5D), and significantly more gaze shifts during the performance of the reaction task than the prediction task, *t*(13)=4.560, *p*=.001. The number of gaze shifts did not change significantly across blocks during the performance of the reaction tasks, but decreased during the performance of the prediction task (see Appendix 9.6 for more details).

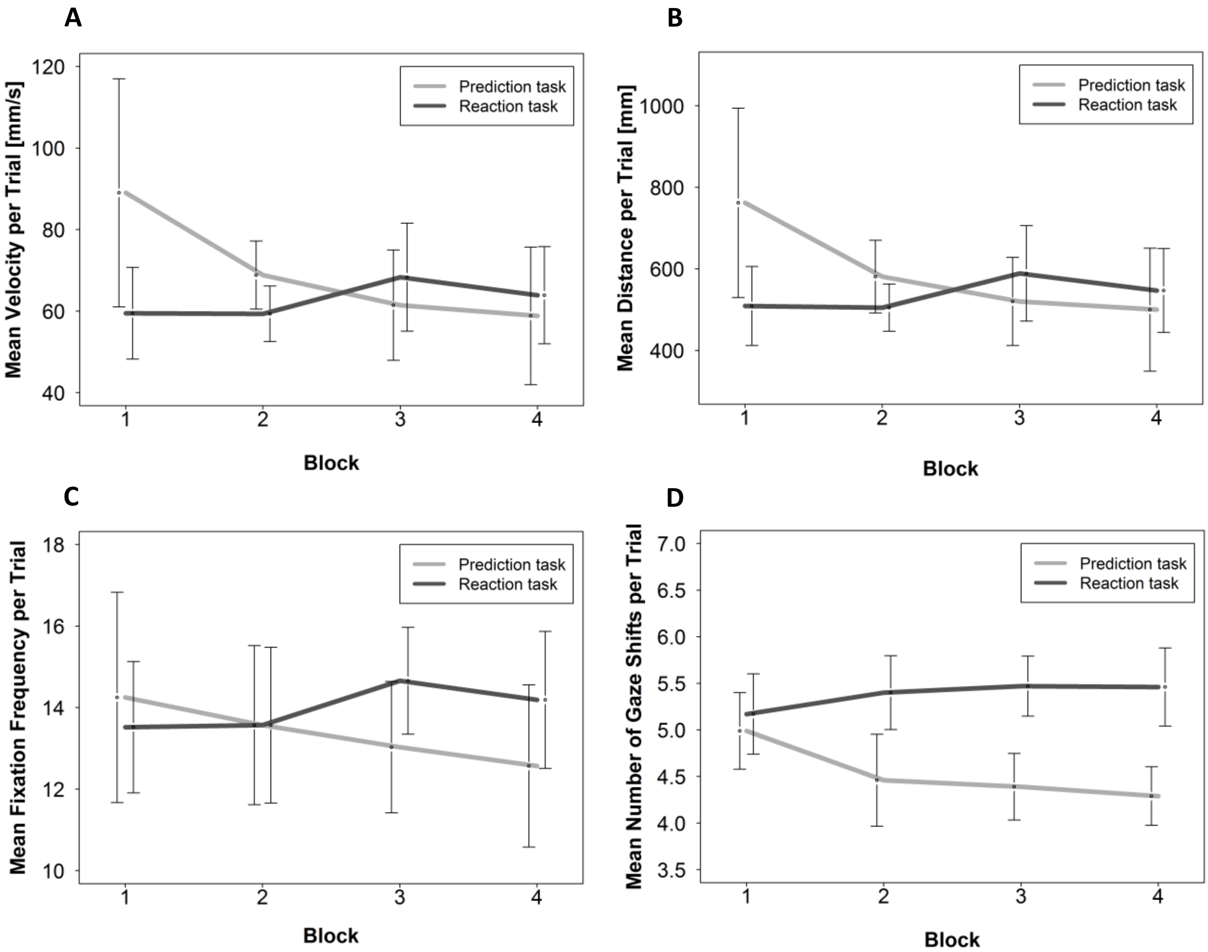


Figure 6.5: Comparison of eye movement parameters in the prediction and the reaction task in Experiment IV. Results of scanpath velocity (A), scanpath distance (B), fixation frequency (C) and number of gaze shifts (D) in the prediction and the reaction task as a function of *block* and *task.*

<u>*Analysis of Control Variables:*</u>

As reported in all previous experiments as well, results of the D2 test showed a significant increase of selected cases after running the experiment, *t*(13)=6.48, *p*<.001, presumably due to learning effects. Participants selected on average 60% (*SD*=15.68%) of the cases before and 70% (*SD*=16.87%) after performing the OVSST. In addition, the error rate did not differ between pre- and post-test (*p*>.30). Thus, attention seems not to be affected by the OVSST.

Analyzing the subscale interest of the QCM (*M*=3.4; *SD*=1.3) showed a significant one-tailed correlation (*r*=562, *p*=.018) with task performance by using Spearman's rank correlation coefficient. Participants who were interested in the task seem to make more correct predictions than participants who were less interested in the task. Planned *t*-test analyses showed no differences between the Canadian sample of Experiment IV and the German samples from Experiment I to Experiment III (interest: M=3.4, SD=1.1; D2 pre-test: M_{total}=188.48, SD_{total}=35.23; D2 post-test: M_{total}=210.61, SD_{total}=39.11) concerning the number of cases selected in the pre- and post-test of the D2 and interest (p>.24).

6.4 Discussion

The aim of Experiment IV was to assess the effect of the prediction and the reaction task on the development of the mental model in more detail. For this reason, the prediction and reaction task of the OVSST were separated and tested respectively. The subjective probability concept for the prediction task was quite accurate whereas it was not accurate for the reaction task. Thus, the mental representation of the participants developed during the performance of the OVSST seems to be mainly based on the prediction task. Eye movement parameters fostered this assumption as they reflected learning effects in the prediction task as well as a lack of learning in the reaction task. Furthermore, reaction times did not decrease significantly over blocks during the performance of the reaction task providing no support for implicit learning. These results rather confirmed the assumption that top-down processes and bottom-up processes are involved during the performance of the prediction and the reaction task.

Mental model development:

The results of the subjective probability concept indicated that participants obviously developed a more accurate mental representation of the prediction task than the reaction task. For the reaction tasks, even likely exits of the probability concept were incorrectly recognized for all object-exit associations. The learned probability concept in the prediction task seemed to be transferred to the reaction task indicated by similar probability estimations in the Concept Awareness Questionnaire. The reason for this finding might be deeper cognitive processing in the prediction task due to the decision making process that relied on working memory involvement. As participants were presumably not able to retrieve any probability concept for the reaction task, they used the probability concept of the prediction task for the estimation of the probabilities.

As mentioned in the introduction of Experiment IV, Evans and Stanovich (2013) discussed in their article dual-process theories and found support for a fast intuitive Type 1 process for default responses and a slow reflective Type 2 process for a higher cognitive processing and reasoning. The former process requires no working memory, but the latter one does. The results of Experiment IV seemed to corroborate this assumption. The prediction task based on decision making between three choices involving higher cognitive processes, and thus, resulted in a conscious knowledge about the underlying concept. This task seemed to be a reflective Type 2 process whereas the reaction on salient stimuli in the reaction task might be rather an intuitive Type 1 process resulting in an inaccurate mental representation of the underlying probability structure as responses relied on simple change detection. In the same way, the prediction task might be rather goal-driven and top-down controlled whereas the reaction task might be stimulus driven and bottom-up controlled. This assumption could be further reinforced by the findings of Orquin et al. (in press). In their article, they suggested that predictable locations increased top-down control whereas unpredictable locations decreased top-down control. This might be comparable with the learning of the probability concept in the prediction task and the missing learning during the performance of the reaction task. Nevertheless, there also seemed to be plausible arguments that not all cognitive processes could be classified into bottom-up and top-down processes, for example, if neither physical salience nor current goals are provided (Awh, Belopolsky, & Theeuwes, 2012; Evans & Stanovich, 2013). For an alternative framework discussing this issue cf. Awh et al. (2012).

Effects on cognitive processing:

The differences in the development of the mental representation was also indicated by judgment times and reaction times. Participants might be able to learn the probability concept explicitly on the basis of prediction tasks. They, however, might also be able to learn the probability concept implicitly on the basis of the reaction task. These benefits had to be apparent in decreasing reaction times. However, the current results showed no significant decrease in reaction times across blocks. In contrast, judgment times decreased significantly across blocks during the performance of the prediction task in Experiment IV as well as in all previous experiments. These findings reinforced the assumption that the prediction task enhanced information processing and allowed to build up an accurate mental representation of the probability concept whereas the task characteristics of the reaction task seemed not to support the mental model development.

Further, judgment times could also be analyzed in detail for correctly predicted unlikely exits in comparison to likely exits for the first time as this combination occurred for enough participants. However, results were not significant and previous findings indicating longer judgment times for unlikely than for likely exit predictions could not be replicated, presumably due to the small sample size. Nevertheless, the actual reasons for these results remained open. Descriptive statistics might rather suggest that judgment times for correctly predicted unlikely exits were faster than all other combinations. The reason for this result, however, remained unclear.

Visual search behavior:

It was expected that eye movement parameters indicated different cognitive processing in the reaction and the prediction task. In fact, the time course of scanpath velocity, scanpath distance, fixation frequency and number of gaze shifts reflected the learning process of the probability concept in the prediction task and the lack of learning in the reaction task. However, the different processes of the prediction and the reaction task seemed not to be reflected by fixation duration, presumably due to the same experimental sequence for both tasks resulting in the same task complexity of the environment. In contrast, other authors already found longer fixation durations with higher task complexity (cf. Horstmann, Ahlgrimm, & Glöckner, 2009; Velichkovsky, Rothert, Kopf, Dornhöfer, & Joos, 2002; Venkatraman et al., 2014).

Furthermore, the number of gaze shifts was significantly higher in the reaction task than in the prediction task. The higher number of gaze shifts between the AOIs in the reaction task might reflect the misunderstanding of the probability concept according to the general finding that low performers showed more gaze shifts than high performers in the prediction task. Participants performing the reaction task might not be able to anticipate the correct exit and thus, had to shift their gaze more often to detect the target and to react to color changes. This might be due to the missing integration of feedback. It was not necessary to learn the object-exit associations for the accurate performance of the reaction task. Interestingly, in the reaction task participants showed more gaze shifts in no-go trials than in go trials. One possible reason for this might be that salient stimuli in go trials guided attention to the target stimulus whereas in no-go trials participants did not need to focus further on the target and thus might move their gaze randomly on the screen waiting for the start of the next trial. However, the actual reasons for this effect were unclear. Overall, it seemed that eye movements reflect the learning processes.

Gaze shifts between the AOIs were also informative in the prediction task as less gaze shifts indicated better task performance. These results replicated the findings of Experiment I even if no reaction had to be performed. Further, participants seemed to anticipate the predicted exit and shifted their gaze to the stimulus if it reappeared at another exit in order to get performance feedback. In addition, scanpath distance decreased significantly across blocks in the prediction task indicating more efficient scanning of the display across blocks (Ehmke & Wilson, 2007; Goldberg et al., 2002; Goldberg & Kotval, 1999). Thus, one might argue that the OVSST also works without the second reaction task as eye movement parameters are still conclusive. However, fixation frequency and fixation duration showed no effects concerning the factor *block* in this separated task condition and the reaction task did not seem to affect the mental model development. Due to these missing findings and reasons of comparability with Experiment I to III, it was plausible to maintain the combined-task structure of the OVSST for the next experiment.

If the prediction and the reaction task of the OVSST could be seen as a dual-process, slower eye movements were expected for the prediction task and vice versa faster eye movements were expected for the reaction task according to van Zoest et al. (2004). However, this is not the case for the data set of the current sample. Scanpath velocity did not differ significantly between both tasks presumably again due to the identical experimental sequence of the OVSST used for both tasks.

Sample-specific features:

Besides data analysis, it was noteworthy that one third of the tested participants did comprehend any probability structure and thus, were excluded from data analysis except for the cluster analysis. There might be different reasons why this default rate was higher than in the experiments before. First, there was no performance feedback in the current version of the OVSST that might motivate participants. Second, studies in the testing environment of the University of British Columbia usually did not last longer than one and a half hour. Thus, participants often reported that the experiment was too long and exhausting. This might additionally lead to less motivation and further to less effort reporting the subjective probability concept, presumably not reported in the motivational questionnaire due to social desirability. Third, participants might report a missing understanding of any object-exit association solely based on the reaction task. As they were not able to acquire an accurate mental representation of the concept used in this task, they probably generalized this state and finally, reported a missing understanding in general.

Influence of confounding variables:

The findings of the current experiment showed that attention did not seem to affect learning results. Thus, the findings of all previous experiments could be replicated. In accordance with Experiment II, but in contrast to Experiment I and II, results revealed a positive effect of motivation on task performance. As mentioned before, motivation did not seem to systematically influence task performance. However, motivation was also used as control variable in the next experiment due to reasons of consistency.

Further research:

Another aspect mentioned in the general introduction might be the degree of uncertainty influencing visual search behavior and the development of the mental representation that was studied in the next experiment. The OVSST enabled to vary the degree of objective uncertainty by using lower and higher probabilities of the probability structure.

6.5 Conclusion

Experiment IV showed that participants developed only an accurate mental representation of the probability concept presented in the prediction task but not in the reaction task. These different learning processes in both tasks were reflected by eye movements, namely number of gaze shifts, fixation frequency, scanpath distance and scanpath velocity. In conclusion, mental representations acquired during the performance of the OVSST seem to be mainly based on the prediction task. It seems that both tasks require different cognitive processing, which fits into the dual-process assumption of attention: the reaction task seems to be an intuitive Type 1 process that is bottom-up controlled whereas the prediction task is rather a reflective Type 2 process that is top-down controlled and results in conscious knowledge participants are able to report.

7 Experiment V – Learning Different Probabilities

7.1 Introduction

In Experiment III and IV, distinct eye movement parameters reflected the learning processes during relearning and performing the prediction task as well as the lack of learning during the performance of the reaction task. Parallel to the learning curve eye movement parameters mainly decreased indicating also a reduced subjective uncertainty. However, the degree of objective uncertainty was kept constant in all previous experiments. Another point of concern mentioned before (see Introduction) addresses the question whether the degree of uncertainty influences eye movement parameters in such a way that the underlying shifts in learning processes become visible. Thus, in the last experiment of the experimental series different degrees of objective uncertainty were investigated which can be manipulated by adjusting the probability structure of the OVSST.

A basic model in decision making under uncertainty, as described earlier, is the expected utility theory proposed by Neumann and Morgenstern (1947). This theory assumes that probabilities of the outcomes are known, however, this is generally not the case in real-life. In contrast, the subjective expected utility theory (SEU) by Savage (1954) assumes that people chose the option which maximizes the subjective expected utility. However, it is also problematic to identify the subjective utility of the decision maker as humans often use fast and frugal heuristics to make their decisions. The way we perceive and evaluate the information can also influence the decision making (Gigerenzer & Goldstein, 1996). Furthermore, findings of several studies fail to find supporting evidence for the SEU theory (Slovic, Fischhoff, & Lichtenstein, 1977). Fishburn (1970) extends the SEU and also considers decision strategies and consequences of the decision maker. This last-mentioned extension of the SEU is highly relevant for the OVSST since the task also requires to develop decision strategies and reflects consequences of the decision between the three target objects. Further, during the performance of the OVSST all steps are involved in the decision process, i.e. from the identification of the decision situation to the final feedback after making the decision during the continuous processing in the working memory (see Fig. 1.2).

The OVSST, however, is not comparable with typical real-life situations, as no prior knowledge with regard to the probability distribution exists at the beginning. At first, participants have to encode the presented implicit information and develop a strategy to cope with the uncertain situation. They can only refer to prior experiences gained during the trainings session which does not included different probabilities of the OVSST. Participants acquire knowledge over time that is used for the mental model development and also influences the optimal degree of memory updating and exploration (cf. Doya, 2008). For instance, it is better to ignore the rare occurrence of the object reappearance at the bottom entrance, than trying to integrate this situation in the response strategy as reported in Experiment I (Chapter 3). Thus, the ignorance of irrelevant information and the focus on relevant information has to be learned and anchored in the mental representation.

In previous work, different degrees of uncertainty manipulated by probabilities were already investigated. In the experiment of Shaw and Shaw (1977) participants had to search and identify single target letters whereby some locations had a lower and a higher probability of occurrence. The authors already showed that participants were able to learn associations between target objects and high probability locations and respond more efficiently relative to low probability associations. Richer and Beatty (1987) reported in their study that reaction times increase with response uncertainty during the performance of two-choice and four choice tasks with go and no-go responses. These studies showed that different degrees of objective uncertainty influenced learning and behavioral data. Therefore, it seems to be necessary to examine in which way previous findings can also be applied to a higher degree of uncertainty. The following research questions address eye movements and judgment times related to lower and higher uncertainty and were derived from the aforementioned literature and presented data in the previous experiments.

7.1.1 Research Question and Hypotheses

In the current experiment, we compared the task performance in a high probability condition, used in the previous experiments, with task performance in a low probability condition. Two pilot studies were used to determine the threshold of lower probabilities which participants were still able to discriminate. Results of the first pilot study showed that Gabor figures with lines caused an object bias in the way that only the preferred exit of this object was learned correctly. Thus, Gabor figures were improved once again and counterbalanced across all participants to avoid biases. In the final experiment participants performed the OVSST twice, in one session they had to perform the probability concept used in the previous experiments and in another session a concept with higher uncertainty. The initial uncertainty of the participant due to the innocence of the experimental concept was reduced as participants were explicitly instructed to learn the probability concept.

Three main expectations can be proposed: First, it is supposed that that prediction accuracy is reduced in the low probability condition due to the higher degree of uncertainty. Even if participants use the optimal decision strategy, TTB, task performance should be reduced and learning times extended, because it takes longer to detect the best decision strategy since more evidence accumulation steps are necessary in order to build a realistic mental model of the probability structure. Second, it is expected that judgment times in the low probability condition are longer according to Richer and Beatty (1987), because of the higher degree of uncertainty and thus different processing demands. Finally, a higher degree of objective uncertainty during decision making would probably lead to more ambiguous information and thus to a more widespread use of coping strategies. Thus, it is expected that higher uncertainty in the low probability condition is reflected by more visual search behavior to reduce uncertainty according to Lipshitz and Strauss (1997).

7.2 Pilot Studies

In research two-choice problems (cf. Mattes, Ulrich, & Miller, 2002; Miller, 1998) with an underlying probability concept were more often considered in detail than multiple choice or in this case three-choice problems (cf. Swensson, 1965). Thus, there were no hints from earlier studies for an appropriate reduction of the probability distribution. In the following, two pilot studies were run to estimate which lower probabilities participants are able to learn.

7.2.1 Participants

5 students (3 female) with mean age 24 years (*SD*=3 years) participated in the first pilot study and 7 participants, 6 of them students, with mean age 24 years (*SD*=4 years) were tested in the second pilot study. All were dominantly right handed and had normal vision without glasses or contact lenses.

7.2.2 Procedure

In a first pilot study, participants had to learn a probability concept within six blocks (81 trials each) of the last version of the OVSST (see Chapter 6.2.2) with 52% higher probability and 22% lower probabilities in one session and in another session a second concept with 44% higher probability and accordingly 26% lower probabilities. Six blocks were chosen because no previous experience could be used to indicate the course of the learning curve, however, the accumulation of evidence might take more time under higher uncertainty than lower uncertainty. The procedure was the same as in Experiment IV and thus the 100% condition was excluded but a trainings session had to be performed before the experimental session. In contrast to the previous experiments, participants had to attend two separated test sessions on different dates due to reasons of time as 6 blocks had to be performed lasting 1 1/2 hours in total. In the first test session participants had to learn one probability distribution and in the second test session the other probability distribution. The sequences were counterbalanced, i.e. participants alternately started with the 52-22-22 probability distribution or the 44-26-26 probability distribution. Participants were initially instructed to learn a probability concept and they were asked to complete the Concept Awareness Questionnaire at the end of each session. Thus, initial uncertainty was reduced by clarifying the aim of the task.

As probabilities in the first pilot study seemed to be too low to be able to learn, participants in the second pilot study had to learn a probability concept with 59% higher probability and 18.5% lower probabilities in one sessions (lower probability concept) and the former concept with 74% higher probability and 11% lower probabilities in another session (higher probability concept). The formerly used high probability concept was included to test a setting for the final experiment with a lower and a higher probability distribution. The higher probability distribution was tested only four blocks because all previous experiments showed that learning reached a saturation after four blocks for this probability distribution. The sequences of the probability concepts were again counterbalanced. Furthermore, new objects were created as results of the first pilot study indicated an object bias (Fig. 7.1). As lines comprise directional information, new patterns without any directional information were chosen. All other conditions were the same as in the first pilot study.

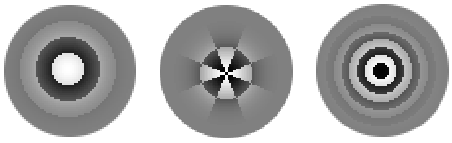


Figure 7.1: The OVSST with a new set of Gabor figures in Experiment V. Gabor figures with non-directional information were used in the second pilot study and in the final version of Experiment V.

7.2.3 Data Analysis

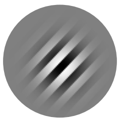
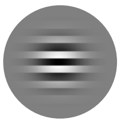
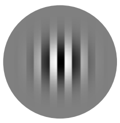
Before analyzing data statistically, eye movement data were checked for drifts on the basis of heat maps. 4.4% of the trials were excluded from data analysis in the first pilot study due to missing judgments and additionally, 0.4% of the trials were excluded as less than 65% of the eye movement data of the trial were valid. In the second pilot study only 2.1% of the trials were excluded due to missing judgments and 0.1% of the additional trials were excluded due to less than 65% of valid eye movement data. The subjective probability concept measured by the Concept Awareness Questionnaire was compared to the objective values. Therefore, the prediction error was calculated (the deviation from the target value) and tested against zero by using one-sample *t*-tests.

7.2.4 Results

Pilot Study 1:

Even if participants were initially instructed to learn a probability concept, 3 of 5 participants reported to recognize the 52-22-22 probability concept and 2 of 5 participants could report the 48-24-24 probability concept. Error rates for both probability concepts were high: 75% (52-22-22 probability distribution) and 76% (48-24-24 probability distribution) of the trials were incorrectly predicted. In total, all object-exits associations estimated by the participants via the Concept Awareness Questionnaire were not accurate except for the Gabor figures with horizontal lines (Tab. 7.1). The prediction error showed significant deviations from the target values for the 52-22-22 probability distribution (p<.04) as well as for the 48-24-24- probability distribution (p<.04). The association between the top exit and the Gabor figure with vertical stripes might indicate an object bias. Post-hoc correlation analysis also supported the assumption of an object bias: results showed a significant correlation (*r*=.040, *p*=.006) of object types (diagonal, vertical and horizontal stripes) and predictions (correct and incorrect). Thus, it was recommended to use different objects in the second pilot study and to test another low probability distribution which might be more recognizable, i.e. lower and higher probabilities should be distinguishable in a clearer manner.

Table 7.1: Memory representation of the probability concept of two lower probability distributions in Pilot Study I

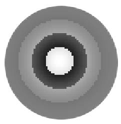
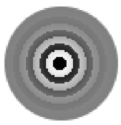
Gabor figures	Exit	Subjective Probability Concept: 52-22-22	Subjective Probability Concept: 48-24-24
	left	**37 % (18.0)**	29 % (8.9)
	top	19 % (9.5)	41 % (16.6)
	right	44 % (20.3)	**30 % (12.7)**
	left	33 % (8.3)	35 % (3.2)
	top	35 % (22.9)	**31 % (3.5)**
	right	**33 % (14.8)**	34 % (3.6)
	left	33 % (22.0)	**25 % (10.2)**
	top	**48 % (21.4)**	41 % (19.4)
	right	19 % (8.9)	34 % (15.4)

Note. The object-exit associations with higher probabilities are shown in bold. Values in brackets show the standard deviation.

<u>*Pilot Study 2:*</u>

In the second pilot study, all participants stated that they understand the higher probability concept (74-11-11 probability distribution) but only 4 of 7 participants confirmed to understand also the lower probability concept of the 59-18.5-18.5 probability distribution. 54% of the trials were incorrectly predicted in the high probability condition and 64% in the low probability condition. Results of the subjective probability concept showed no highly significant prediction error for the high probability concept (p>.08) but marginal signficant to signficant effects (.02 > p <.07) for the low probability distribution (Tab. 2). Nevertheless, the low probabiltiy distribution of the second pilot study (59-18.5-18.5) showed correct tendencies with regard to the estimated probabilites and thus seems to be more likely to learn than both lower probability distributions in the first pilot study. This time, correlation analysis showed a non-significant correlation (r=.008, p=.273) of the new object types (few circles, many circles and symmetric pattern) and predictions (correct and incorrect). Thus, the newly designed objects seem to be more suitable for the final experiment.

Table 7.2: Memory representation of the probability concept of two lower probability distributions in Pilot Study II

Gabor figures	Exit	Subjective Probability Concept: 59-18.5-18.5	Subjective Probability Concept: 74-11-11
	left	**53 % (28.0)**	19 % (23.2)
	top	36 % (27.6)	**65 % (24.2)**
	right	12 % (7.2)	16 % (8.5)
	left	19 % (4.2)	11 % (6.1)
	top	**62 % (17.5)**	25 % (21.6)
	right	19 % (14.6)	**64 % (22.8)**
	left	20 % (19.0)	**70 % (25.0)**
	top	24 % (20.1)	11 % (4.0)
	right	**56 % (32.1)**	19 % (22.9)

Note. The object-exit associations with higher probabilities are shown in bold. Values in brackets show the standard deviation.

7.2.5 Discussion & Conclusion of the pilot studies

Two pilot studies were used to investigate which lower probability distributions are able to be learned or too close to chance. In total, three different lower probability distributions were tested. In the first pilot study participants had to learn a 52-22-22 and a 48-24-24 probability distribution. Results showed a high error rate with regard to the correctness of the predictions and highly significant prediction errors for the estimated probabilities via the Concept Awareness Questionnaire. Thus, participants did not seem to learn the probability distributions. In the second pilot study, a 59-18.5-18.5 probability distribution was tested which seems to be easier to learn than the probability distributions in the first pilot study. Finally, this probability distribution was chosen for the final experiment.

Besides the learnability, an object bias was found in the first pilot study. This bias was probably not existing in Experiment IV while using the same Gabor figures due to the larger sample size and the lower uncertainty. The newly developed Gabor figures in the second pilot study with no directional information seem to be less confounding and thus were also used in the final experiment.

In conclusion, the setting used in the second pilot study worked well and can be applied in the final experiment. In more detail, the 59-18.5-18.5 probability distribution should be used for low probability concept and Gabor figures with few circles, many circles and a symmetric pattern as target objects.

7.3 Method

7.3.1 Participants

26 students participated in the study at the IfADo. Two participants had to be excluded due to technical issues and language barriers. Finally, 24 participants (14 female) with mean age 24 years (*SD*=3 years) entered data analysis. All of them had a dominant right hand and no vision impairment.

7.3.2 Procedure

Participants in Experiment V had to performer a 59-18.5-18.5 lower and a 74-11-11 higher probability condition due to the results of the pilot studies. The procedure was the same as in the second pilot study. In general, there were three main differences of Experiment V in comparison to the previous experiments. First, participants had to learn two different probability concepts separately each within a test session. Thus, participants had to perform two sessions on different dates. The high probability concept lasted four blocks as in the previous experiments whereas the low probability concept lasted six blocks as few information was available about how much time participants need to learn a concept with a lower probability concept. At the end of each session participants had to complete the Concept Awareness Questionnaire for the learned concept. The sessions were counterbalanced in the way that all participants with even number started with the lower probability concept and all participants with odd numbers started with the higher probability concept. Second, for the first time participants were explicitly instructed that every object is associated with the exits to a distinct probability to control prior knowledge. Participants had to complete the Concept Awareness Questionnaire after each session and would be biased concerning a possible probability concept in the second session. Therefore, we decided to inform them about the presentation of a probability structure. Thus, the initial uncertainty regarding the task was reduced. Third, all object-exit associations were counterbalanced across participants in order to avoid potential confounding effects caused by an object bias.

As the findings of the D2 attention test were consistent in all of the previous experiments, this test was excluded in the current experiment. Instead, the “Inventar zur Messung der Ambiguitätstoleranz” by Reis (1996) was used at the end of the experiment to control for ambiguity tolerance which might influence the learning process. Tymula et al. (2012) already reported in their study that participants with a higher tolerance for ambiguity, i.e. options with consequences inhering unknown probabilities, showed increased risk-seeking behavior. Thus, participants' performance might also be influenced by tolerance of ambiguity in the current experiment. Participants filled in the questionnaire with 40 items on a scale ranging from 1 ("trifft sehr zu", "I strongly agree") to 6 ("trifft gar nicht zu", "I strongly disagree"). This inventory contains five subscales measuring ambiguity tolerance for problems that seems to be unsolvable (1), ambiguity tolerance for social conflicts (2), ambiguity tolerance of the parents' image (3), ambiguity tolerance for stereotyped roles (4) and ambiguity tolerance for new experiences (5). The internal consistency for these subscales lies between Cronbach's alpha α = .74 and α = .86 and for the full scale α =.87.

7.3.3 Data Analysis

Before analyzing data statistically, again systematical drifts in eye movement data were checked and drifts were found for the first time. Thus, a correction was necessary by calculating the mean of the scatter plot per block and then, centering the scatter plot. 2.2% of the trials were excluded from data analysis of the high-probability concept due to missing predictions and additionally, 0.5% of the trials were excluded as less than 65 % of the eye movement data of the trial were valid. In the low-probability concept condition, 1.9% of the trials were excluded as predictions were missing and additionally, 0.2% of the trials had to be excluded due to insufficient validity of the data (less than 65%). Distinct AOIs were defined as shown in table 7.3. These additional classifications of AOIs were necessary due to the counterbalancing.

Table 7.3: AOI classifications and definitions in Experiment V

AOI	**Definition**
AOI_{target}	AOI around the exit at which the target object appears
$AOI_{predict}$	AOI around the exit which is predicted by the participant
AOI_{low}	AOI around the exits with the lower probability
AOI_{high}	AOI around the exit with the higher probability

The further analyses are comparable with those in the previous experiments. Two-way repeated measures ANOVAs were calculated with the within-subject factors *block* (1-4) and *judgment* (correct, incorrect) for each condition in order to analyze the effect of the learning development. Dependent variables were fixation frequency, fixation duration, number of gaze shifts, scanpath length, saccadic velocity as well as judgment time and reaction time. Now, saccadic velocity could be calculated as eye movements were recorded with a higher sampling rate of 500 Hz. Additionally, task performance, viz. the number of correct predictions was used as dependent variable. ANOVAs were also run for the newly defined AOIs to analyze the participants visual search behavior within the distinct AOIs. For this reason, fixation duration, fixation frequency and a new variable, the dwell time were analyzed. The dwell time is the time participants spend in the AOI independent of the fixations (see Chapter 1). For the comparison of both conditions, the high and low probability concept, planned *t*-tests were run as well as repeated measures ANOVA with the between factor probability condition.

Stimuli used in the experiments (Experiment I-V) were gradually developed due to object biases. The type of object (geometric figures, Gabor figures with lines, Gabor figures symmetric patterns) as well as the background (patterned vs. single-colored) differed in the experiments. Planned *t*-tests were used to investigate the influence of the stimulus presentation on the subjective probability concept, i.e. the prediction error, eye movements, i.e. fixation frequency, as well as performance. Finally, correlation analyses were used to test the impact of the confounding variables on task performance.

7.4 Results

In the following, only significant (p<.05) results or trends (p<.10) were reported, except if the results were relevant for the aforementioned research questions.

7.4.1 High Probability Concept

<u>*Task performance:*</u>

Analysis of *block* revealed a significant main effect, $F(3,69)=10.39$, $p<.001$, $\eta_p^2=0.311$, indicating a significant increase of task performance increased across blocks (Fig. 7.6).

<u>*Judgment times:*</u>

We observed a main effect of *block*, $F(3,69)=9.69$, $p=.001$, $\eta_p^2=0.296$, suggesting a significant decrease of judgment times across blocks. In addition, we found a main effect of *judgment*, $F(1,23)=6.25$, $p=.020$, $\eta_p^2=0.214$, indicating shorter judgment times during correctly than incorrectly predicted trials (Fig. 7.3A). A detailed analysis of judgment time showed, in line with the previous experiments, that the prediction of unlikely exits (correct: $M=0.551$; $SD=0.145$, incorrect: $M=0.573$; $SD=0.273$) was generally slower than the prediction of likely exits (correct: $M=0.436$; $SD=0.197$, incorrect: $M=0.437$; $SD=0.195$) as depicted in Figure 7.2. The effects were partly significant for judgment times during incorrect likely and incorrect unlikely predictions, $t(13)=2.66$, $p=.020$, and during correct likely and incorrect unlikely predictions $t(13)=2.41$, $p=.032$.

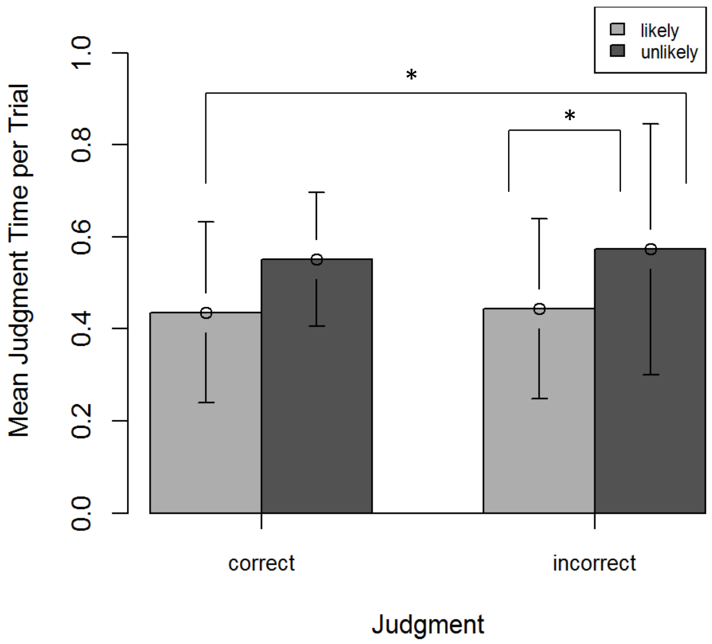


Figure 7.2: Judgment times in Experiment V in the high probability condition: Correct and incorrect judgments for likely or unlikely exits. Error bars depict the standard deviation.

Reaction time:

Analysis of *block* revealed a significant main effect, $F(3,69)=2.57$, $p=0.096$, $\eta_p^2=0.100$, suggesting a significant decrease of reaction times across blocks. In addition, there was a main effect of *judgment*, $F(1,23)=140.85$, $p<.001$, $\eta_p^2=0.860$, suggesting shorter reaction times during correctly than incorrectly predicted trials (Fig. 7.3B).

Fixation frequency:

We did not observe a main effect of *block*, $F(3,69)=0.25$, $p=.748$, $\eta_p^2=0.011$, indicating no change of fixation frequency across blocks. However, we observed a main effect of *judgment*, $F(1,23)=13.34$, $p=.001$, $\eta_p^2=0.367$, indicating fewer fixations during correctly than incorrectly predicted trials (Fig. 7.3C)

Fixation duration:

There was neither a significant main effect of *block*, $F(3,69)=0.72$, $p=.484$, $\eta_p^2=0.030$, nor a significant main effect of *judgment*, $F(1,23)=1.10$, $p=.306$, $\eta_p^2=0.046$, indicating no significant change. However, there was a trend for a main effect of *block* in AOI_{low}, especially for incorrectly predicted trials, $F(3,69)=2.91$, $p=.084$, $\eta_p^2=0.112$ (see Appendix C, Table 9.6 for details).

Number of gaze shifts:

A main effect *block* was found, $F(3,69)=6.58$, $p=.005$, $\eta_p^2=0.222$, indicating a significant decrease of the number of gaze shifts across blocks. We found also a main effect of *judgment*, $F(1,23)=86.70$, $p<.001$,

η_p^2=0.790, indicating that participants showed less gaze shifts for correctly predicted than for incorrectly predicted trials (Fig. 7.3D).

Gaze velocity:

Analysis of *block* revealed a significant main effect, $F(3,69)=4.04$, $p=.046$, η_p^2=0.149, indicating a significant decrease of gaze velocity across blocks. There was no significant main effect of *judgment*, $F(1,23)=0.08$, $p=.781$, η_p^2=0.003, indicating no difference between gaze velocity in correctly and incorrectly trials (Fig. 7.3E).

Scanpath distance:

We observed a main effect of *block* $F(3,69)=3.89$, $p=.050$, η_p^2=0.145, indicating that the distance if the scanpath decreased across blocks. However, we observed no main effect *judgment*, $F(1,23)=0.18$, $p=.673$, η_p^2=0.008, indicating no difference between scanpath distance in correctly and incorrectly trials (Fig. 7.3F).

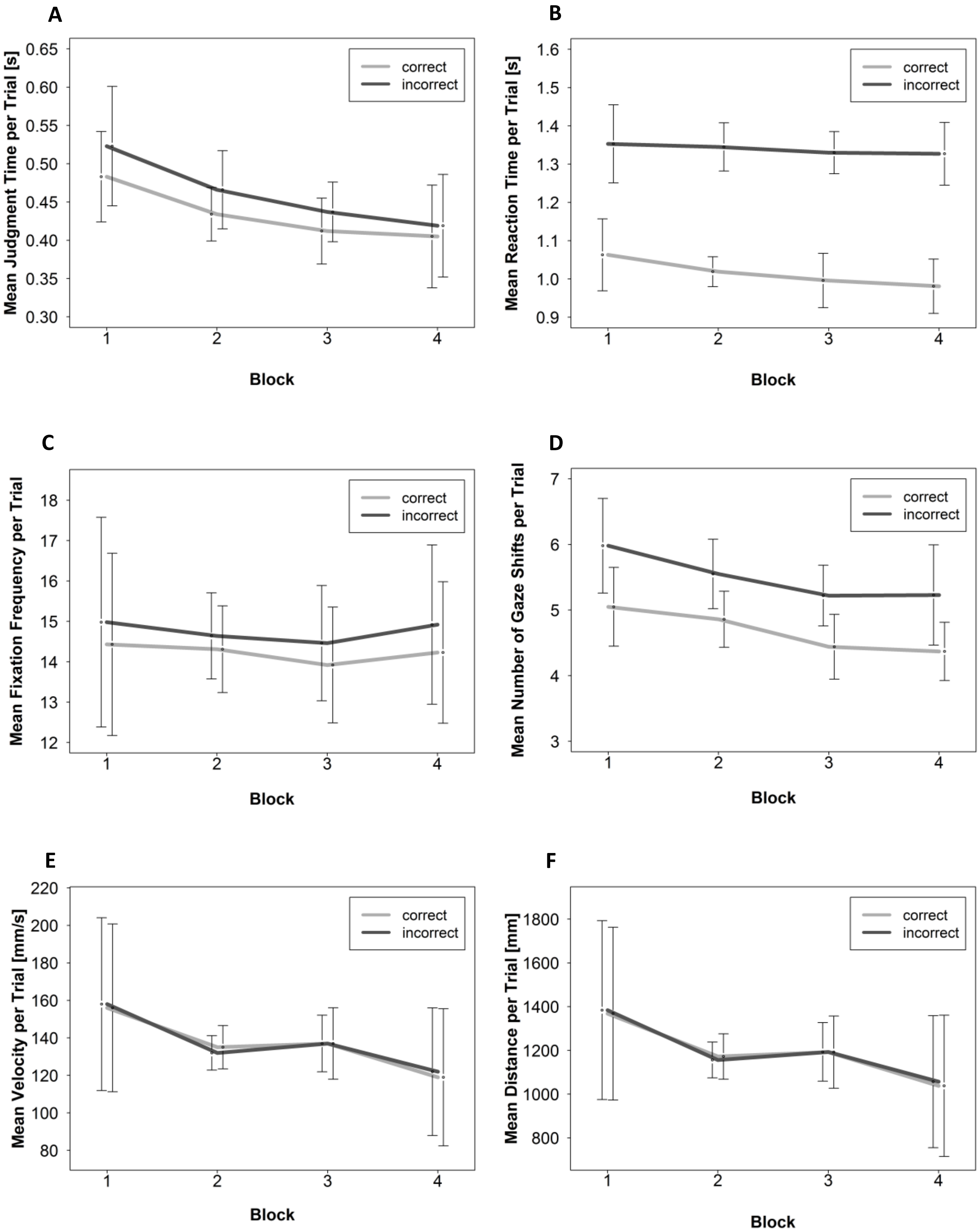


Figure 7.3: Results of variables of the high-probability condition in Experiment V: Judgment time (A), reaction time (B), number of gaze shifts (C), fixation frequency (D), scanpath distance (E) and saccadic velocity (F) across blocks for correctly and incorrectly predicted trials as a function of *block* and *judgment*.

AOIs:

The analysis of the AOIs generally showed that values during correctly predicted trials were higher in AOI_{target}, $AOI_{predict}$ and AOI_{high} and lower in AOI_{low} in comparison to incorrectly predicted trials. For reasons of clarity the results of the AOIs are shown in table 7.4.

Table 7.4: Results of eye movements in the AOIs in the high probability condition for correctly in comparison to incorrectly predicted trials of Experiment V

AOI	Variable	Higher ⬆ or lower ⬇ values for correct predictions	ANOVA
AOI_{target}	dwell time	⬆	$F(1,23)=53.363$, $p<.001$, $\eta_p^2=0.699$
	fixation duration	⬆	$F(1,23)=48.715$, $p<.001$, $\eta_p^2=0.679$
	fixation frequency	⬆	$F(1,23)=27.592$, $p<.001$, $\eta_p^2=0.545$
$AOI_{predict}$	dwell time	⬆	$F(1,23)=261.979$, $p<.001$, $\eta_p^2=0.919$
	fixation duration	⬆	$F(1,23)=163.107$, $p<.001$, $\eta_p^2=0.876$
	fixation frequency	⬆	$F(1,23)=120.944$, $p<.001$, $\eta_p^2=0.840$
AOI_{low}	dwell time	⬇	$F(1,23)=177.927$, $p=.020$, $\eta_p^2=0.886$
	fixation duration	⬇	$F(1,23)=123.446$, $p<.001$, $\eta_p^2=0.843$
	fixation frequency	⬇	$F(1,23)=120.944$, $p<.001$, $\eta_p^2=0.840$
AOI_{high}	dwell time	⬆	$F(1,23)=188.507$, $p=.020$, $\eta_p^2=0.891$
	fixation duration	⬆	$F(1,23)=140.644$, $p<.001$, $\eta_p^2=0.859$
	fixation frequency	⬆	$F(1,23)=75.407$, $p<.001$, $\eta_p^2=0.766$

7.4.2 Low Probability Concept

Task performance:

Analysis of *block* revealed a significant main effect, $F(5,115)=6.23$, $p=.001$, $\eta_p^2=0.213$, suggesting a significant increase of task performance across blocks (Fig. 7.6).

Judgment time:

We observed neither a main effect of *block*, $F(5,115)=0.75$, $p=.478$, $\eta_p^2=0.032$, nor a main effect of *judgment*, $F(1,23)=2.86$, $p=.104$, $\eta_p^2=0.111$, indicating no significant change (Fig. 7.5A). The detailed analysis of judgment times in the low probability condition showed no significant effects (p>.10). However, mean values show the same pattern as in the high probability condition (Fig. 7.4.). Likely exit predictions (correct: *M*=0.457; *SD*=0.233, incorrect: *M*=0.452; *SD*=0.220) were descriptively faster than unlikely exit predictions (correct: *M*=0.534; *SD*=0.237, incorrect: *M*=0.547; *SD*=0.245).

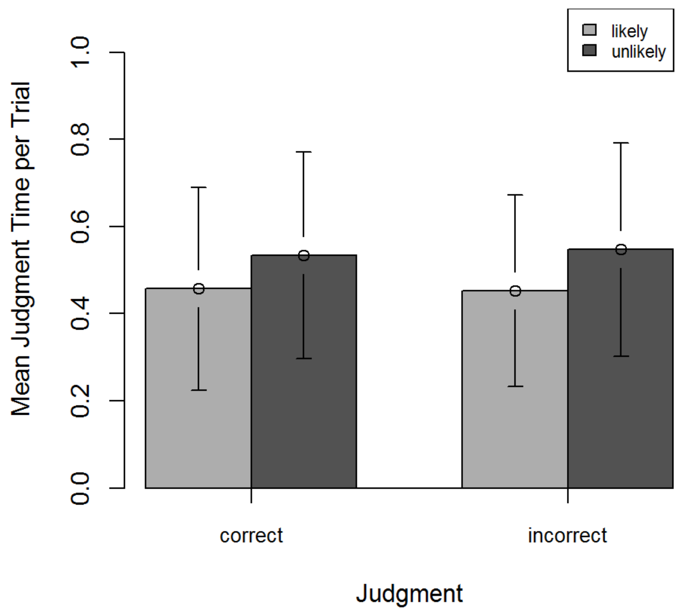


Figure 7.4: Judgment times in Experiment V in the low probability condition: Correct and incorrected judgments for likely or unlikely exits. Error bars depict the standard deviation.

Reaction time:

There was a main effect of *block*, $F(5,115)=5.39$, $p=.004$, $\eta_p^2=0.190$, suggesting a decrease of reaction times across blocks. In addition, there was a main effect of *judgment*, $F(1,23)=95.73$, $p<.001$, $\eta_p^2=0.806$, suggesting shorter reaction times during correctly than incorrectly predicted trials (Fig. 7.5B).

Fixation frequency:

We did not observe a main effect of *block*, $F(5,115)=0.81$, $p=.467$, $\eta_p^2=0.034$. However, analysis of the AOIs revealed a trend for a main effect *block* in $AOI_{predict}$, $F(5,115)=2.51$, $p=.063$, $\eta_p^2=0.098$, indicating a decrease of fixation frequency in $AOI_{predict}$ across blocks (see Appendix C, Table 9.7 for details). Furthermore, we observed a main effect of *judgment*, $F(1,23)=6.58$, $p=.017$, $\eta_p^2=0.222$, indicating fewer fixations during correctly than incorrectly predicted trials (Fig. 7.5C)

Fixation duration:

There was neither a significant main effect of *block*, $F(5,115)=0.72$, $p=.484$, $\eta_p^2=0.030$, nor a significant main effect of *judgment*, $F(1,23)=1.10$, $p=.306$, $\eta_p^2=0.046$, indicating no significant change.

<u>*Number of gaze shifts:*</u>

We found trend for a main effect *block*, *F*(5,115)=6.58, *p*=0.087, η_p^2=0.095, indicating that the number of gaze shifts decreased across blocks. We found also a main effect of *judgment*, *F*(1,23)=48.68, *p*<.001, η_p^2=0.679, indicating that participants showed less gaze shifts for correctly predicted than for incorrectly predicted trials (Fig. 7.5D).

<u>*Gaze velocity:*</u>

We observed neither a significant main effect of *block*, *F*(5,115)=2.04, *p*=.154, η_p^2=0.081, nor a significant main effect of *judgment, F*(1,23)=0.21, *p*=.655, η_p^2=0.009, indicating no significant change.

<u>*Scanpath distance:*</u>

Analysis revealed neither a significant main effect of *block*, *F*(5,115)=2.03, *p*=.154, η_p^2=0.081, nor a significant main effect of *judgment, F*(1,23)=0.04, *p*=.851, η_p^2=0.002, indicating no significant change.

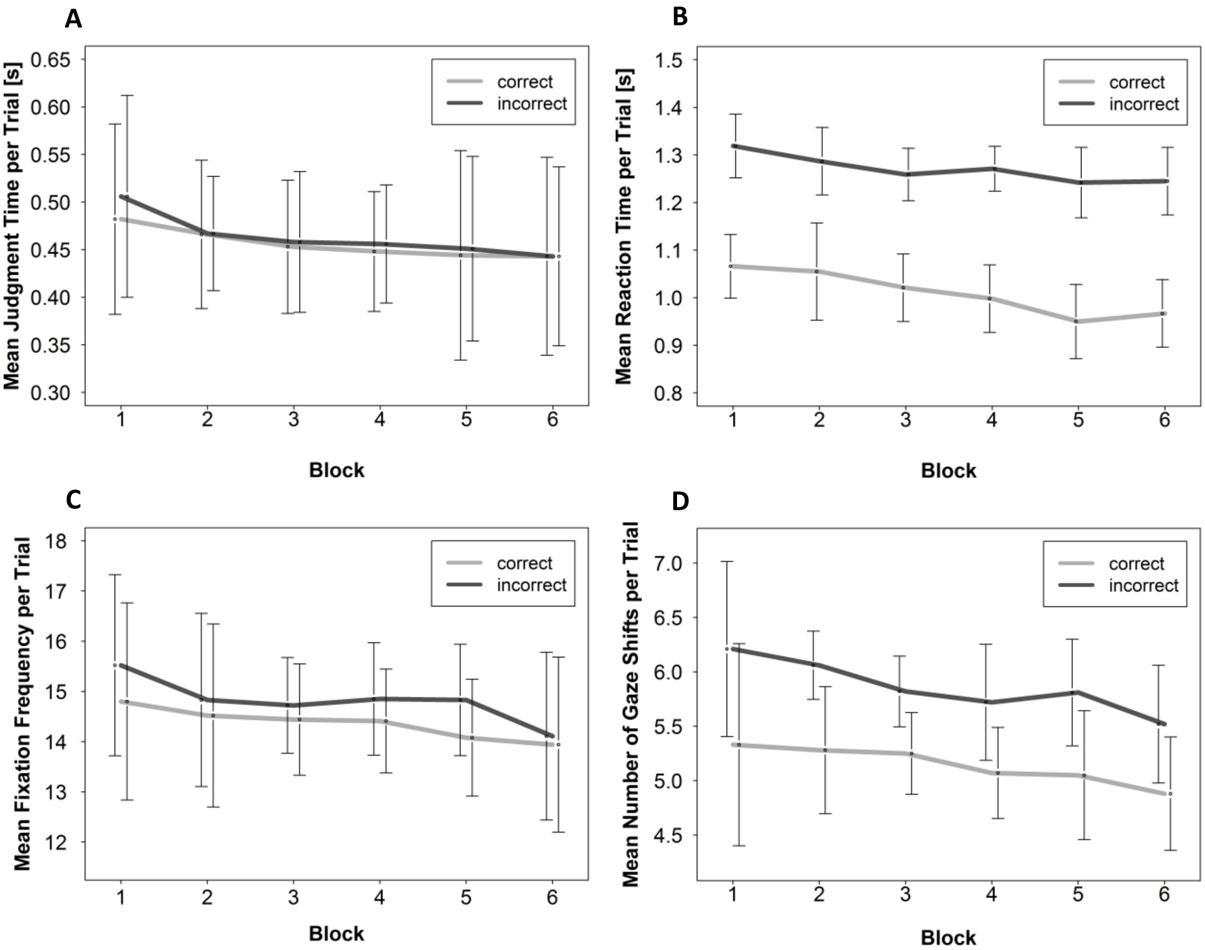


Figure 7.5: Results of variables of the low probability condition in Experiment V: Judgment time (A), reaction time (B), number of gaze shifts (C) and fixation frequency (D) across blocks for correctly and incorrectly predicted trials as a function of *block* and *judgment*.

AOIs:

Table 7.5. shows the results for the analysis of the AOIs which are similar to those in the high probability condition. Values of the eye movement variables during correctly predicted trials were higher in AOI_{target}, $AOI_{predict}$ and AOI_{high} and lower in AOI_{low} in comparison to incorrectly predicted trials.

Table 7.5: Results of eye movements in the AOIs in the low probability condition for correctly in comparison to incorrectly predicted trials of Experiment V

AOI	Variable	Higher ⬆ or lower ⬇ values for correct predictions	ANOVA
AOI_{target}	dwell time	⬆	$F(1,23)=70.548$, $p<.001$, $\eta_p^2=0.754$
	fixation duration	⬆	$F(1,23)=58.799$, $p<.001$, $\eta_p^2=0.719$
	fixation frequency	⬆	$F(1,23)=49.727$, $p<.001$, $\eta_p^2=0.684$
$AOI_{predict}$	dwell time	⬆	$F(1,23)=448.708$, $p<.001$, $\eta_p^2=0.951$
	fixation duration	⬆	$F(1,23)=245.013$, $p<.001$, $\eta_p^2=0.914$
	fixation frequency	⬆	$F(1,23)=129.959$ $p<.001$, $\eta_p^2=0.850$
AOI_{low}	dwell time	⬇	$F(1,23)=80.099$, $p<.001$, $\eta_p^2=0.777$
	fixation duration	⬇	$F(1,23)=70.706$, $p<.001$, $\eta_p^2=0.755$
	fixation frequency	⬇	$F(1,23)=75.803$, $p<.001$, $\eta_p^2=0.767$
AOI_{high}	dwell time	⬆	$F(1,23)=76.489$, $p<.001$, $\eta_p^2=0.769$
	fixation duration	⬆	$F(1,23)=60.082$, $p<.001$, $\eta_p^2=0.723$
	fixation frequency	⬆	$F(1,23)=62.292$, $p<.001$, $\eta_p^2=0.730$

7.4.3 Comparison of the Low and High Probability Concept

As shown in Figure 7.6, the number of correct predictions as well as the prediction of the likely exits increased across blocks in both conditions. However, the learning effect was significantly higher for the high probability concept than for the low probability concept, $t(46)=3.83$, $p=.001$. In addition, the comparison of number of correct prediction for the first 4 blocks showed a significant between-subject factor with regard to the probability condition, $F(1,46)=17.66$, $p<.001$, $\eta_p^2=0.277$, and a trend for an *interaction* effect, $F(3,138)=2.31$, $p=0.093$, $\eta_p^2=0.048$, . Figure 7.6 indicates that the number of correct predictions increases in the high probability condition continuously whereas it seems that the learning curve ends in a plateau between block 2 and block 4 in the low probability condition. A learning curve was fitted to check if participants generally learn in the low probability condition. Data of the power model showed a very good fit (R^2=.99) for the high probability condition, but a poor fit (R^2=.70) for the low probability condition suggesting that participants clearly learn in the high probability condition, but not in the low probability condition. Contrary to the prior assumption, the comparison of eye movements and judgment times between both conditions showed no differences (p>.38).

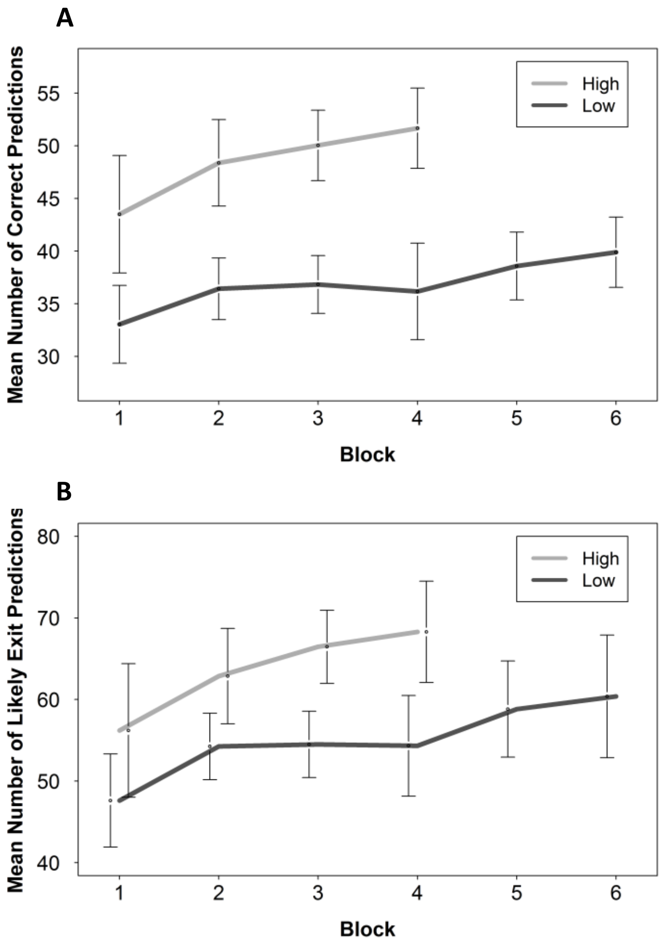
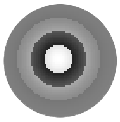
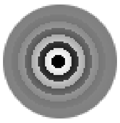


Figure 7.6: Comparison of predictions in the low and high-probability condition in Experiment V: Mean number of correct predictions (A) and mean number of likely exit predictions (B) for the high probability concept (High) and the low probability concept (Low) across blocks.

Results of the Concept Awareness Questionnaire showed that tendencies for both concepts were learned in a correct manner. Accordingly, likely object-exit associations in the low probability condition were estimated lower than in the high probability condition (Tab. 7.6). Behavioral data underline these findings as participants chose the likely exit in 78.8% of the cases while learning the high probability concept whereas the likely exit was chosen in 68.0% of the cases during the performance of the low probability concept. However, three participants stated no understanding of any probability concept even if it was mentioned in the instruction that a probability distribution had to be learned. Further, six participants filled in not to understand one of the concepts, either the lower or higher probability concept.

Table 7.6: Subjective probability concept in the high and low probability condition in Experiment V

Gabor figures	Probability	High Probability Concept	Low probability Concept
	high	62 % (21.8)	53 % (19.1)
	low	19 % (11.0)	23 % (9.4)
	high	60 % (19.4)	55 % (17.5)
	low	20 % (9.6)	23 % (8.7)
	high	66 % (18.1)	53 % (16.4)
	low	17 % (9.0)	23 % (8.1)

Note. Values in brackets show the standard deviation.

7.4.4 Comparison of Different Stimulus Presentations

Stimuli in Experiment I-V were gradually developed due to hints of an object bias. In the following, the influence of the stimulus presentation on the subjective probability concept, eye movements and performance data is reported. The probability concept of 74-11-11 was stable in all of the experiments enabling a comparison of the results. The subjective probability concept for geometric figures used in Experiment I (see Chapter 3) was more accurate than the subjective probability concept for Gabor figures in the current experiment (few circles, many circles and symmetric pattern). The prediction error, viz. the deviation from the target value, differed significantly between geometric figures in Experiment I (*M*=8.37; *SD*=8.66) and Gabor figures in Experiment V (*M*=15.43; *SD*=15.46), $t(39)=1.86$, $p=.035$. It seems that behavioral data are closely related to the self-estimation of the probabilities as in Experiment I with geometric figures likely exits were chosen in 87.7% of the cases and probability estimations of the participants were more accurate (on average 69%) for the likely object-exit association. In the current experiment with Gabor figures the average estimation of the likely exit was only 63% in the condition with the high probability concept which employed the same probability distribution used in Experiment I. In addition, participants showed more fixations during correctly predicted trials in Experiment V (*M*=14.22; *SD*=3.84) than in the original version with geometric figures (*M*=11.15; *SD*=4.22), $t(39)=2.47$, $p=.018$, presumably due to the degradation of the stimuli by designing the stimuli with more detailed patterns and the fade in of the Gabor figures.

In comparison to Experiment II an unstructured gray-white patterned was used for the stimulus degradation of the geometric figures. Fixation frequency during correctly predicted trials in Experiment II (M=14.51; SD=3.48) did not differ significantly from the current experiment, $t(39)=0.24$, $p=.811$. However, the prediction error was significantly smaller in Experiment II (M=8.43; SD=5.43) than in Experiment V (M=15.43; SD=15.46), $t(39)=1.86$, p=.035. Thus, there seems to be a difference in degrading the stimulus by including a pattern in the background or by making the stimulus more detailed with regard to the learning process. However, the results have to be interpreted with caution as these analyses base on a between-subject design and participants were explicitly instructed to learn a probability concept in the Experiment V. Thus, also other experimental conditions were manipulated additionally.

7.4.5 Analysis of Control Variables

Correlational analysis showed different effects of the tested control variables. Interest measured by the QCM (M=3.28, SD=0.92) showed a significant effect on task performance only during the presentation of the high probability concept (r=.547, p=.003). The effect is missing in the low probability condition (r=.224, p=.146).

Tolerance of ambiguity measured by the IMA (full scale: M=143, SD=16.72) indicated an overall low tolerance of ambiguity, as values for all scales are below a percentile rank of 50 compared to the table of standard values for the age group of 20-29 years (Reis, 1996). Overall, participants' tolerance of ambiguity did not affect task performance (p>.13).

7.5 Discussion

In the last experiment of this experimental series, participants had to learn a higher and a lower probability distribution in order to investigate if previous results can also be applied to a higher degree of uncertainty. Performance data and the self-reported probability concept showed that participants were able to learn both concepts, viz. they correctly estimated the tendencies of the probability concepts but made less correct predictions during the presentation of the low probability concept as expected. However, eye movement data as well as judgment times could not reflect differences in the learning processes, presumably due to same degree of subjective uncertainty in both conditions and the usage of the same decision strategy, namely probability maximizing instead of probability matching. Thus, the manipulation of the subjective uncertainty by manipulating the probabilities did not seem to work.

Mental model development:

Participants made significantly less correct predictions in the low probability condition than in the high probability condition in line with the given lower probabilities. In the low probability condition participants' performance increased at the beginning and was rather stable from Block 2 to Block 4 indicated by a trend of an interaction and the poor fit of the learning curve. After Block 4, the mean values of participants' task performance suggest again an increase until the last block. Thus, it was useful to run six blocks of the low probability condition. The plateau in the middle of the experiment might occur due to the high objective uncertainty as more information has to be stored before an accurate mental representation can be developed than in the high probability condition. A similar learning curve for trainings in the context of working environments was described by Reichel (1985, p. 60). Generally, the performance of the training increased at the beginning of the training sessions until a performance plateau occurs. A further increase of the performance could only be reached by an extension of the training workload.

Effects on cognitive processing:

Unexpectedly, results showed no differences between judgment times in the low probability and the high probability condition. However, there were differences with regard to the development over time. In the low probability condition judgment times did not change significantly across blocks in contrast to the high probability condition. This might base on the higher objective uncertainty which is closer to chance and the resulting less steep learning curve in comparison to the high probability condition as mentioned before. The detailed analysis of judgment times, divided into correct and incorrect and likely and unlikely exit predictions, showed only significant effects in the high probability condition but not in the low probability condition. The pattern of results seems to be comparable with previous experiments. Thus, unlikely exit predictions might need more cognitive resources presumably due to a change in strategy as reported earlier. Nevertheless, the inconsistency of significance did not allow a concrete conclusion, but rather a presumption.

In accordance with previous experiments, reaction times clearly showed a preparation benefit in correctly predicted trials leading to faster reactions. However, the temporal development of the reaction time was only significant in the low probability condition, presumably due to the longer learning process. In the high probability condition, the learning effect was usually highest in the first block as shown in all previous experiments suggesting a ceiling effect that is well known in statistics. This might be the reason why reaction times did not decrease to a greater extent in the following blocks.

Visual search behavior:

It was also expected that participants show more visual search behavior in the low probability condition. Results showed that scanpath length and saccadic velocity decreased significantly across blocks in the high-probability condition, but not in the low-probability condition. A longer scanpath often indicates less efficient searching of the environment according to Goldberg et al. (2002). In the high-probability condition participants might be able to improve information search across blocks entailing a reduced scanpath length whereas the low-probability condition might be too demanding for this development due to the high uncertainty.

However, results of a direct comparison showed no significant differences in the visual search behavior between conditions. This might be due to the same decision strategy participants used in both of the conditions. Parallel to the increasing number of correct predictions, also the number of likely exit predictions increased according to TTB indicating that there was a strong influence of the decision strategy on task performance. Participants tended to use the optimal probability maximizing strategy, which is equal to TTB, in both conditions. Based on dual cognitive process theories, Schulze et al. (2015) suggested that probability matching was related to an intuitive cognitive system whereas probability maximizing relied on cognitive capacities to correct the first impulse. This might be one possible explanation for the process of choosing probability maximizing in line with the correct understanding of the probability concepts due to explicit learning. Here, participants rather seemed to interpret objective uncertainty as certainty indicated by their rather unilateral response behavior using the TTB strategy to minimize task uncertainty in both conditions. However, the low probability condition seemed to be more demanding and thus, participants presumably needed more time to identify a likely exit for every object and to build up an appropriate decision strategy reflected by the less steep learning curve. According to the degree of objective uncertainty, participants chose the likely exit in the low probability condition overall less often than in the high probability condition. Instead, task uncertainty might be reduced similarly over time in both conditions due to the initial information about learning a probability concept. Further, the experimental task did not differ in both of the conditions requiring no changed attention processes according to the attentional set which is a bias towards stimuli that are learned to be relevant (Scherrmann et al., 2010).

Besides the main expectations, results generally showed more eye movements in the relevant AOIs during correct predictions than incorrect predictions in both conditions. In more detail, participants showed longer dwell time in the relevant AOIs according to Jacob and Karn (2003) and Poole et al. (2005) as well as longer fixation frequency and fixation duration in AOI_{target}, $AOI_{predict}$ and AOI_{high} during correct predictions indicating a general understanding of the objective probability in both conditions reinforcing the importance of focusing on relevant aspects. In line with these results participants estimated the tendencies of all object-exit associations in a correct manner. Nevertheless, the comparison of the prediction error showed that learning the geometric objects in Experiment I (Chapter 3) resulted in a more accurate estimation of the high probability condition than using Gabor figures in the current experiment even if the participants were explicitly instructed to learn probabilities. Prior knowledge usually facilitates learning (Hewson & Hewson, 1983). There might be different explanations for the reported differences in task performance. It might be that the new target objects did not facilitate learning in the same way as before, because of their complexity. New target objects in Experiment V presumably took more time to be encoded, also indicated by a higher fixation frequency, leading to slower learning. However, it might also be possible that the initial instruction caused misleading and exaggerated expectations regarding the underlying probability concept (Hilbert, 2012), eventually influencing the learning process.

Stimulus presentation:

There seemed to be also a difference of the subjective probability concept if the background was fuzzy and degraded the stimuli (Experiment II) or if the stimulus itself was degraded by having no clear contour lines and a more detailed pattern (e.g., Experiment V) indicated by the prediction error. However, the degradation led to an overall high fixation frequency which did not change across blocks in both cases. Presumably, participants always needed more time to perceive and accurately encode degraded stimuli independent of the learning process. The results and corresponding interpretations had to be considered carefully as many parameters (instruction, target objects, counterbalancing, separated sessions) were changed in the current experiment, making it difficult to compare the results with the previous experiments in an accurate manner (see General Discussion for more details).

Influence of confounding variables:

The analysis of control variables showed that objective uncertainty did not only seem to influence task performance directly, but also influenced task performance via interest as provided by the questionnaire results. Results showed that high interest in the task could only be related to an improvement of task performance in the high probability condition, but not in the low probability condition. This might be due to the high objective uncertainty in the low probability condition which was too close to chance to consciously control the performance. In contrast to the variable “interest”, the newly added control variable tolerance of ambiguity did not seem to influence task performance. One reason for this might be the low variability of the data indicating an overall low tolerance of ambiguity of the participants. Another reason might be the laboratory situation which might be in contrast to the questionnaire asking for everyday life situations.

7.5 Conclusion

In the current experiment, the probability structure of the OVSST was manipulated to induce two different degrees of objective uncertainty. Participants had to learn each in a separate test session. Participants made more correct predictions within the high probability condition presumably due to the lower objective uncertainty, however, this difference was neither reflected by eye movement parameters nor by judgment times probably due to the same decision strategy they used in both conditions, i.e. probability maximizing. This strategy led to the assumption that participants rather interpreted objective uncertainty as certainty to reduce their task uncertainty. In sum, it has to be critically reviewed what eye movements actually reflect.

8 General Discussion

In a series of five experiments, we investigated visual search behavior during the acquisition of mental representations under uncertainty. In the following, the research questions mentioned in Chapter 1 are answered to summarize the results. Then, different questions are derived from the reported results to discuss the issues on a more general level. Thereafter, practical relevance and limitations of the experimental studies are discussed and possible future research introduced. Finally, the chapter ends with a conclusion of the thesis.

8.1 Summary

Research question 1 - The meaning of eye movements in the development of mental representations: Eye movements gave insights into the state of learning concerning the mental representation of the task and the degree of subjective uncertainty.

Although the objective probability remained constant across trials, participants showed distinct behavioral variability in their response to uncertainty. In the beginning of the task, participants seemed to be uncertain due to the missing prior knowledge about the underlying concept of the task and shown by the extensive visual search behavior. Visual search behavior became more focused over time in parallel to an increasing learning rate. Behavioral data suggested that participants preferred probability maximizing strategies and thereby considered only the likely exits, which equaled the take the best strategy (TTB). Thus, objective uncertainty seemed to be rather interpreted as certainty by the participants during all of the described experiments in the way that participants interpreted the 74% probability as 100% probability and ignored lower probabilities (Chapter 3 -7).

Research question 2 - The effect of degraded stimuli on the development of mental representations: Degraded stimuli did not hamper learning per se, but might even facilitate learning depending on the kind of distraction.

In Chapter 4 (Experiment II), the OVSST was altered by introducing an unstructured background to force participants to use more attentional resources to encode the target objects. Indeed, fixation frequency was higher in Experiment II with degraded stimuli than in Experiment I. However, participants reported in part a more accurate mental representation of the probability concept, presumably due to the more focused attention or perhaps even more perceived stimulus quality. In the latter case, the pattern in the background of the screen might be used as landmarks facilitating learning and memory retrieval. However, using degraded target objects in the form of detailed Gabor figures led to less accurate subjective probability concepts even if fixation frequency was similarly high.

In this case, degraded target objects rather hampered learning and the high fixation frequency might indicate difficulties during encoding of the target objects (Chapter 7) which complied with the experiments of Sternberg (1969).

<u>Research question 3 - Eye movement patterns as an indicator of relearning a probability concept:</u> Eye movement patterns indicated different phases of relearning but entailed specific characteristics.

Participants had to relearn a probability concept, i.e. they had to learn a probability concept and then, they had to switch to another probability concept without prior knowledge. This unknown change of the concept led to decreased performance accuracy and was reflected by a higher fixation frequency. Thus, the beginning of the relearning phase was directly signalized by a higher fixation frequency. In contrast, the characteristics of fixation duration responded on the relearning with a time delay (Chapter 5). However, generalizability of the results was questionable.

<u>Research question 4 - The differential influence of the prediction and the reaction task on the mental representation:</u> The developed mental model during the performance of OVSST is based on the prediction task and not on the reaction task.

Initially, the prediction task of the OVSST was developed to enable the acquisition of a mental representation about the underlying probability concept. The reaction task should force participants to show visual search behavior at the exits. Experimental results actually showed that participants were only able to report an accurate mental representation of the prediction task but not in the reaction task when testing both tasks separately (Chapter 6). Thus, deeper cognitive processing involved in the decision making of the prediction task seemed to affect memory retrieval. However, eye movement parameters seemed not to reflect different cognitive processes during the performance of the prediction task and the reaction task.

<u>Research question 5 - Eye movements as an indicator of different degrees of objective uncertainty:</u> Eye movement patterns did not reflect different learning processes during the performance of the OVSST under high and low uncertainty.

Objective uncertainty was varied by choosing a low and high probability distribution of the OVSST. Participants had to learn both probability concepts in separate sessions and were explicitly informed to learn a probability concept before the experiment. Results showed a better task performance in the high probability condition than in the low probability condition according to the objective uncertainty. However, eye movement patterns did not differ between the performances of both concepts (Chapter 7). One reason for this might be that participants used the same decision strategy in both conditions: probability maximizing. Furthermore, task uncertainty might be similar in both conditions due to the initial information about learning a probability concept. Another reason might be that participants were not able to perceive differences of the high and low probabilities in the low probability condition as probabilities were too close to chance. However, the quite accurate mental representation of the subjective probability concept in the low probability condition argued against this reasoning.

Are learning processes indicated by behavioral characteristics?

This question seemed to be highly relevant to ensure that learning processes occurred during the performance of the OVSST. As already mentioned in the main introduction, in the third stage of the perceptual recognition process by Jacob and Hochstein (2009), reaction times decreased due to explicit knowledge which led directly to an attention allocation on the target stimulus. In almost all experiments judgment times as well as reaction times decreased across blocks indicating the ongoing learning process and thus, the development of a mental representation also reported in literature (e.g. Hunt & Aslin, 2001).

Another evidence for the participants' successful learning process was the choice of the probability maximizing strategy. Likely exit predictions were preferred in order to improve task performance. This reflected the "Take The Best" decision strategy which seemed to be the most adequate decision strategy participants could use to increase task performance in OVSST (cf. Yu & Huang, 2014).

Further characteristics of the learning process, found in the current experiments, were preparation benefits, i.e. during correctly predicted trials participants were able to anticipate the correct target exit of the object reappearance. This preparation benefit resulted in faster reaction times and fewer gaze shifts during correctly than during incorrectly predicted trials. If participants failed to anticipate the target exit, more visual search was necessary to focus on the relevant stimulus and to react in an appropriate way.

Interestingly, judgment times were also faster during correctly predicted trials even if the feedback was given after the prediction. For a deeper understanding, judgment times were analyzed separately for correct and incorrect predictions and likely and unlikely exits. Overall, results rather showed slower judgment times for unlikely exit predictions. An explanation for this might be the costs that occurred during the change of strategy also called strategy switch costs (Lemaire & Lecacheur, 2010). Actually, likely exits were preferred as mentioned earlier in the section. If another strategy was selected, the preferred strategy had to be inhibited first. Thus, a new selection seemed to be more costly and is finally reflected by longer judgment times.

What do eye movements reflect during the mental model acquisition of choices?

As described in the previous paragraph, evidence for learning processes and thus, the development of a mental representation of the OVSST was found. Thereby, the basic assumption of the experimental paradigm could be confirmed. However, the crucial issue of this thesis was the role of eye movements during the development of the mental representation discussed in the following.

Earlier theories like the eye-mind assumption by Just and Carpenter (1980) assumed that eye movements were a mirror of the brain and reflect cognitive processes. However, the limits of this theory were already described by Irwin (2004) as mentioned in the main introduction. The validity of the assumption is further questionable for the following reasons. Kok and Jarodzka (2016) concluded recently that "Eye movements reflect cognitive processes, but cognitive processes cannot be directly inferred from eye tracking data." (p.1). Irwin (2004) extended this general conclusion by noting that fixations are "[…] not sufficient to specify precisely what information a subject is processing from a visual display and how effectively it is represented and interpreted by the cognitive system." (p.110). This criticism was also brought forward by other authors stating that there was still little knowledge of how cognitive process could be deduced from eye movements (Feng, 2003; Wedel & Pieters, 2008; Wedel & Pieters, 2008). Further, eye movements should always be interpreted in the scope of the current task to avoid misleading conclusions (Gidlöf, Wallin, Dewhurst, & Holmqvist, 2013; Kok & Jarodzka, 2016; van der Gijp et al., 2016; Yarbus, 1967). We could thus assume that eye movement patterns in the current study might reflect cognitive processes but additional information were necessary to gain access to the underlying content.

The results of the experiments presented in this thesis underlined and specified the statements by saying that rather general cognitive processes were reflected by eye movements, i.e. learning effects. More precisely, the degree of subjective uncertainty and the overall state of learning was indicated by decreasing visual search behavior. Thus, eye movements seemed to be strongly coupled with processes that were relevant for learning, for example attentional processes, information accumulation and the familiarity of the environment.

However, visual search behavior did not reflect cognitive processing depth of uncertain concepts. Here, eye movements seemed not to inform about the depth of cognitive processing during the separate testing of the prediction and the reaction task. Although, longer fixation duration was often related to deeper processing in literature (Holmqvist et al., 2011), this was not observed in the current experiments.

Anderson et al. (2004) proposed to redefine the eye-mind assumption in the way that this assumption only applied for the encoding of information. Thus, eye movements reflected only ongoing processes that depended on the encoding of information. The relevance of eye movements for ongoing processes was confirmed by results of the well-studied “looking at nothing” phenomenon. This phenomenon clearly showed that the gaze was directed to empty spatial locations that were associated with the relevant information earlier. Thus, the retrieval process was facilitated and eye movements informed about ongoing processes (Scholz et al., 2015). However, this effect diminished with practice (Scholz, Melhorn, Bocklisch, & Krems, 2011). The results of the current study showed the same tendency: If the mental representation was enriched, participants no longer showed extensive visual search behavior and retrieval processes were no longer reflected by eye movements.

However, it was important to consider eye movements not as one homogenous entity but to differentiate between specific characteristics as listed, for instance, by Ehmke and Wilson (2007) as well as by Poole and Ball (2006) and shown in the relearning experiment (Experiment III). Here, fixation frequency directly indicated the beginning of the relearning phase whereas fixation duration did not. Next to breaking down eye movements into specific parameters, it was also relevant to consider eye movement parameters in the light of interindividual differences. The results of the experiments already indicated that the interindividual variability was high which has also been observed by Goldberg and Wichansky (2003). Every person showed a distinct average fixation duration (Johansson, Holmqvist, Mossberg, & Lindgren, 2011; Rayner, Li, Williams, Cave, & Well, 2007) and either people made lots of fixations or only a small number of fixations described as a kind of personality in visual search (Holmqvist et al., 2011). In addition, memory capacities as well as attentional and cognitive abilities varied individually (Jipp, 2016; Yi & Davis, 2003) also influencing visual search behavior.

Nevertheless, in this thesis we mainly tried to examine general common features of visual search behavior. In conclusion, eye movement patterns seemed to reflect the state of the user viz. the degree of subjective uncertainty and the overall state of learning of a human decision maker. Thus, individual troublesome situations can be derived from eye movement patterns.

Are the three-stage models of perceptual and learning processes applicable to uncertain situations?

To be able to explain the relations of eye movement parameters and behavioral measures on a more general level, three stages of eye movements during learning under uncertainty were introduced. The three-stage model of perceptual processes (Jacob & Hochstein, 2009) and learning processes (Ackerman, 1988; Schumacher & Czerwinski, 2014) compared in the main introduction seemed to be plausible for the current findings. The three stages of the learning models seemed also to be evident in the current study: (1) Participants processed the information given by the OVSST, (2) they developed a mental representation of the task via information accumulation and finally, (3) an adequate decision strategy was developed and actions became more automatic reflected by faster judgment and reaction times. The three-stage structure was also true for eye movements: First, (1) participants showed an extensive search behavior which (2) became more precise over time so that finally, (3) mainly relevant stimuli were focused. As an extension, there might also be three stages of subjective uncertainty: (1) The degree of subjective uncertainty was higher at the beginning of the experiment due to the innocence of the underlying probability concept, (2) diminished as a result of information accumulation until (3) a lower degree of subjective uncertainty was reached due to a comprehension of the situation.

Based on this comparison of the three-stages and the results of the current experiments, I proposed a general three-stage model for eye movements during learning which also applies to learning processes in uncertain situations as an extension (Fig. 8.1). Thus, a higher degree of subjective uncertainty at the beginning of the experiment was indicated by extensive visual search and the accumulation of information. In the next stage, subjective uncertainty as well as visual search behavior was reduced due to the advanced learning and mental model acquisition leading to an improved performance. Finally, uncertainty was low indicated by the visual focus on relevant information and automated actions that were precisely stored in the mental model. A difference between learning processes in certain and uncertain environments might be the speed of learning. Uncertain information probably led to slower learning than certain information.

A limitation of the model was that the objective uncertainty was not depicted, but rather the subjective uncertainty. As shown in Experiment V eye movements could only reflect the subjective uncertainty but not the objective uncertainty of the two probability concepts with lower and higher uncertainty. A further limitation referred to the segregation of the second stage in the three-stage models. The degree of subjective uncertainty in the second stage seemed to be vague and could not be separated clearly in the current experiments, especially due to the quick learning of the OVSST mainly within the first block. The learning curve in the low probability condition of Experiment V indicated three-stages due to the plateau in the middle of the learning process but clear evidence for a second stage was not given. Nevertheless, a combination of the three-stage models seemed to be adequate.

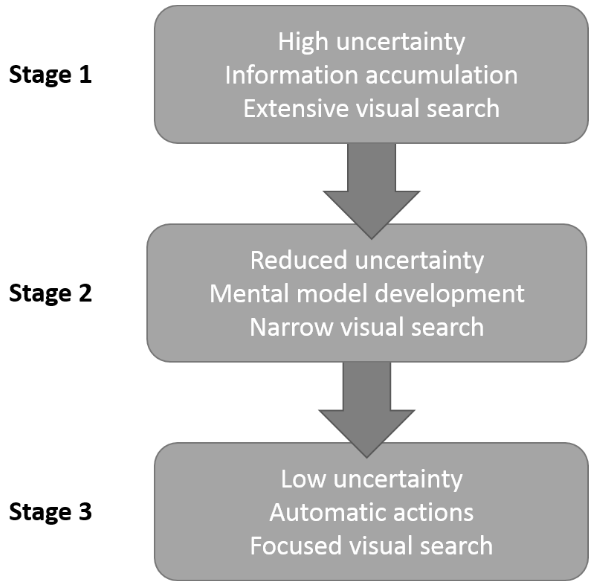


Figure 8.1: An extension of three-stage models: Three stages of eye movements during learning under uncertainty. The degree of uncertainty, the state of learning and the visual search behavior develop in parallel.

How does the visual presentation affect the mental model acquisition?

The proposed three-stage model described a learning process, however, it did not consider the accuracy of the learned content. The accuracy of the developed mental representation seemed to be strongly influenced by the visual quality of the stimuli and is sensitive to changes as shown by the results of the experiments and the continuous adaptation to the stimuli due to biases. An unstructured, but stable patterned background seemed to facilitate learning of single geometric target stimuli whereas more detailed patterned target stimuli seemed to hamper learning. Furthermore, even abstract stimuli used in the current study seemed to comprise information that bias results and thus, could not be considered independently. Boucheix and Lowe (2010) already showed in their study that subtle differences between stimuli influenced learning. Using spreading-color cues to highlight relevant information of an animated piano mechanism led to a higher attention on the relevant stimuli and improved the comprehension as well as the quality of the mental model compared to arrow cues.

Further, feedback affected the learning process and thus, the mental model acquisition. The OVSST provided feedback about the task directly without a long time delay and thereby directed attention to the relevant stimuli. This enabled to build a connection between the stimuli and the exit probability and putting forward an expectation used for the next prediction. Further, feedback about the task performance was depicted after every block. In literature the impact of feedback was already confirmed when misinterpretations of the feedback could be precluded and learners were motivated to learn (Hattie & Timperley, 2007; Kluger & DeNisi, 1996). The effect of motivation on the learning process during decision making was discussed in the next paragraph. In conclusion, only task relevant information should be available in a clear manner to evoke accurate mental models (Canham & Hegarty, 2010).

What is the role of interest as a motivational aspect in the decision process?

As state above motivational aspects were relevant in the learning process and were taken into account in the current study by measuring interest in the task. Results showed that motivational processes influenced the decision process, however in an unsystematic way. A higher level of motivation led to a performance improvement as well as to a performance decline depending on the task characteristics. If people tried harder, they might complicate their decision strategies, for example, by using probability matching (Fantino & Esfandiari, 2002). Finally, they made less correct predictions as shown in the first experiment. Laude et al. (2012) reported similar results comparing choices between a more likely (75%) and less likely (50%) option for food of hungry motivated pigeons and less hungry thus, less motivated pigeons. Highly motivated pigeons chose the suboptimal option more often presumably due to a greater impulsivity whereas less motivated pigeons preferred the optimal reinforcement more often. In studying human gambling behavior Molet et al. (2012) also emphasized a tendency for suboptimal strategies for participants with higher gambling motivation, viz. more self-reported gambling activities.

However, if people were highly motivated and tried to integrate the presented feedback in their mental model, task performance could also be positively affected as shown by Experiment I and IV. There also seemed to be a limit concerning the extent to which task performance could be influenced as shown in the last experiment. If the task is too difficult, even higher motivation to learn the underlying concept could not cause better task performance. In literature negative and positive effects of motivation were already reported (Bekkering & Neggers, 2002; Laude et al., 2012; Molet et al., 2012). In the current study negative, positive as well as no effects of motivation on task performance were reported even if the same paradigm was used in all experiments. Thus, motivation seemed to be quite sensitive and presumably dependent on the sample. Further studies might measure motivation with a broader scale to get a clearer statement.

8.2 Practical Relevance

This section provides a possible application of the research findings presented within this thesis and briefly discusses the practical feasibility. The reported eye movement patterns might be relevant for practitioners in the context of detecting individual learning problems and task uncertainty. They could improve human-computer interaction even under uncertain conditions by adapting the information content to the individual user's learning curve and cognitive capacities. The information of the state of the user might be collected by eye tracking measures and used to adapt the system to the individual user's state (Goldberg & Wichansky, 2003). Lai et al. (2013) already had the idea of an "[…] adaptive learning system with eye tracking embedded technology [...]" (p.99) to improve learning processes. A more optimal level of cognitive load, i.e. the mental effort to operate working memory efficiently, might be reached by reducing the amount of information and offering help systems. Cognitive load theory was based on limited capacities of working memory for information and emphasizes the importance of instructional design to reduce extraneous cognitive load which was especially of high relevance in practice with multiple stimuli (Sweller, 1988; van Merriënboer & Sweller, 2005).

Another possibility to reduce cognitive load might be the guidance of eye movements. Earlier studies already reported that attention of participants can be guided via a model's eye movements, i.e. the expert's eye movements were presented to the novice while solving a task. This intervention fostered learning by improving the visual search behavior and thus enhancing the interpretation of relevant information (Jarodzka, van Gog, Dorr, Scheiter, & Gerjets, 2013; van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009). The guidance of attention through experts' eye movements might be an adequate intervention if a troublesome situation was previously detected via visual search behavior.

In order to realize the practical applications, a few issues have to be solved. Due to the high interindividual variability, an individual baseline is needed before assumptions of the user profile can be made. As earlier discussed, the context specificity and task dependency of eye movements have to be considered additionally. There might also be failures in eye tracking that can occur due to glasses or dark eye colors as well as inconvenient lightning conditions (Goldberg & Wichansky, 2003). Therefore, eye tracking data should be used carefully and results checked with other measures if possible. Hyönä (2010) noted that offline measures are necessary to complement eye tracking data emphasizing the necessity of the Concept Awareness Questionnaire used in the current study. In addition, the acceptance of the technique due to data protection should not be disregarded.

8.3 Limitations

In the following, some methodological issues were discussed that led to limitations of this thesis. The generalizability of the reported results was reduced as participants represented only a specific target group, namely students. Also external validity of the laboratory experiments was low due to the high level of abstraction and the high task dependency of eye movements (Gidlöf et al., 2013). Further research is necessary to test the applicability on real situations.

Another limitation referred to the definition of the eye movement parameters and emphasized the need for standardized parameters. Studies in literature were often based on different parameter definitions and thus, every study was limited to these definitions and the comparability was missing. Results of the current studies might be different if other definitions of the tested eye movement parameters were used. For instance, fixations in the current study were defined by the saccade detection algorithm provided by SR Research based on velocity and acceleration criteria (Tatler, 2007). Nyström and Holmqvist (2010) already addressed the problem of a missing standard and proposed a new adaptive velocity algorithm that should be more robust. Furthermore, the display was separated into four AOIs around the exits and the entrance with a line acting as a rigorous threshold. However, the validation study mentioned earlier showed that eye tracking data was always scattered to a certain extent and not completely accurate. Thus, a corridor acting as a border area might be better to separate the AOI's instead of a line. Eye movement data within this corridor should be excluded due to the lack of a clear allocation to the AOIs.

There might also be some restrictions involved in the OVSST. The prediction task was processed offline as participants had to choose the exit when the target object is in the room and thus, when the relevant stimulus was not visible. Due to this absence of the stimuli, visual search behavior might be reduced informing less about the ongoing cognitive processes. In addition, visual search behavior might be reduced because of the simplified task that differs only in one feature, namely the direction of the exits. This might be advantageous for the interpretation of the manipulated variables, but also led to a simplification that is not common in the real world. Another limitation referred to the comparability of the stimuli that was restricted as some object-exit associations seemed to be easier to learn and to retrieve than others. The last version of target stimuli seemed to be still not optimal as Gabor figures with few and many circles also seemed to be less distinguishable in comparison to the third Gabor figure with a symmetric pattern. A solution might be the use of different letters or numbers ordered on a line to exclude a higher order. In the current experiments, the top exit had some kind of higher order due to its location. Further, it might be useful to ask participants about the ad-hoc associations of the target stimuli before running the experiment to identify any biases. Finally, the measurement of the subjective probability concept, the Concept Awareness Questionnaire, has to be improved as the 4% of the rare occurrence, the reappearance of the target object at the bottom entrance, were not included in the questionnaire. Thus, participants could not consider this event during the estimation of the probabilities, but rather tried to split 100 percent into three parts, the reappearance of the target object at the left, top and right exit. The rare occurrence and the Concept Awareness Questionnaire remained the same in all experiments to be able to compare data. However, the rare occurrence might be excluded in further experiments or the questionnaire expanded by asking for the percentage of the rare occurrence to have a more reliable set of data for the analysis.

Another aspect considered that uncertainty was operationalized and manipulated via probabilities. This manipulation did not evoke the intended effects in the way that the manipulation of the objective uncertainty did not affect subjective uncertainty, and thus might not be an appropriate way to vary the degree of subjective uncertainty. Further, objective uncertainty seemed to be rather interpreted as certainty as a kind of rationalization strategy (Lipshitz & Strauss, 1997): participants developed a decision strategy to ignore unlikely events and focused only on likely events. The question arose whether objective uncertainty can be operationalized in another way, for example, by varying the sequence of the trial (temporal uncertainty).

8.4 Directions for Future Research

Future research might follow two different directions: an extended experimental paradigm and a more real world setting for a better generalizability of the results. As mentioned in the previous section uncertainty was created by varying probabilities called objective uncertainty. However, uncertainty could also be varied by varying the time intervals in the trials and thus, creating a kind of temporal uncertainty (cf. Alegria & Bertelson, 1970; Rolke & Hofmann, 2007). The comparability of objective uncertainty and temporal uncertainty could be tested and discussed. In addition, it was still not clear which probability distributions could be learned by participants or were too close to chance as shown in Experiment V. The probability distribution finally chosen in Experiment V was tested in a pilot study, however, the probabilities were chosen by chance. The threshold for a random response behavior due to the inability to distinguish between the likelihood of different options and the ability for statistical learning might be investigated in further experiments. In the context of perception, experience about determining a perceptual threshold was already gained and used to study perceptional learning without attention and awareness (Watanabe, Náñez, & Sasaki, 2001).

Another aforementioned adjustment would be the design of a more complex OVSST, especially concerning the visual search, to provoke more eye movements in line with common eye movement literature reporting more visual search within more complex settings (Duchowski, 2007; Ehmke & Wilson, 2007; Holmqvist et al., 2011). It could be tested if results of the thesis can be also transferred to more complex test environments. It would be also interesting to investigate when people tend to use the TTB strategy or switch to other strategies like probability matching (Fantino & Esfandiari, 2002), Take The Last or minimalist strategy (Dougherty et al., 2008). An experiment could be conducted which contains a condition with rewards for the prediction of the unlikely exits or costs for incorrect predictions. The Prospect Theory by Kahneman (2011, pp. 278–288) already showed that humans deal with benefits and costs in different ways. The results could be compared with previous studies and thus, the influence of rewards and costs on the selection of the strategy could be analyzed.

In order to extent the external validity of the current results, mentioned as a limitation of the current experiments, it might be interesting to use real workspace scenarios. For instance, users of compute clusters were modelled with data of a newly developed questionnaire for user habits of compute clusters to simulate the human-machine interaction (Schlagkamp, Ferreira da Silva, Renker, & Rinkenauer, 2016). In the same way, the eye movement technique could be used while interacting with a specific interface or a robot to investigate if eye movements are a reliable indicator for uncertainty when interacting with real interfaces or robots. This could be used, for instance, to extent and stress the results concerning approach-avoidance tendencies of the interaction between robots and humans studied by Rinkenauer et al. (2017).

Besides these experimental suggestions for future research, a theoretical issue might relate to the developed three-stage model of learning under uncertainty inspired by earlier developed three-stage models (Ackerman, 1988; Jacob & Hochstein, 2009; Schumacher & Czerwinski, 2014). More evidence should be gathered for the confirmation of the model. Further, the model could be enriched by clarifying the specific role of different eye movement variables regarding the distinct processing steps and the learning environment. For instance, it could be tested if fixation frequency is more sensitive to uncertain situations than fixation duration in accordance with the results in Experiment III.

Another relevant aspect might a clearer differentiation of uncertainty and stress as the theoretical construct might overlap to a great extent. Uncertainty might be a trigger for stress which can be seen as a reaction on a higher cognitive load. The development of stress depended on the evaluation of the situation containing an estimation of the likelihood for a high or low cognitive load in the situation (Ulich, 2011, p. 473). Thus, there seemed to be a strong coupling between stress and subjective uncertainty defined as the perceived uncertainty by the individual (see Chapter 1). In literature, a higher level of stress was associated with shorter fixations (Holmqvist et al., 2011, p. 383) comparable with the experimental results for a higher level of subjective uncertainty in this thesis. In future studies, stress might be tested separately, for example, via questionnaire to be able to differentiate the constructs in a clearer manner.

The field of eye movements during decision making under uncertainty seemed to be a new research field that just started to develop which provides a lot of different directions for future research that should also consider the limits of eye movement analysis. This thesis provided a lot of additional aspects that were beyond the scope of the current study but should be investigated in more detail like eye blinks (see Experiment 1) or recurrence plots (see General Method) in the context of uncertain learning environments. Especially, the analysis of eye blinks seems to be a promising approach to extend the investigation of eye movement behavior.

8.5 Conclusion

This thesis tried to shed light on the development of mental models under uncertainty by investigating visual search behavior. In conclusion, this thesis showed that learning processes as well as the subjective uncertainty were reflected by eye movements. The findings could be integrated into a three-stage model of eye movements during learning under uncertainty providing an extension of existing models. Uncertainty as well as visual search behavior is reduced with increased learning progress. Importantly, eye movement parameters entailed specific characteristics and were depended on various aspects like the design of the task, initial knowledge or individual characteristics that have to be considered when interpreting results.

Reference List

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*(3), 288–318.

Ackerman, P. L. (1990). A correlational analysis of skill specificity: Learning, abilities, and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(5), 883–901.

Adamson, R. E. (1952). Functional fixedness as related to problem solving: A repetition of three experiments. *Journal of Experimental Psychology, 44*(4), 288–291.

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature, 387*(6631), 401–406.

Alegria, J., & Bertelson, P. (1970). Time uncertainty, number of alternatives and particular signal-response pair as determinants of choice reaction time. *Acta Psychologica, 33*, 36–44.

Anderson, J. R., Bothell, D., & Douglass, S. (2004). Eye movements do not reflect retrieval processes: Limits of the Eye-Mind Hypothesis. *Psychological Science, 15*(4), 225–231.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*(4), 369–406.

Anderson, N. C., Bischof, W. F., Laidlaw, K., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavior research methods, 45*(3), 842–856.

Arieli, A., Ben-Ami, Y., & Rubinstein, A. (2011). Tracking decision makers under uncertainty. *American Economic Journal: Microeconomics, 3*(4), 68–76.

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*(8), 437–443.

Bammer, G., & Smithson, M. (2008). *Uncertainty and risk: Multidisciplinary perspectives. Earthscan risk in society series.* London, Sterling, VA: Earthscan.

Bates, M. E., & Lemay, E. P. (2004). The d2 Test of attention: Construct validity and extensions in scoring techniques. *Journal of the International Neuropsychological Society: JINS, 10*(3), 392–400.

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language, 83*, 62–78.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (2005). The Iowa Gambling Task and the Somatic Marker Hypothesis: Some questions and answers. *Trends in Cognitive Sciences, 9*(4), 159–162.

Bekkering, H., & Neggers, S. F. (2002). Visual search is modulated by action intentions. *Psychological Science, 13*(4), 370–374.

Bennett, K. B., & Flach, J. (2011). *Display and interface design: Subtle science, exact art.* Boca Raton, London: CRC Press.

Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and Take-The-Best Models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(1), 107–129.

Betsch, T., Funke, J., & Plessner, H. (2011). *Denken – Urteilen, Entscheiden, Problemlösen.* Berlin, Heidelberg: Springer.

Beuer-Krüssel, M., & Krumpal, I. (2009). Der Einfluss von Häufigkeitsformaten auf die Messung von subjektiven Wahrscheinlichkeiten. *Methoden – Daten – Analysen: Zeitschrift für Empirische Sozialforschung, 3*(1), 31–57.

Boehm-Davis, D. A., Gray, W. D., & Schoelles, M. J. (2016). The eye blink as a physiological indicator of cognitive workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 44*(33), 6-116-6-119.

Boucheix, J.-M., & Lowe, R. K. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction, 20*(2), 123–135.

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry, 52*(10), 976–986.

Bowden, E. M. (1997). The effect of reportable and unreportable hints on anagram solution and the Aha! experience. *Consciousness and Cognition, 6*(4), 545–573.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision, 11*(5), 4.

Brickenkamp, R. (1994). *Test D2: Aufmerksamkeits-Belastungstest* (2nd ed.). Göttingen: Hogrefe.

Brickenkamp, R., & Zillmer, E. (1998). *The D2 test of attention.* Toronto: Hogrefe & Huber.

Brown, C. M. (1999). *Human-computer interaction design guidelines.* Exeter, England: Intellect.

Brunyé, T. T., & Gardony, A. L. (2017). Eye tracking measures of uncertainty during perceptual decision making. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology, 120*, 60–68.

Buelow, M. T., & Suhr, J. A. (2009). Construct validity of the Iowa Gambling Task. *Neuropsychology review, 19*(1), 102–114.

Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology, 4*(3), 121–138.

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science (New York), 315*(5820), 1860–1862.

Camerer, C., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty, 5*(4), 325–370.

Canham, M., & Hegarty, M. (2010). Effects of knowledge and display design on comprehension of complex graphics. *Learning and Instruction, 20*(2), 155–166.

Chalmers, P. A. (2003). The role of cognitive theory in human–computer interface. *Computers in Human Behavior, 19*(5), 593–607.

Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & Cognition, 29*(8), 1146–1152.

Chein, J. M., & Schneider, W. (2012). The brain's learning and control architecture. *Current Directions in Psychological Science, 21*(2), 78–84.

Chen, S. Y., Ross, B. H., & Murphy, G. L. (2016). Eyetracking reveals multiple-category use in induction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(7), 1050–1067.

Chi, M. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *Handbook of research on conceptual change* (pp. 61–82). Hillsdale, NJ: Erlbaum.

Chi, M. T. H., & Roscoe, R. D. (2002). The processes and challenges of conceptual change. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 3–27). Dordrecht, Boston: Kluwer Academic Publishers.

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science, 10*(4), 360–365.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*(1), 28–71.

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: Bottom-up versus top-down. *Current biology, 14*(19), 850–852.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in quantitative methods for psychology, 1*(1), 42–45.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268–294.

Craik, K. (1943). *The nature of explanation.* Cambridge, UK: Cambridge University Press.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience, 18*, 193–222.

Dole, J. A., & Sinatra, G. M. (1998). Reconceptalizing change in the cognitive construction of knowledge. *Educational Psychologist, 33*(2-3), 109–128.

Donnell, M. L. (1996). Human cognition and the expert systems interface mental models and explanation facilities. In C. A. Ntuen & E. H. Park (Eds.), *Human Interaction with Complex Systems* (pp. 343–349). Boston, MA: Springer US.

Dougherty, M. R., Franco-Watkins, A. M., & Thomas, R. (2008). Psychological plausibility of the theory of Probabilistic Mental Models and the fast and frugal heuristics. *Psychological Review, 115*(1), 199–213.

Doya, K. (2008). Modulators of decision making. *Nature neuroscience, 11*(4), 410–416.

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice* (2nd ed.). London: Springer.

Durso, F. T., & Gronlund, S. D. (1999). Situation Awareness: Chapter 10. In *F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindays & M. T. H. Chi (Eds.), Handbook of applied cognition* (pp. 283–314). Chichester: John Wiley & Sons.

Dutke, S. (1994). *Mentale Modelle: Konstrukte des Wissens und Verstehens.* Göttingen: Verlag für Angewandte Psychologie.

Eckstein, M. P. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science, 9*(2), 111–118.

Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision, 11*(5), 14.

Edwards, W. (1961). Probability learning in 1000 trials. *Journal of Experimental Psychology, 62*(4), 385–394.

Ehmke, C., & Wilson, S. (2007). *Identifying web usability problems from eye-tracking data.* Paper presented at the British HCI Conference 2007. University of Lancaster, U. K.

Ellis, J. (2012). *Using eye movements to investigate insight problem solving.* Doctoral Dissertation, Toronto. Retrieved February 9, 2017, from https://tspace.library.utoronto.ca/handle/1807/33987.

Engelkamp, J. (2006). *Lehrbuch der kognitiven Psychologie.* Göttingen: Hogrefe.

Epelboim, J., & Suppes, P. (1997). Eye movements during geometrical problem solving. *Perception, 26*(1), 138.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*(3), 215–251.

Evans, J., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science: a journal of the Association for Psychological Science, 8*(3), 223–241.

Evans, J. S. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*(10), 454–459.

Fantino, E., & Esfandiari, A. (2002). Probability matching: Encouraging optimal responding in humans. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 56*(1), 58–63.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Method, 39*(2), 175–191.

Fellows, L. K., & Farah, M. J. (2007). The role of ventromedial prefrontal cortex in decision making: Judgement under uncertainty or judgement per se? *Cerebral Cortex, 17*(11), 2669–2674.

Feng, G. (2003). From eye movement to cognition: Toward a general framework of inference comment on Liechty et al., 2003. *Psychometrika, 68*(4), 551–556.

Fischhoff, B., & Beyth, R. (1975). I knew it would happen. *Organizational Behavior and Human Performance, 13*(1), 1–16.

Fishburn, P. C. (1970). *Utility theory for decision making.* New York: Wiley.

Flombaum, J. I., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance, 32*(4), 840–853.

Fogarty, C., & Stern, J. A. (1989). Eye movements and blinks: Their relationship to higher cognitive processes. *International Journal of Psychophysiology, 8*(1), 35–42.

Fox, J., Merwin, D., Marsh, R., McConkie, G., & Kramer, A. (1996). Information extraction during instrument flight: An evaluation of the validity of the eye-mind hypothesis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 40*(2), 77–81.

Galitz, W. O. (2007). *The essential guide to user interface design: An introduction to GUI design principles and techniques* (3rd ed.). Indianapolis: Wiley.

Gentner, D. (2001). Mental Models, Psychology of. In N. J. Smelser (Ed.), *International encyclopedia of the social & behavioral sciences* (1st ed., pp. 9683–9687). Amsterdam: Elsevier, Pergamon.

Gentner, D., & Stevens, A. L. (1983). *Mental models. Cognitive science.* Hillsdale, N.J.: Erlbaum.

Gidlöf, K., Wallin, A., Dewhurst, R., & Holmqvist, K. (2013). Using eye tracking to trace a cognitive process: Gaze behaviour during decision making in a natural environment. *Journal of Eye Movement Research, 6*(1).

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology, 62*, 451–482.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*(4), 650–669.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic Mental Models: A Brunswikian theory of confidence. *Psychological Review, 98*(4), 506–528.

Gilchrist, I. D. (2011). Saccades. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 85–94). Oxford, New York: Oxford University Press.

Glaholt, M. G., & Reingold, E. M. (2009). The time course of gaze bias in visual decision tasks. *Visual Cognition, 17*(8), 1228–1243.

Glaholt, M. G., Wu, M., & Reingold, E. M. (2009). Predicting preference from fixations. *PsychNology Journal, 7*(2), 141–158.

Glaholt, M. G., Wu, M.-C., & Reingold, E. M. (2010). Evidence for top-down control of eye movements during visual decision making. *Journal of Vision, 10*(5), 15.

Glöckner, A., & Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making, 24*(1), 71–98.

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics, 24*(6), 631–645.

Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002). *Eye tracking in web search tasks: Design implications. In ETRA '02 Proceedings of the 2002 symposium on Eye tracking research & applications,* pp. 51–58.

Goldberg, J. H., & Wichansky, A. M. (2003). Eye tracking in usability evaluation: A practitioner's guide. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 493–516). Amsterdam, Boston: North-Holland.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14*(5), 462–466.

Haider, H., & Frensch, P. A. (1999). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(1), 172–190.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Hauer, B. J., & MacLeod, C. M. (2006). Endogenous versus exogenous attentional cuing effects on memory. *Acta Psychologica, 122*(3), 305–320.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7*(2), 185–207.

Hegarty, M., & Just, A. M. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32*(6), 717–742.

Hegarty, M., Mayer, R. E., & Green, C. E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. *Journal of Educational Psychology, 84*(1), 76.

Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 210–228.

Hewson, M. G., & Hewson, P. W. (1983). Effect of instruction using students' prior knowledge and conceptual change strategies on science learning. *Journal of Research in Science Teaching, 20*(8), 731–743.

Hilbert, M. (2012). Toward a synthesis of cognitive biases: how noisy information processing can bias human decision making. *Psychological Bulletin, 138*(2), 211–237.

Hoffman, A. B., & Rehder, B. (2010). The costs of supervised classification: The effect of learning task on conceptual flexibility. *Journal of Experimental Psychology: General, 139*(2), 319–340.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57*(6), 787–795.

Holmqvist, K., Nyström, M., & Andersson, R. (2011). *Eye tracking: A comprehensive guide to methods and measures.* Oxford, New York: Oxford University Press.

Horstmann, N., Ahlgrimm, A., & Glöckner, A. (2009). How distinct are intuition and deliberation? An eye-tracking analysis of instruction-induced decision modes. *Judgement and Decision Making, 4*(5), 335–354.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science, 310*(5754), 1680–1683.

Huang, L., & Pashler, H. (2005). Attention capacity and task difficulty in visual search. *Cognition, 94*(3), 101–111.

Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of experimental psychology. General, 130*(4), 658–680.

Hyönä, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction, 20*(2), 172–176.

Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 105–133). New York: Psychology Press.

Ishihara, S. (1917). *Tests for Colour Blindness.* Tokyo: Handaya Hongo Harukich.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(11), 1254–1259.

Jacob, M., & Hochstein, S. (2009). Comparing eye movements to detected vs. undetected target stimuli in an Identity Search Task. *Journal of Vision, 9*(5), 20–2016.

Jacob, R. J., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573–605). Amsterdam, Boston: North-Holland.

Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*, 146–154.

Jarodzka, H., van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see: Guiding students' attention via a model's eye movements fosters learning. *Learning and Instruction, 25*, 62–70.

Jipp, M. (2016). Expertise development with different types of automation a function of different cognitive abilities. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 58*(1), 92–106.

Johansson, R., Holmqvist, K., Mossberg, F., & Lindgren, M. (2011). Eye movements and reading comprehension while listening to preferred and non-preferred study music. *Psychology of Music, 40*(3), 339–356.

Johnson, J. (2014). *Designing with the mind in mind: Simple guide to understanding user interface design guidelines.* Amsterdam: Morgan Kaufmann.

Johnson, J., & Henderson, A. (2012). *Conceptual models: Core to good design.* San Rafael, Calif.: Morgan & Claypool.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, UK: Cambridge University Press.

Johnson-Laird, P. N., & Savary, F. (1996). Illusory inferences about probabilities. *Acta Psychologica, 93*(1-3), 69–90.

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences of the United States of America, 107*(43), 18243–18250.

Jones, N., Ross, H., Lynam, T., Perez, P., & Leitch, A. (2011). Mental models: An interdisciplinary synthesis of theory and methods. *Ecology and Society, 16*(1), 46.

Jones, S. H., Johnson, M. L., & Campbell, B. D. (2015). Hot factors for a cold topic: Examining the role of task-value, attention allocation, and engagement on conceptual change. *Contemporary Educational Psychology, 42*, 62–70.

Jungé, J. A., Scholl, B. J., & Chun, M. M. (2007). How is spatial context learning integrated over signal versus noise? A primacy effect in contextual cueing. *Visual Cognition, 15*(1), 1–11.

Jungermann, H., Fischer, K., & Pfister, H.-R. (2010). *Die Psychologie der Entscheidung: Eine Einführung* (3. Aufl.). Heidelberg, Neckar: Spektrum, Akad.-Verl.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*(4), 441–480.

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2002). Perspective effects on online text processing. *Discourse Processes, 33*(2), 159–173.

Kabata, T., & Matsumoto, E. (2012). Cueing effects of target location probability and repetition. *Vision Research, 73*, 23–29.

Kahneman, D. (2011). *Thinking fast and slow.* New York: Farrar, Straus, and Giroux.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*(4), 237–251.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22*(3), 374–419.

Kim, S., & Rehder, B. (2011). How prior knowledge affects selective attention during category learning: An eyetracking study. *Memory & Cognition, 39*(4), 649–665.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163–182.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences, 4*(4), 138–147.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition, 29*(7), 1000–1009.

Koch, I., Gade, M., Schuch, S., & Philipp, A. M. (2010). The role of inhibition in task switching: A review. *Psychonomic Bulletin & Review, 17*(1), 1–14.

Kok, E. M., & Jarodzka, H. (2016). Before your very eyes: The value and limitations of eye tracking in medical education. *Medical Education*, 1–9.

Kristjánsson, Á. (2011). The intriguing interactive relationship between visual attention and saccadic eye movements. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 455–469). Oxford, New York: Oxford University Press.

Kristjánsson, Á., Mackeben, M., & Nakayama, K. (2001). Rapid, object-based learning in the deployment of transient attention. *Perception, 30*(11), 1375–1387.

Kumada, T. (1999). Limitations in attending to a feature value for overriding stimulus-driven interference. *Perception & Psychophysics, 61*(1), 61–79.

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., et al. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review, 10*, 90–115.

Langan-Fox, J., Armstrong, K., Balvin, N., & Anglim, J. (2002). Process in skill acquisition: Motivation, interruptions, memory, affective states, and metacognition. *Australian Psychologist, 37*(2), 104–117.

Laude, J. R., Pattison, K. F., & Zentall, T. R. (2012). Hungry pigeons make suboptimal choices, less hungry pigeons do not. *Psychonomic bulletin & review, 19*(5), 884–891.

Lee, J. D., Wickens, C. D., Liu, Y., & Boyle, L. N. (2017). *Designing for people: An introduction to human factors engineering* (3rd Edition). Charleston: CreateSpace.

Lemaire, P., & Lecacheur, M. (2010). Strategy switch costs in arithmetic problem solving. *Memory & Cognition, 38*(3), 322–332.

Lipshitz, R., & Strauss, O. (1997). Coping with uncertainty: A naturalistic decision-making analysis. *Organizational Behavior and Human Decision Processes, 69*(2), 149–163.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4*(1), 6–14.

Liversedge, S. P., Gilchrist, I. D., & Everling, S. (Eds.) (2011). *The Oxford handbook of eye movements.* Oxford, New York: Oxford University Press.

Locke, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational Behavior and Human Performance, 3*(2), 157–189.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic bulletin & review, 1*(4), 476–490.

Maisto, A. A., & Baumeister, A. A. (1975). A developmental study of choice reaction time: The effect of two forms of stimulus degradation on encoding. *Journal of Experimental Child Psychology, 20*(3), 456–464.

Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences, 9*(6), 296–305.

Marwan, N. (2017). *Recurrence plots and cross recurrence plots.* Retrieved May 05, 2017, from http://recurrence-plot.tk/rqa.php.

Marwan, N., Romano, M., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports, 438*(5-6), 237–329.

Mattes, S., Ulrich, R., & Miller, J. (2002). Response force in RT tasks: Isolating effects of stimulus probability and response probability. *Visual Cognition, 9*(4-5), 477–501.

Miller, J. (1998). Effects of stimulus-response probability on choice reaction time: Evidence from the lateralized readiness potential. *Journal of Experimental Psychology: Human Perception and Performance, 24*(5), 1521–1534.

Mohammed, S., Ferzandi, L., & Hamilton, K. (2010). Metaphor no more: A 15-year review of the team mental model construct. *Journal of Management, 36*(4), 876–910.

Molet, M., Miller, H. C., Laude, J. R., Kirk, C., Manning, B., & Zentall, T. R. (2012). Decision making by humans in a behavioral task: Do humans, like pigeons, show suboptimal choice? *Learning & behavior, 40*(4), 439–447.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4*(2), 61–64.

Müller, H. J., & Krummenacher, J. (2006). Visual search and selective attention. *Visual Cognition, 14*(4-8), 389–410.

Mulligan, N. W. (1998). The role of attention during encoding in implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(1), 27–47.

Najemnik, J., & Geisler, W. G. (2005). Optimal eye movement strategies in visual search. *Nature, 435*(7031), 387–391.

Neumann, J. von, & Morgenstern, O. (1947). *Theory of games and economic behavior.* Princeton: Princeton University Press.

Newell, A., & Rosenbloom, P. S. (2013). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Carnegie Mellon Symposia on Cognition Series. Cognitive Skills and Their Acquisition* (pp. 1–55). London: Taylor and Francis.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175–220.

Nyström, M., Andersson, R., Holmqvist, K., & van de Weijer, J. (2013). The influence of calibration method and eye physiology on eyetracking data quality. *Behavior Research Methods, 45*(1), 272–288.

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior research methods, 42*(1), 188–204.

O'Brien, F., & Cousineau, D. (2014). Representing error bars in within-subject designs in typical software packages. *The Quantitative Methods for Psychology, 10*(1), 56–67.

Olson, M. H., & Hergenhahn, B. R. (2013). *An introduction to theories of learning* (9th ed.). Upper Saddle River, N.J.: Pearson Prentice Hall.

Orquin, J. L., Chrobot, N., & Grunert, K. G. (in press). Guiding decision makers' eye movements with (un)predictable object locations. *Journal of Behavioral Decision Making*.

Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica, 144*(1), 190–206.

Paulus, M. P., Hozack, N., Zauscher, B., McDowell, J. E., Frank, L., Brown, G. G., & Braff, D. L. (2001). Prefrontal, parietal, and temporal cortex networks underlie decision-making in the presence of uncertainty. *NeuroImage, 13*(1), 91–100.

Piaget, J. (1952). *The origins of intelligence in children.* New York: International Universities Press.

Platt, M. L., & Huettel, S. A. (2008). Risky business: The neuroeconomics of decision making under uncertainty. *Nature neuroscience, 11*(4), 398–403.

Pomerol, J.-C., & Brézillon, P. (2003). Context proceduralization in decision making. In P. Blackburn (Ed.), *Lecture Notes in Computer Science. International and Interdisciplinary Conference on Modeling and Using Context* (pp. 491–498). Berlin, New York: Springer.

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. *Enceyclopedia of human computer interaction 1*, 211–219.

Poole, A., Ball, L. J., & Phillips, P. (2005). In search of salience: A response-time and eye-movement analysis of bookmark recognition. In S. Fincher (Ed.), *People and computers XVIII. Design for life: Proceedings of HCI 2004* (pp. 363–378). London: Springer.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Preim, B., & Dachselt, R. (2010). *Interaktive Systeme.* Berlin, Heidelberg: Springer Berlin Heidelberg.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9*(7), 545–556.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.

Rayner, K., Li, X., Williams, C. C., Cave, K. R., & Well, A. D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research, 47*(21), 2714–2726.

Reichel, G. (1985). *Grundlagen der Arbeitsmedizin.* Stuttgart: Kohlhammer.

Reis, J. (1996). *Inventar zur Messung der Ambiguitätstoleranz (IMA).* Heidelberg: Asanger.

Renker, J., & Rinkenauer, G. (2016). The acquisition of mental representations under uncertainty. *Kognitive Systeme, 3*(1). Retrieved January 02, 2017, from http://duepublico.uni-duisburg-essen.de/servlets/DocumentServlet?id=42388.

Rheinberg, F., Vollmeyer, R., & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktuller Motivation in Lern- und Leistungssituationen. *Diagnostica, 47*(2), 57–66.

Richer, F., & Beatty, J. (1987). Contrasting effects of response uncertainty on the task-evoked pupillary response and reaction time. *Psychophysiology, 24*(3), 258–262.

Rieskamp, J., & Hoffrage, U. (2005). The use of simple heuristics: Inferences and preferences under time pressure. *Manuscript, Max Planck Institute for Human Development, Berlin.*

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of experimental psychology. General, 135*(2), 207–236.

Rinkenauer, G., Böckenkamp, A., & Weichert, F. (2017). Man-robot collaboration in the context of industry 4.0: Approach-avoidance tendencies as an indicator for the affective quality of interaction? In C. M. Schlick (Ed.), *Advances in Ergonomic Design of Systems, Products and Processes. Proceedings of the Annual Meeting of GfA 2016* (pp. 335–348). Berlin: Springer.

Rolke, B., & Hofmann, P. (2007). Temporal uncertainty degrades perceptual processing. *Psychonomic bulletin & review, 14*(3), 522–526.

Savage, L. J. (1954). *The foundations of statistics.* New York: Wiley.

Scherrmann, J.-M., Wolff, K., Franco, C. A., Potenza, M. N., Uzbay, T., Bizarro, L., et al. (2010). Attentional set. In I. P. Stolerman (Ed.), *Encyclopedia of psychopharmacology* (pp. 181–182). Berlin, Heidelberg: Springer.

Schlagkamp, S., Ferreira da Silva, R., Renker, J., & Rinkenauer, G. (2016). Analyzing users in parallel computing: A user-oriented study. In W. W. Smari & V. Zeljkovic (Eds.), *Proceedings of the 2016 International Conference on High Performance Computing & Simulation (HPCS 2016). July 18-22, 2016, Innsbruck, Austria* (pp. 395–402). Piscataway, NJ: IEEE.

Schneider, W., & Chein, J. M. (2003). Controlled & automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science, 27*(3), 525–559.

Scholz, A., Helversen, B. von, & Rieskamp, J. (2015). Eye movements reveal memory processes during similarity- and rule-based decision making. *Cognition, 136*, 228–246.

Scholz, A., Melhorn, K., Bocklisch, F., & Krems, J. F. (2011). Looking at nothing diminishes with practice. In *Proceedings of the 33rd annual meeting of the Cognitive Science Society* (pp. 1070–1075).

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General, 122*(2), 166–183.

Schulze, C., van Ravenzwaaij, D., & Newell, B. R. (2015). Of matchers and maximizers: How competition shapes choice under risk and uncertainty. *Cognitive Psychology, 78*, 78–98.

Schumacher, R. M., & Czerwinski, M. P. (2014). Mental models and the acquisition of expert knowledge. In R. R. Hoffman (Ed.), *The Psychology of Expertise: Cognitive Research and Empirical Ai* (pp. 61–79). Taylor & Francis.

Shaw, M. L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception and Performance, 3*(2), 201–211.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127–190.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6*(12), 1317–1322.

Shuper, P. A., Sorrentino, R. M., Otsubo, Y., Hodson, G., & Walker, A. M. (2004). A theory of uncertainty orientation: Implications for the study of individual differences within and across cultures. *Journal of Cross-Cultural Psychology, 35*(4), 460–480.

Simon, H. A. (2000). Bounded rationality in social science: Today and tomorrow. *Mind & Society, 1*(1), 25–39.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual review of psychology, 28*(1), 1–39.

Smithson, M. (2009). Psychology's ambivalent view of uncertainty. In G. Bammer & M. Smithson (Eds.), *Earthscan risk in society series. Uncertainty and risk. Multidisciplinary perspectives* (pp. 205–217). London, Sterling, VA: Earthscan.

Solman, G., & Smilek, D. (2010). Item-specific location memory in visual search. *Vision research, 50*(23), 2430–2438.

Speelman, C., & Kirsner, K. (2008). *Beyond the learning curve: The construction of mind. Oxford Scholarship Online.* Oxford: Oxford Univ. Press.

Spivey, M. J., & Dale, R. (2011). Eye movements both reveal and influence problem solving. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 551–562). Oxford, New York: Oxford University Press.

Starling, S. J. (2012). *Learning and behavior in an uncertain world: Probability learning in children and adults.* Doctoral Dissertation, University of Rochester, New York. Retrieved November 10, 2016, from http://hdl.handle.net/1802/26701.

Stern, J. A., Boyer, D., & Schroeder, D. (1994). Blink rate: A possible measure of fatigue. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 36*(2), 285–297.

Sternberg, S. (1967). Two operations in character recognition: Some evidence from reaction-time measurements. *Perception & Psychophysics, 2*(2), 45–53.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285.

Swensson, R. G. (1965). Incentive shifts in a three-choice decision situation. *Psychonomic Science, 2*(1-12), 101–102.

Tai, R. H., Loehr, J. F., & Brigham, F. J. (2006). An exploration of the use of eye-gaze tracking to study problem-solving on standardized science assessments. *International Journal of Research & Method in Education, 29*(2), 185–208.

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision, 7*(14), 4.

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica, 135*(2), 77–99.

Theeuwes, J., Belopolsky, A., & Olivers, C. N. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica, 132*(2), 106–114.

Thompson, J. D. (1967). *Organizations in action: Social science bases of administrative theory.* New Brunswick, NJ: Transaction Publishers.

Thorndike, E. L. (1927). The law of effect. *The American Journal of Psychology, 39*(1/4), 212.

Thornton, T. L., & Gilden, D. L. (2007). Parallel and serial processes in visual search. *Psychological Review, 114*(1), 71–103.

Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 353*(1373), 1295–1306.

Treisman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing, 31*(2), 156–177.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136.

Tsai, M.-J., Hou, H.-T., Lai, M.-L., Liu, W.-Y., & Yang, F.-Y. (2012). Visual attention for solving multiple-choice science problem: An eye-tracking analysis. *Computers & Education, 58*(1), 375–385.

Tymula, A., Belmaker, L. A. R., Roy, A. K., Ruderman, L., Manson, K., Glimcher, P. W., & Levy, I. (2012). Adolescents' risk-taking behavior is driven by tolerance to ambiguity. *Proceedings of the National Academy of Sciences, 109*(42), 17135–17140.

Ulich, E. (2011). *Arbeitspsychologie* (7.th ed.). Zürich: vdf Hochschulverlag.

van der Gijp, A., Ravesloot, C. J., Jarodzka, H., van der Schaaf, M. F., van der Schaaf, I. C., van Schaik, J. P. J., & Cate, T. J. (2016). How visual search relates to visual diagnostic performance: A narrative systematic review of eye-tracking research in radiology. *Advances in health sciences education: theory and practice*, 1–23.

van Dijk, T. A., & Kintsch, W. (1992). *Strategies of discourse comprehension* (3.th ed.). San Diego: Academic Press.

van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Computer in Human Behaviour, 25*(3), 785–791.

van Merriënboer, J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*(2), 147–177.

van Zoest, W., & Donk, M. (2004). Bottom-up and top-down control in visual search. *Perception, 33*(8), 927–937.

van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 746–759.

Velichkovsky, B. M., Rothert, A., Kopf, M., Dornhöfer, S. M., & Joos, M. (2002). Towards an express-diagnostics for level of processing and hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour, 5*(2), 145–156.

Venkatraman, V., Payne, J. W., & Huettel, S. A. (2014). An overall probability of winning heuristic for complex risky decisions: Choice and eye fixation evidence. *Organizational Behavior and Human Decision Processes, 125*(2), 73–87.

Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics, 56*(5), 495–500.

Wascher, E., Heppner, H., Möckel, T., Kobald, S. O., & Getzmann, S. (2015). Eye-blinks in choice response tasks uncover hidden aspects of information processing. *EXCLI journal, 14*, 1207–1218.

Watanabe, T., Náñez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature, 413*(6858), 844–848.

Weber, M., & Camerer, C. (1987). Recent developments in modelling preferences under risk. *OR Spektrum, 9*(3), 129–151.

Wedel, M., & Pieters, R. (2008). A review of eye-tracking research in marketing. In N. K. Malhotra (Ed.), *Review of marketing research: Vol. 4. Review of marketing research* (pp. 123–147). Armonk, N.Y., London: M.E. Sharpe.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3. ed.). Upper Saddle River, NJ: Prentice-Hall.

Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202–238.

Wolford, G., Miller, M. B., & Gazzaniga, M. (2000). The left hemisphere's role in hypothesis formation. *The Journal of neuroscience, 20*(6), 1–4.

Yarbus, A. L. (Ed.) (1967). *Eye Movements and Vision.* Boston, MA: Springer US; Imprint; Springer.

Yi, M. Y., & Davis, F. D. (2003). Developing and validating an observational learning model of computer software training and skill acquisition. *Information Systems Research, 14*(2), 146–169.

Young, I., & Veen, J. (2008). *Mental models: Aligning design strategy with human behavior.* Brooklyn, N.Y.: Rosenfeld Media.

Yu, A. J., & Huang, H. (2014). Maximizing masquerading as matching in human visual search choice behavior. *Decision, 1*(4), 275–287.

Ziori, E., & Dienes, Z. (2012). The time course of implicit and explicit concept learning. *Consciousness and Cognition, 21*(1), 204–216.

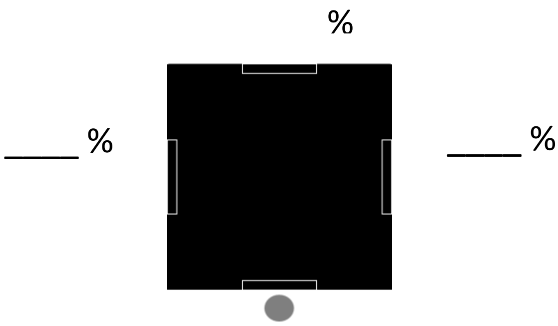
Zuckerman, M. (1979). *Sensation seeking: Beyond the optimal level of arousal.* Hillsdale, N.J., New York: L. Erlbaum Associates.

Appendix A: Concept Awareness Questionnaire

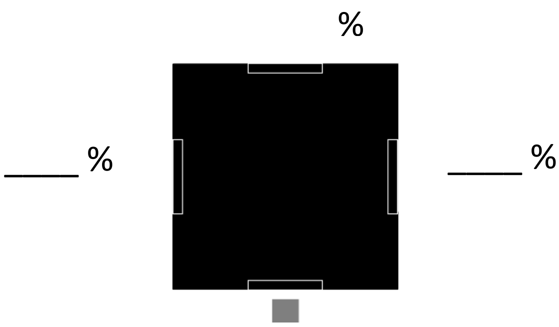
Do you think the object (circle, square, triangle) influences the exit where the symbol reappears?

Yes ☐ No ☐

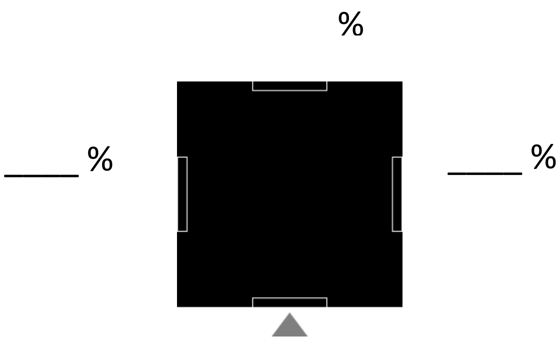
Please estimate the probability of the circle reappearing at each of the exits. Write your answer in the following figure (in sum it has to be 100%).



Please estimate the probability of the square reappearing at each of the exits. Write your answer in the following figure (in sum it has to be 100%).



Please estimate the probability of triangle reappearing at each of the exits. Write your answer in the following figure (in sum it has to be 100%).



Appendix B: Procedure and Results of the Eye Tracking Validation Study

In the following, the method and the results of the validation study, comparing the SMI Red 500 and EyeLink 1000 eye trackers are briefly described. After a 5-point calibration and 4-point validation, the experimental condition started. A central cross appeared for 1 sec followed by one of the three dark gray objects (circle, triangle, square) appearing randomly at one of nine different positions (in the middle of the screen, 11.5 cm left, top, right and down from the center and 45°, 135°, 225°, 315° in between). As shown in Figure 9.1 either objects remained dark gray and participants were instructed to focus the eyes on the appeared object (no-go trial) or the object became light gray after 1 sec and participants had to react as quickly as possible by pressing the space bar (go trial). Trials with a longer reaction time than 2 sec were declared as errors. Every object appeared two times at the same position, so that 54 trials had to be performed subdivided into three blocks. One-third of the trials were go trials. Participants performed the experiment two times, once with the SMI eye tracker and once with the EyeLink. The order of the eye tracker model was randomized. In total 10 participants (7 female) with mean age of 24 years (SD=2.4) were tested. Results of correlation analysis with Spearman's Rho showed that fixation time (r=-.020, p=.645), fixation count (r=.010, p=.824) as well as reaction time (r=-.083, p=.266) was not significantly associated with the object type. The reaction time of the participants to the circle (M=.458, SD=.045), the triangle (M=.458, SD=.060) and the square (M=.458, SD=.065) did not differ significantly: t(9)=0.01, p=.995 for the difference between triangle and circle, t(9)=-.015, p=.988 concerning circle and square and t(9)=-.018, p=.986 for the pairing triangle and square. The comparison of the SMI and EyeLink eye tracker showed that eye movement data recorded by the SMI eye tracker were more precise as scatter plots of fixations on the target objects at the nine positions showed lower standard deviations M=.567, SD=.080 for x-coordinates and M=.764, SD=.142 for y-coordinates than data resulted from the EyeLink (x-coordinate: M=.738, SD=.094, y-coordinate: M=.956, SD=.200). Planned t-tests showed that these differences were highly significant for the x-coordinates, t(8)=5.33, p<.001, and marginal significant (p<.10) for the y-coordinates, t(8)=1.90, p=.093. However, the drift defined as the difference between start and end position of the fixations, viz. the deviation from the mean of the scatter plots for the data recorded by the EyeLink (x-coordinate: M=.012, SD=.059, y-coordinate: M=.126, SD=.218) and the SMI (x-coordinate: M=.030, SD=.116, y-coordinate: M=.021, SD=.105) did not differ significantly (x-coordinates: t(8)=0.36, p=726; y-coordinates: t(8)=1.41, p=.197). In conclusion objects seem to be comparable and thus reliable for the OVSST. Additionally, the precision of the SMI eye tracker seems to be sufficient as well as the overall accuracy. However, a drift check of the eye movement data was implemented before analyzing data of all experiments executed for the thesis to be sure that data points are accurate and thus reliable.

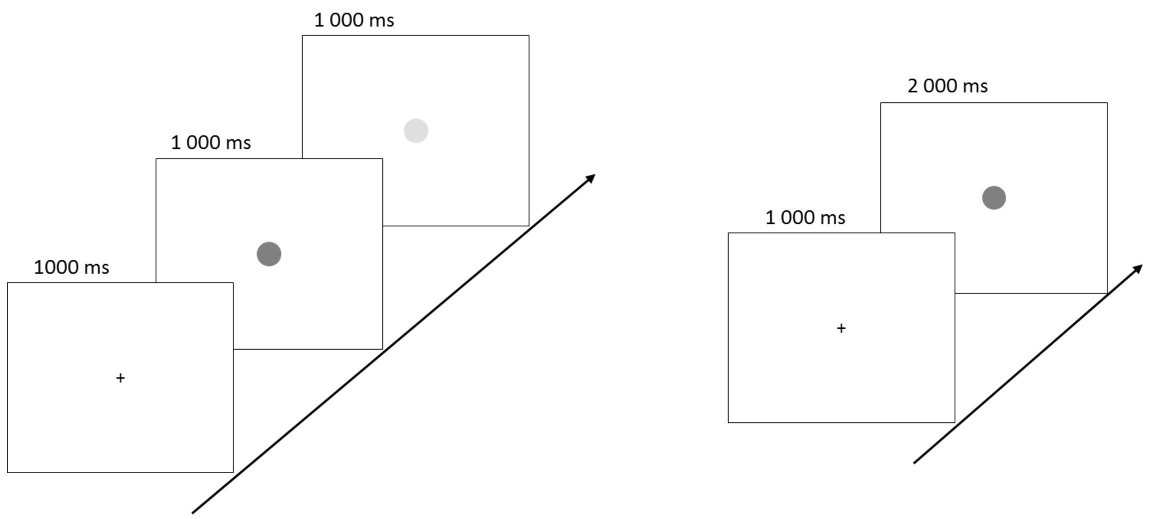


Figure 9.1: Schematic description of the validating task: Go Trials (left) and no-go Trials (right). All trials last 3 seconds (1 sec central cross and 2 sec target object presentation) regardless whether the target object changes the color intensity or not.

Appendix C: Means and Standard Deviations for Trends across Blocks

Experiment I

Table 9.2: Means and standards values for correctly and/or incorrectly predicted trials of marginal or partly significant effects across blocks in Experiment I

Variable	Judgment	Block	Mean	Standard Deviation
Reaction time (ms)	correct	1	953ms	141ms
		2	914ms	132ms
		3	885ms	128ms
		4	880ms	121ms
	incorrect	1	1.148ms	188ms
		2	1.113ms	130ms
		3	1.102ms	137ms
		4	1.100ms	137ms
Number of Predictions	correct	1	48.53	7.87
		2	55.24	4.48
		3	56.41	4.84
		4	56.59	3.99
Blinks per trial	correct	1	2.25	1.51
		2	2.38	1.32
		3	2.66	1.46
		4	3.01	1.94
	incorrect	1	2.13	1.41
		2	2.29	1.30
		3	2.84	1.52
		4	3.14	2.17

Experiment II

Table 9.3: Means and standards values for correctly and/or incorrectly predicted trials of marginal or partly significant effects across blocks in Experiment II

Variable	Judgment	Block	Mean	Standard Deviation
Number of Predictions	correct	1	48.06	9.08
		2	54.29	4.96
		3	56.88	3.52
		4	57.94	2.08

Experiment III

Table 9.4: Means and standards values for correctly and/or incorrectly predicted trials of marginal or partly significant effects across blocks in Experiment III

Variable	Judgment	Block	Mean	Standard Deviation
Reaction time (ms)	correct	1	888ms	148ms
		2	827ms	112ms
		3	808ms	126ms
		4	797ms	124ms
		5	839ms	119ms
		6	809ms	146ms
		7	830ms	135ms
		8	801ms	128ms
	incorrect	1	1.097ms	128ms
		2	1.079ms	128ms
		3	1.089ms	103ms
		4	1.099ms	158ms
		5	1.090ms	133ms
		6	1.073ms	142ms
		7	1.094ms	118ms
		8	1.084ms	132ms

Experiment IV

Table 9.5: Means and standards values for the number of gaze shifts in the reaction and the prediction task across blocks in Experiment IV

Variable	Task	Block	Mean	Standard Deviation
Number of gaze shifts	Prediction	1	4.99	1.11
		2	4.46	1.14
		3	4.39	1.14
		4	4.29	1.11
	Reaction	1	5.17	1.05
		2	5.40	1.13
		3	5.47	1.05
		4	5.46	1.16

Experiment V

Table 9.6: Means and standards deviations for the fixation duration in AOI_{low} for incorrectly and correctly predicted trials across blocks in the high probability condition of Experiment V

Variable	Task	Block	Mean	Standard Deviation
Fixation Duration AOI_{low}	correct	1	0.74ms	0.55ms
		2	0.96ms	0.89ms
		3	0.82ms	0.90ms
		4	0.92ms	1.01ms
	incorrect	1	2.00ms	0.65ms
		2	2.22ms	0.65ms
		3	2.27ms	0.70ms
		4	2.32ms	0.61ms

Table 9.7: Means and standards deviations for the fixation frequency in $AOI_{predict}$ for incorrectly and correctly predicted trials across blocks in the low probability condition of Experiment V

Variable	Task	Block	Mean	Standard Deviation
Fixation Frequency $AOI_{predict}$	correct	1	8.09	3.17
		2	7.79	3.28
		3	7.59	2.59
		4	7.48	3.10
		5	7.31	2.57
		6	7.05	2.64
	incorrect	1	3.93	1.54
		2	3.70	1.43
		3	3.66	1.24
		4	3.61	1.41
		5	3.70	1.21
		6	3.41	1.08

Selbstständigkeitserklärung

Hiermit versichere ich **schriftlich** und **eidesstattlich** gemäß § 11 Abs. 2 PromO v. 08.02.2011/08.05.2013:

1. Die von mir vorgelegte Dissertation ist selbstständig verfasst und alle in Anspruch genommenen Quellen und Hilfen sind in der Dissertation vermerkt worden.
2. Die von mir eingereichte Dissertation ist weder in der gegenwärtigen noch in einer anderen Fassung an der Technischen Universität Dortmund oder an einer anderen Hochschule im Zusammenhang mit einer stattlichen oder akademischen Prüfung vorgelegt worden.

Dortmund, den 05.12.2017 ________________________________

Unterschrift

3. Weiterhin erkläre ich **schriftlich** und **eidesstattlich**, dass mir der „Ratgeber zur Verhinderung von Plagiaten“ und die „Regeln guter wissenschaftlicher Praxis der Technischen Universität Dortmund“ bekannt und von mir in der vorgelegten Dissertation befolgt worden sind.

Dortmund, den 05.12.2017 ________________________________

Unterschrift