

TECHNISCHE UNIVERSITÄT
CHEMNITZ

Philosophische Fakultät
Institut für Psychologie

Diplomarbeit zum Thema

Blickbewegungsmessung als Prozessmaß bei gedächtnisbasierten Multi-Attribut-Entscheidungen

Agnes Scholz

betreut von:

Prof. Dr. Josef F. Krems
Prof. Dr. Georg Jahn
Dr. Frank Renkewitz

Berlin, August 2009

Inhaltsverzeichnis

Zusammenfassung	4
1 Einleitung	5
2 Theoretischer Hintergrund	7
2.1 Gedächtnisbasierte Multi-Attribut-Entscheidungen.....	7
2.1.1 Kompensatorische vs. nichtkompensatorische Strategien	9
2.1.2 Empirische Befunde zu Multi-Attribut-Entscheidungen	12
2.1.3 Empirische Befunde zu gedächtnisbasierten Entscheidungen	13
2.1.4 Zusammenfassung.....	14
2.2 Methoden der deskriptiven Entscheidungsforschung	15
2.2.1 Outcome basierte Methoden	16
2.2.1.1 Structural Modeling.....	16
2.2.1.2 Regressionsbasierte Strategieklassifikation.....	17
2.2.1.3 Bayesianische Strategieklassifikation.....	18
2.2.2 Prozessbasierte Methoden.....	19
2.2.2.1 Informationstafeln	19
2.2.2.2 Mouselab und Eye Tracking.....	20
2.2.2.3 Lautes Denken	22
2.2.2.4 Entscheidungszeiten	23
2.2.3 Diskussion der Methoden.....	23
2.2.4 Anforderungen an ein multimethodales Vorgehen	26
2.3 Blickbewegungsmessung als Prozessmaß.....	27
2.3.1 Looking at nothing-Phänomen.....	28
2.4 Fragestellung und experimentelle Umsetzung	32
3 Methode	36
3.1 Design.....	36
3.2 Material	36
3.3 Prozedur	39
3.3.1 Lernphase	39
3.3.2 Entscheidungsphase	40

3.4	Apparatur.....	42
3.5	Versuchsteilnehmer	42
3.6	Bayesianische Strategieklassifikation	43
4	Ergebnisse.....	49
4.1	Zeiten und Durchgänge der Lernphase	49
4.2	Zeiten der Entscheidungsphase	49
4.3	Bayesianische Strategieklassifikation	49
4.4	Blickwechsel zwischen Alternativen pro Sekunde	51
4.5	Blickwechsel in Abhängigkeit der Position des ersten differenzierenden Cues.....	53
4.6	Entscheidungszeiten	56
4.7	Anzahl der Fixationen vor einem Blickwechsel	57
4.8	Verweildauern vor einem Blickwechsel	59
5	Diskussion.....	61
5.1	Ziel der Untersuchung	61
5.2	Zusammenfassung und Interpretation der Ergebnisse	61
5.3	Forschungspraktische Konsequenzen und Ausblick	65
	Literatur.....	69
	Anhang.....	77
	Anhang A: Untersuchungsmaterial der Lernphase.....	77
	Anhang B: Untersuchungsmaterial der Entscheidungsphase	78
	Anhang C: Funktionsweise des Eye Trackers	79
	Anhang D: Areas of Interest	81

Zusammenfassung

Die Informationssuche und –integration konnten bei gedächtnisbasierten Multi-Attribut-Entscheidungen bislang nicht untersucht werden. Die vorliegende Arbeit beschäftigt sich mit der Entwicklung eines neuen prozessbasierten Verfahrens, das auf einer Analyse von Blickbewegungen beruht. Das Verfahren wurde in einer experimentellen Untersuchung erprobt. In einer vorangestellten Lernphase eigneten sich die Versuchsteilnehmer Eigenschaftsausprägungen (Cue-Ausprägungen) von sechs verschiedenen Alternativen an. Die Eigenschaftsausprägungen waren in Textform in rechteckigen Kästen um eine bildliche Darstellung einer Alternative angeordnet. In der darauf folgenden Entscheidungsphase führten die Teilnehmer gedächtnisbasierte Entscheidungen zwischen jeweils zwei der Alternativen durch. Die eine Hälfte der Versuchsteilnehmer bekam die nichtkompensatorische Take the best (TTB) Heuristik instruiert, die andere Hälfte die kompensatorische Dawes Rule (DR). Während der Entscheidungsphase wurden die Blickbewegungen der Versuchsteilnehmer aufgezeichnet. Es wurde angenommen, dass sich die Blickbewegungen zwischen den zwei Gruppen instruierter und nach ihrem Entscheidungsverhalten klassifizierter Versuchsteilnehmer hinsichtlich der strategieinduzierten Suchmuster unterscheiden würden. Diese Annahme beruht auf Befunden, die zeigen, dass Versuchsteilnehmer beim Informationsabruf an die Orte zurückblicken, an denen sie zuvor Informationen aufgenommen haben, selbst wenn diese Informationen dort nicht mehr stehen. Verwender der TTB Heuristik zeigten mehr Blickwechsel zwischen den Alternativen, je später der erste Cue differenzierte, und sie verweilten kürzer in einer Alternative, bevor sie den Blick auf die andere Alternative richteten. Bei DR Verwendern fanden sich erwartungsgemäß keine Unterschiede in den Blickwechselhäufigkeiten, je später der erste Cue differenzierte, und sie verweilten länger innerhalb einer Alternative vor einem Blickwechsel. Die Ergebnisse stützen die Idee, dass Blickbewegungen wertvolle Informationen über die kognitiven Prozesse während der Gedächtnissuche liefern können. Das Potenzial der Blickbewegungsmessung als Prozessmaß zur Untersuchung gedächtnisbasierter Entscheidungen wird diskutiert.

1 Einleitung

Die Forschung im Bereich der gedächtnisbasierten Multi-Attribut-Entscheidungen beschäftigt sich mit der Frage, welche kognitiven Prozesse bei Entscheidungen zwischen zwei und mehr Alternativen ablaufen und durch welche Faktoren diese Prozesse beeinflusst werden. Die bis dato identifizierten Prozesse oder Entscheidungsstrategien unterscheiden sich in der Anzahl und Art der Informationen, die für die Entscheidungsfindung ausgewählt und integriert werden. Beim Informationsabruf resultieren strategiebedingt unterschiedliche Suchmuster (Jungermann, Pfister & Fischer, 2005; Jungermann & Lütge, 2009).

Um den „gap between theory and data“ (Bröder, 2000b, S. 1332) schließen zu können, müssten die durch die verschiedenen Entscheidungsstrategien produzierten Suchmuster identifiziert werden. Es gibt bereits einige Verfahren, die die Informationssuche und -integration bei Vorliegen der Eigenschaftsausprägungen beobachten können (Bröder, 2005; Payne, Bettman & Johnson, 1993; Harte & Koele, 2001; Russo & Doshier, 1983). Die meisten alltäglichen Entscheidungen beinhalten allerdings neben der Informationssuche in der Umwelt auch gedächtnisbasierte Suchprozesse (Bröder & Schiffer, 2003b; Weber & Johnson, 2009). Für diesen Bereich sind Maße rar, die geeignet sind, den Entscheidungsprozess zu beobachten. (Renkewitz & Jahn, 2009). Ein Grund dafür ist, dass es bisher für unmöglich gehalten wurde, die Gedächtnissuche bei multi-attributen Entscheidungsaufgaben zu beobachten: „In any case, process-tracing techniques cannot be used in our paradigm in which information search takes place in memory, where a monitoring of the search processes is impossible.“ (Bröder & Schiffer, 2003b, S. 280). Befunde aus der Aufmerksamkeitforschung geben allerdings Anlass zur gegenteiligen Auffassung. Spivey und Richardsson (2001) konnten zeigen, dass Menschen an Orte zurückblicken, an denen zuvor Informationen gezeigt wurden, die aber zum Zeitpunkt des Informationsabrufs dort nicht mehr verfügbar sind.

Könnte man, ausgehend von diesen Befunden, die Informationssuche und -integration auch bei gedächtnisbasierten Multi-Attribut-Entscheidungen beobachten, würde das einen bedeutenden Fortschritt in der Untersuchung kognitiver Prozesse der Informationsverarbeitung bedeuten. Die vorliegende Untersuchung hat

zum Ziel, herauszufinden, ob die Gedächtnissuche bei unterschiedlichen Entscheidungsstrategien zu strategiebedingten Unterschieden in den Suchmustern führt. Dazu wurde eine multimethodale Vorgehensweise entwickelt, deren zentraler Bestandteil die Blickbewegungsmessung darstellt.

Der Forschungsstand zu gedächtnisbasierten Multi-Attribut-Entscheidungen und der Hintergrund der zu erprobenden Methode werden in Kapitel zwei dargestellt. Es werden zwei Klassen von kognitiven Prozessen vorgestellt und die Bedeutung gedächtnisbasierter Entscheidungen im Gegensatz zu Entscheidungen, bei denen die Informationen vorgegeben sind, herausgearbeitet. Der zweite Teil des zweiten Kapitels beschäftigt sich mit den momentan verfügbaren Methoden der Strategieklassifikation. Neben deren Stärken und Schwächen wird das Potenzial eines multimethodalen Vorgehens diskutiert, dessen wesentlicher Bestandteil die Neuanwendung der Eye Tracking Methode als Prozessmaß in einer gedächtnisbasierten Entscheidungsaufgabe darstellt. Der dritte Teil des zweiten Kapitels erläutert die theoretischen Grundlagen der neuen Methode und endet mit der Fragestellung und den experimentellen Anforderungen an das Untersuchungsdesign.

Im dritten Kapitel werden das verwendete Material und die eingesetzte Apparatur, sowie der Ablauf des Experiments beschrieben. Weiterhin wird die Logik der Bayesianischen Strategieklassifikation nach Bröder und Schiffer (2003a) erläutert, die als ein von Blickbewegungsmaßen unabhängiger Indikator Teil des multimethodalen Vorgehens ist. Kapitel vier stellt die Ergebnisse zu fünf Hypothesen dar, die Aufschluss geben sollen, ob das Verfolgen der Blickbewegungen als Prozessmaß bei gedächtnisbasierten Multi-Attribut-Entscheidungen die bisherigen Maße der Strategieklassifikation sinnvoll ergänzen kann. Kapitel fünf fasst die Ergebnisse zusammen und diskutiert deren Implikationen.

“Whether this matter can be resolved by inventing clever testing methods remains to be investigated.” (Newell & Bröder, 2008, S. 201)

2 Theoretischer Hintergrund

2.1 Gedächtnisbasierte Multi-Attribut-Entscheidungen

Stellen sie sich vor, sie unterhalten sich mit einem Freund oder einer Freundin über ihre beiden letzten Kinobesuche. Beim ersten Besuch sahen sie einen Dokumentarfilm („Unsere Erde“) und beim zweiten Besuch eine Komödie („Willkommen bei den Sch’tis“). Ihr Freund oder ihre Freundin fragen sie, welcher Film ihrer Meinung nach insgesamt mehr in die Kinokassen einspielen wird. Sie wissen, dass neben dem Film-Genre auch Merkmale wie die Menge erlangter Auszeichnungen und die Höhe des Filmbudgets auf den Erfolg eines Films hinweisen. In einer Kinozeitschrift haben sie gelesen, dass Komödien im Moment mehr einspielen als Dokumentarfilme. Sie haben auch erfahren, dass der Dokumentarfilm bereits zwei Auszeichnungen erhalten hat und ihm ein Budget von 30 Mio. Euro zur Verfügung stand. Die Komödie ist dagegen erst für eine Auszeichnung nominiert worden und musste mit 11 Mio. Budget auskommen. Für welche der beiden Alternativen entscheiden sie sich?

Solche und ähnliche Entscheidungssituationen treffen Menschen täglich an. Sie sind Gegenstand der Untersuchungen im Forschungsbereich der *multiple-cue probabilistic inferences*. Im Fokus des Interesses steht dabei die Frage, wie Menschen vorgehen, wenn sie sich zwischen mindestens zwei Alternativen entscheiden müssen, die durch mehrere Eigenschaften, die in einer statistischen Beziehung zur Zielvariablen stehen (Cues), und ihren Ausprägungen hinsichtlich dieser Eigenschaften beschrieben werden können. Die statistische Beziehung oder Prädiktionskraft der Cues wird durch ihre Validitäten ausgedrückt. Zum Beispiel ist das Film-Genre ein zuverlässigerer Prädiktor für den Kinoerfolg als die Anzahl der erhaltenen Auszeichnungen. Die Cues deuten mit ihren Validitäten auf das Kriterium hin. Zum Beispiel wird eine Komödie mit mehreren Auszeichnungen und einem hohen Budget am ehesten erfolgreich werden. Die Cue-Validitäten bestimmen gemeinsam mit den kritischen Cue-Ausprägungen die Entscheidungsrelevanz einer Alternative.

Gedächtnisbasiert sind Entscheidungen, wenn die Attributinformationen beim Fällen der Entscheidung nicht mehr verfügbar sind und aus dem Gedächtnis abgerufen werden müssen. Zum Beispiel fragt sie ihr Freund oder ihre Freundin, welchem Film sie den größeren Erfolg versprechen. In diesem Moment liegt das Kinoheft, in dem sie die Informationen gelesen haben, nicht mehr vor ihnen, und sie versuchen sich zu erinnern, welche Eigenschaftsausprägungen die Filme aufwiesen.

Zunächst wurde angenommen, dass Menschen bei Entscheidungen dieser Art unbegrenzt rational handeln (Laplace 1814/1986). Das heißt, der Entscheider integriert die Informationen nach mathematischen und wahrscheinlichkeitstheoretischen Modellen. Gewählt wird dann die Alternative mit dem höchsten Erwartungsnutzen (Edwards, 1954; von Neumann & Morgenstern, 1944), der höchsten Wahrscheinlichkeit für eine Alternative gegeben ein bestimmtes Datenmuster gemäß dem Bayes-Theorem (Baron, 2000; Dawes, 1988) oder dem höchsten Vorhersagewert als Ergebnis linearer Integration multipler Cues gemäß regressionsbasierter Verfahren (Fiedler, 1978; Hoffman, 1960). Solche Modelle werden allerdings, angewandt auf reale Umwelten, schnell zu komplex und ohne technische Hilfsmittel nicht lösbar (Gigerenzer & Goldstein, 1996). Der Mensch funktioniert nicht wie ein „intuitive statistician“ (Pachur, 2009) und handelt nicht nach den Grundpostulaten rationalen Denkens (Jungermann & Lütge, 2009). Vielmehr geht er nur „begrenzt rational“ vor (Simon, 1955; 1957). Das heißt, die kognitiven Ressourcen für die Informationsaufnahme, -verarbeitung und -speicherung sind begrenzt. Deshalb wird heute davon ausgegangen, dass der Mensch Informationen im Sinne eines „information processor“ (Newell & Bröder, 2008; Newell & Simon, 1972) verarbeitet, der intelligent mit seinen kognitiven Ressourcen umgeht.

Wie könnte diese in der kognitiven Psychologie weit verbreitete Annahme auf den Bereich der Multi-Attribut-Entscheidungen angewendet werden? Simon ging davon aus, dass Menschen Heuristiken (Daumenregeln) ¹ verwenden. Diese einfachen Verfahren zur Lösungsfindung führen nicht notwendigerweise zur besten

¹ Eine ausführliche Diskussion des Begriffs und die sich daraus ergebenden Implikationen für dessen Verständnis und Untersuchung finden sich bei Shah und Oppenheimer (2008).

Lösung (Bröder, 2005). Dennoch erlauben sie eine Restrukturierung und Vereinfachung komplexer Problemsituationen (Newell & Simon, 1972). Ausgehend von Simon wurde eine Vielzahl von Heuristiken postuliert und getestet. Zu den wichtigsten Beiträgen gehören das *heuristics and biases* Programm von Kahnemann, Slovic & Tversky (1982), der *adaptive decision maker* von Payne, Bettman und Johnson (1993) und die Heuristiken der *adaptiven Werkzeugkiste* von Gigerenzer, Todd und der ABC Research Group (1999).

2.1.1 Kompensatorische vs. nichtkompensatorische Strategien

Im Folgenden werden prototypische Vertreter zweier „Klassen“ (Bröder, 2005) von Heuristiken erklärt und für diese Untersuchung relevante Eigenschaften herausgearbeitet. Zum einen wird *Dawes Rule* (auch *Equal Weighted Adding* oder *EQW*) vorgestellt, die nach Robyn M. Dawes benannt wurde, der ihre Effizienz im Vergleich zu optimal gewichteten linearen Modellen wie der multiplen Regression zeigen konnte (Dawes, 1979; Dawes & Coorigan, 1974). Zum anderen wird *Take The Best* erläutert (eine lexikographische Strategie, daher auch *LEX*), die empirisch am häufigsten getestete Strategie (Bröder, 2005) der adaptiven Werkzeugkiste von Gigerenzer et al. (1999; siehe auch Gigerenzer, 1996; Gigerenzer, Hoffrage & Kleinböling, 1991). Weitere Entscheidungsstrategien finden sich bei Jungermann, Pfister und Fischer (2005).

Dawes Rule (DR) betrachtet alle Cue-Ausprägungen jeder der beiden Alternativen. Für jede Alternative wird die Anzahl kritischer Cue-Ausprägungen bestimmt. Die Entscheidung erfolgt zu Gunsten der Option mit der größeren Summe. Denken sie zurück an das Kinobeispiel zu Beginn dieses Absatzes: Für welchen Film würden sie sich entscheiden, wenn sie DR anwenden würden? Zur Veranschaulichung sind die Cues und ihre Ausprägungen in Tabelle 1 dargestellt. Wenn sie nach DR vorgehen, bestimmen sie für jede Alternative die Anzahl positiver Cue-Ausprägungen. Sie erinnern sich, dass eine Komödie, die bereits mehrfach ausgezeichnet wurde und ein hohes Budget zur Verfügung hatte, den größten Erfolg erzielen müsste. Alternative 1 stimmt mit diesem Kriterium nur im Hinblick auf den

ersten Cue Film-Genre überein. Alternative 2 besitzt dagegen zwei positive Cue-Ausprägungen: Der Dokumentarfilm wurde bereits ausgezeichnet und besaß ein Budget von 30 Mio. Euro. Sie entscheiden sich für Alternative 2, da diese die größere Summe positiver Cue-Ausprägungen besitzt.

DR ist ein Vertreter kompensatorischer Strategien. Das heißt, einzelne Cue-Informationen können sich gegenseitig ausgleichen. Weiterhin gehen alle verfügbaren Informationen in das Urteil ein (Bröder, 2005). Bei Payne, Bettman & Johnson (1993) wird diese Heuristik auch als Equal Weighted Adding Strategie bezeichnet. Der Grund dafür ist, dass DR nicht die Cue-Validitäten berücksichtigt. Das heißt, sie gewichtet die Attribute nicht gemäß der Prädiktionskraft einzelner Cues. Im Beispiel wird das Attribut Film-Genre genauso gewichtet, wie das Attribut Auszeichnung oder Budget. Weiterhin wird angenommen, dass DR die Informationen der Alternativen optionsweise verarbeitet. Das heißt, es werden erst alle Informationen einer Option integriert, bevor zu einer anderen Alternative gewechselt wird.

Tabelle 1. Kinofilme mit Cue-Ausprägungen und resultierendem Cue-Muster

Kinofilm	Cue 1: Film-Genre	Cue 2: Auszeichnung	Cue 3: Budget	Cue- Muster	TTB wählt	DR wählt
Kriterium	Komödie	ausgezeichnet	30 Mio	+++		
1	Komödie	nominiert	11 Mio	+-		
2	Dokumentarfilm	ausgezeichnet	30 Mio	+++	1	2

Dem gegenüber ordnet die Take The Best Heuristik (TTB) Informationen über Cue-Ausprägungen entlang ihrer Validitäten. Die Informationssuche beginnt beim validesten Cue, und die Optionen werden attributweise verglichen. Das heißt, die Ausprägung des validesten Cues der einen Alternative wird mit der entsprechenden Ausprägung der anderen Alternative verglichen (Russo & Doshier, 1983). Falls bezüglich der Ausprägungen keine Alternative günstiger als die andere

erscheint, wird der nächst-valide Cue betrachtet. Sobald ein Cue eine Differenzierung zwischen den Alternativen erlaubt, wird die Informationssuche und -integration abgebrochen und eine Entscheidung gefällt. Am Kinobeispiel verdeutlicht, würden sie die Alternativen zunächst im Hinblick auf das Film-Genre vergleichen, da dieses den validesten Cue darstellt. Alternative 1 war die Komödie, Alternative 2 der Dokumentarfilm. In diesem Beispiel differenziert bereits der erste Cue. Unabhängig von den Ausprägungen weiterer Attribute entscheiden sie sich für Alternative 1.

TTB gehört zu den nichtkompensatorischen Strategien, da sie die Attribute mit geringerer Validität als der differenzierende Cue „ignoriert“ (Gigerenzer & Todd, 1999). Für die Entscheidung werden also nicht alle Informationen herangezogen. Darin besteht auch der wichtigste Unterschied zu der zuvor eingeführten kompensatorischen DR Heuristik. Eine Ausnahme stellt der Fall dar, in dem erst der letzte Cue in der Hierarchie der Cue-Validitäten differenziert. Aufgrund der seriellen Berücksichtigung der Cue-Validitäten wird TTB auch als „Lexikographische Regel“ bezeichnet (Payne, Bettmann, & Johnson, 1993). Die Attribute werden gemäß ihrer Validitäten geordnet und dann entsprechend dieser Reihenfolge seriell integriert.

In dem Kinobeispiel kommt TTB zu einem anderen Ergebnis, als DR. Verschiedene Entscheidungsstrategien können wie in diesem Beispiel zu unterschiedlichen Ergebnissen führen (Bröder, 2005; Harte & Koele, 2001). Es drängt sich vielleicht die Frage auf, was die richtige oder optimale Entscheidung darstellt. Damit beschäftigen sich eigenständige Forschungsbereiche zum Entscheiden und Urteilen, weshalb an dieser Stelle nur auf weiterführende Literatur verwiesen wird: Einhorn und Hogarth (1981) setzen sich kritisch mit der Definition optimaler Entscheidungen auseinander. Jungermann und Lütge (2009) diskutieren die Rolle von Emotionen beim Entscheidungsprozess. Hastie und Dawes (2001) setzen sich mit der Problematik der Bestimmung rationalen Entscheidens auseinander. Eine allgemeine Einführung in den komplexen Gegenstandsbereich der Forschung zum Entscheiden und Urteilen findet sich bei Goldstein und Hogarth (1997). Einen aktuellen Überblick über die wichtigsten Diskussionspunkte bieten Weber & Johnson (2009). Die Forschung zu gedächtnisbasierten Multi-Attribut-

Entscheidungen beschäftigt sich mit den psychologischen Aspekten der Informationsintegration, welche in komplexen Alltagssituationen gezeigt werden (Jungermann & Lütge, 2009).

2.1.2 Empirische Befunde zu Multi-Attribut-Entscheidungen

In der Literatur finden sich sowohl Belege für den Einsatz kompensatorischer Strategien, die alle verfügbaren Informationen optionsweise verarbeiten als auch nichtkompensatorischer, selektiver und attributweiser Prozesse (Bröder 2000b, Lee & Cummins, 2004). Der Prozentsatz von TTB Verwendern lag in früheren Untersuchungen zwischen 20 und 72% (Bröder & Gaissmaier, 2007, Weber & Johnson, 2009). Dieser Wert schwankt stark, denn die Auswahl einer Strategie hängt von verschiedenen Faktoren ab und geschieht adaptiv an die Gegebenheiten der Umwelt (Bröder; 2005; Gigerenzer & Todd, 1999; Payne, Bettman & Johnson, 1993). Die empirisch am besten belegten Einflussfaktoren werden im Folgenden genannt.

Bröder (2000b) zeigte, dass hohe monetäre Kosten für den Informationsabruf einzelner Merkmalsausprägungen eher zum Einsatz nichtkompensatorischer Strategien führen. Christensen-Szalanski (1980) sowie Rieskamp und Hoffrage (2008) demonstrierten, dass Zeitdruck zum verstärkten Einsatz nichtkompensatorischer Strategien führt. Dem gegenüber zeigten Newell und Shanks (2003) und Newell, Weston & Shanks (2003), dass Personen mehr Informationen kaufen als für die Anwendung der TTB Heuristik erforderlich wäre und das sogar, wenn der zusätzliche Informationskauf keinen Nutzen bringt. Letzteres zeigten sie an einer fiktiven Börsenspielaufgabe mit nur zwei Cues. Dabei konnte der wenig valide Cue den valideren nicht kompensieren. Das hatte zur Folge, dass DR immer dieselbe Vorhersage wie TTB traf. Ein Kauf zusätzlicher Informationen erbrachte demnach keinerlei Nutzen. Im Gegenteil, es verursachte Kosten. Dennoch kauften 9 von 24 Versuchsteilnehmern zusätzliche Informationen. Offensichtlich bevorzugen es Entscheider, Urteile aufgrund vollständiger Informationen zu treffen (Bröder, 2005).

Dieses Ergebnis ist kritisch für die TTB Heuristik, die von einem Abbruch der Informationssuche ausgeht, sobald ein Cue zwischen zwei Optionen differenziert.

Befunde von Hausmann (2004) und Newell, Rakow, Weston und Shanks (2003) zeigten, dass Personen bei der Informationssuche nicht nur das Kriterium der Validität, sondern auch die Diskriminationsrate (Wahrscheinlichkeit, mit der ein Cue zwischen Objekten diskriminiert) oder den Success (Wahrscheinlichkeit einer korrekten Antwort mit nur einem Cue) berücksichtigen. TTB bezieht bei der Informationsintegration allerdings nur die Validität und nicht die Diskriminationsrate oder den Success ein.

2.1.3 Empirische Befunde zu gedächtnisbasierten Entscheidungen

Bei der Untersuchung multi-attributer Entscheidungssituationen blieb eine wesentliche Einflussgröße auf die Strategiewahl bisher unberücksichtigt: Einfache Heuristiken wie TTB werden vor allem dann ausgewählt, wenn die internen Kosten für die Informationsauswahl groß sind (Bröder & Newell, 2008). Das ist der Fall, wenn Informationen aus dem Langzeitgedächtnis abgerufen werden müssen (*inferences from memory*) und nicht vorgegeben sind (*inferences from givens*, Gigerenzer & Todd, 1999). Bröder bezeichnete die Annahme, dass das Verfügbarmachen der relevanten Attributinformationen aus dem Langzeitgedächtnis erhöhte Kosten verursacht und deshalb zu einem vermehrten Einsatz nichtkompensatorischer Strategien führe als *Gedächtnisabruf-Hypothese* (Bröder, 2005, S. 196). Zu ihrer Überprüfung entwickelte er ein eigenes Paradigma, das im Folgenden erläutert wird.

Die Experimente zur Gedächtnisabruf-Hypothese bestanden immer aus einer Lern- und einer Entscheidungsphase. In der Lernphase eigneten sich die Versuchsteilnehmer Wissen über fiktive Objekte an. Dieses sollte drei Bedingungen erfüllen: Es durfte nicht mit Vorwissen konfundiert sein, am Beginn der Entscheidungsphase sollten alle Versuchsteilnehmer die gleiche Wissensbasis besitzen, und die Informationen über die Cues mussten für eine Testung im experimentellen Setting in einer angemessenen Zeit erlernbar sein. Diese beschränkenden Bedingungen waren erforderlich, um überhaupt Schlussfolgerungen

über die gewählte Entscheidungsstrategie ziehen zu können. Die Entscheidungsrelevanz der Cue-Ausprägungen sollte den Versuchsteilnehmern erst am Beginn der Entscheidungsphase mitgeteilt werden, damit die Teilnehmer nicht schon während der Lernphase Urteile über die Alternativen fällten, denn das hätte die Anwendung verschiedener Entscheidungsstrategien überflüssig gemacht (Bröder, 2005).

Als inhaltliche Einkleidung wählte Bröder ein Kriminalspiel. Dabei sollte die Mörderin (Ex-Liebhaberin) eines berühmten Sängers gefunden werden. In der Lernphase eigneten sich die Versuchsteilnehmer die Ausprägungen von vier Cues an, die die Verdächtigen beschrieben. Nach der Lernphase wurde ihnen mitgeteilt, welche Indizien (Cues) am Tatort vorhanden waren. In der Entscheidungsphase bekamen die Teilnehmer je zwei Verdächtige präsentiert. Ihre Aufgabe war es zu entscheiden, wer mit größerer Wahrscheinlichkeit die Tat begangen hatte. In insgesamt fünf Experimenten fanden sich mehr TTB Verwender als Verwender kompensatorischer Strategien. Darin sah Bröder eine Bestätigung der Gedächtnisabruf-Hypothese: „Der Gedächtnisabruf der Attributinformationen scheint mit kognitiven Kosten verbunden zu sein und die Verwendung einfacher und vor allem sparsamer Heuristiken zu fördern.“ (Bröder, 2005, S. 229). Untersuchungen, die Aussagen über den Einsatz kompensatorischer vs. nichtkompensatorischer Strategien unter Bedingungen mit Informationsabrufkosten treffen wollen, sollten gedächtnisbasierte Entscheidungen beinhalten.

2.1.4 Zusammenfassung

Zusammenfassend lässt sich für die Untersuchungen zu gedächtnisbasierten Multi-Attribut-Entscheidungen festhalten: Die Forschung in diesem Bereich beschäftigt sich mit der Frage, wie Menschen vorgehen, wenn sie sich zwischen mindestens zwei Alternativen entscheiden müssen, die sich hinsichtlich der Ausprägungen auf mehreren Cues (Multi-Attribut) unterscheiden. Dabei besteht eine Möglichkeit in der Verwendung von Heuristiken. Je nach Anforderungen der Situation gehen Menschen manchmal kompensatorisch und optionsweise (wie DR)

und manchmal nichtkompensatorisch und attributweise (wie TTB) vor. Ein wichtiger Faktor für die Auswahl einer Heuristik besteht darin, ob die Informationen aus dem Gedächtnis abgerufen werden. Denn das führt zum verstärkten Einsatz nichtkompensatorischer Strategien und sollte bei Untersuchungen, die verschiedene Strategieklassen gegeneinander testen, berücksichtigt werden.

Welche Entscheidungsstrategien unter welchen Bedingungen zum Einsatz kommen, wird derzeit noch viel diskutiert (Bröder & Newell, 2008; Dougherty, Thomas & Francko-Watkins, 2008; Gigerenzer, Hoffrage & Goldstein, 2008; Newell & Bröder, 2008). Um die verschiedenen theoretischen Ansätze einer empirischen Überprüfung zugänglich zu machen, bedarf es ausgefeilter „diagnostischer Verfahren, die einen Rückschluss auf die von Vpn in einer experimentellen Situation verwendeten Strategien ermöglichen“ (Bröder, 2005, S. 84). Diese Herausforderung wird in der kognitiven Psychologie als „Theorie-Empirie-Überbrückungsproblem“ (Steyer & Eid, 2001) bezeichnet. Im Folgenden werden die bisher verwendeten Methoden zur Strategieklassifikation vorgestellt und diskutiert.

2.2 Methoden der deskriptiven Entscheidungsforschung

Ziel der Methoden der deskriptiven Entscheidungsforschung sollte es sein, „Rückschlüsse über die von Vpn verwendeten Entscheidungsstrategien zu ziehen, wenn multiple Attributinformationen zu berücksichtigen sind“ (Bröder, 2005, S. 85). Bisher wurde eine ganze Reihe von Verfahren entwickelt, die mehr oder weniger Klarheit über die verwendeten Entscheidungsstrategien geben. Von diesen sind nicht alle für die Untersuchung gedächtnisbasierter Multi-Attribut-Entscheidungen geeignet. Ziel dieser Arbeit soll es sein, das Methodenrepertoire sinnvoll mit einer Methode zur Untersuchung gedächtnisbasierter Entscheidungen zu ergänzen. Deshalb werden zunächst die bisher am häufigsten verwendeten Verfahren zur Untersuchung multi-attributer Entscheidungssituationen vorgestellt und auf ihre Stärken und Schwächen hin betrachtet.

Die Methoden der deskriptiven Entscheidungsforschung unterteilen sich in solche, die das Ergebnis der Entscheidungen, also den *Outcome*, fokussieren und

solche, die versuchen, Annahmen über die zu Grunde liegenden kognitiven Prozesse bei Verwendung verschiedener Strategien zu testen (Bröder, 2005). Letztere werden daher als *prozessbasierte Methoden* bezeichnet (Payne, Bettman & Johnson, 1993). Zu den Outcome basierten Methoden gehören das *Structural Modeling (SM)*, die *regressionsbasierte* sowie die *Bayesianische Strategieklassifikation*. Bekannte Vertreter des prozessbasierten Ansatzes sind die *Informationstafeln* mit ihren Weiterentwicklungen dem *Mouselab* und dem *Eye Tracking*, *Lautes Denken* sowie die *Entscheidungszeiten*.

2.2.1 Outcome basierte Methoden

2.2.1.1 Structural Modeling

Als Grundlage des SM dienen die Entscheidungen der Versuchsteilnehmer in den experimentell kontrollierten Entscheidungsdurchgängen. Ziel ist es, kompensatorische von nichtkompensatorischen Strategien zu unterscheiden. Dazu werden die Entscheidungen mit Hilfe mathematischer Modelle wie der multiplen linearen Regression² in Beziehung zu den Attributinformationen gesetzt. Die Attributwerte dienen als Prädiktoren und die Entscheidungen werden als Kriterium eingesetzt. Die Regressionsgewichte werden mit der Gewichtung einzelner Attribute im Entscheidungsprozess in Beziehung gesetzt. Der Multiple Korrelationskoeffizient R dient als Maß der Modellanpassung und der Konsistenz des Entscheiders. Ein hoher multipler Korrelationskoeffizient spricht für den Einsatz kompensatorischer Strategien.

Der Einsatz des SM in der Entscheidungsforschung ist nicht unproblematisch. Zum einen kann es einen perfekten linearen Fit und damit einen hohen Korrelationskoeffizienten auch bei nichtkompensatorischen Strategien geben. Richtet ein Versuchsteilnehmer zum Beispiel die Entscheidungen konsistent nur an einem Cue aus, so erhalten die anderen Regressionsgewichte einen Wert von Null. Dennoch ist der Model Fit gegeben. Zum anderen stellt Bröder fest, dass „regression weights

² Ebenso üblich sind varianzanalytische Verfahren beim Structural Modeling (Harte & Koele, 2001)

may not conform to our intuitions about importance of cues in a particular strategy“ (2000a, S. 652). Sowohl der Korrelationskoeffizient als auch die Regressionsgewichte stellen die Eigenschaften der Entscheidungsregeln nur unzureichend dar. Es ist daher schwierig aus dem Ergebnis des SM Rückschlüsse über die verwendete Strategie zu ziehen. Zudem versucht die Methode mit Hilfe zusammenfassender Maße aus den Entscheidungen induktiv Rückschlüsse über die zu Grunde liegenden Prozesse zu erhalten. Bröder (2000a; 2005) kritisiert dieses methodische Vorgehen daher als nicht valide.

2.2.1.2 Regressionsbasierte Strategieklassifikation

Bröder (2002; 2005) entwickelte eigene Verfahren zur Strategieklassifikation. Anders als das SM leiten diese zunächst formal die zu erwarteten Datenstrukturen verschiedener Heuristiken aus der Theorie her und prüfen anschließend, inwiefern diese zu den empirischen Entscheidungsmustern der Versuchsteilnehmer passen. Das Ziel besteht darin herauszufinden, welche Strategie das Entscheidungsmuster produziert hat.

In der regressionsbasierten Strategieklassifikation werden für verschiedene Strategien (TTB und DR) lineare Entscheidungsregeln formalisiert, die sich hinsichtlich der Struktur der Cue-Gewichtung unterscheiden. Spezifische Untersuchungsbedingungen ermöglichen es, aus den formalen Eigenschaften von TTB und DR eindeutige Vorhersagen über die Regressionsgewichte abzuleiten. Außerdem wird ein Fehlermodell spezifiziert, denn es wird angenommen, dass Versuchsteilnehmer Strategien nicht konsistent anwenden. Das heißt, die verwendeten Strategien wechseln im Verlauf der Entscheidungsdurchgänge. Das Fehlermodell nimmt eine konstante Fehlerwahrscheinlichkeit für alle Entscheidungsdurchgänge eines Experiments an. Im Anschluss prüfen statistische Tests der Varianzaufklärung einzelner Modelle (Regressionsgleichungen) spezifische Vorhersagen über den Einsatz verschiedener Strategien. Die Nullhypothese besagt, dass ein Versuchsteilnehmer TTB verwendet hat. Muss diese verworfen werden, wird davon ausgegangen, dass der Teilnehmer nach DR handelte. Das Problem

dieser Methode besteht in den bereits erwähnten spezifischen Untersuchungsbedingungen. Unter anderem (Bröder, 2002) ist die Klassifikationsmethode auf die Analyse binärer Wahlen beschränkt. Die Alternativen müssen mindestens durch vier Cues beschrieben und vollständig miteinander verglichen werden. Das heißt, die Versuchsteilnehmer müssen die Ausprägungen von $2^4 = 16$ Objekte lernen und 120^3 Items bearbeiten. Für die Teilnehmer sind diese Bedingungen in einer angemessenen Zeit kaum realisierbar.

2.2.1.3 Bayesianische Strategieklassifikation

Eine weitere von Bröder & Schiffer (2003a) entwickelte Methode stellt die Bayesianische Strategieklassifikation dar. Auch sie testet aus vorab formalisierten Modellen der Strategien vorhergesagte Entscheidungsmuster auf ihren Fit mit dem beobachteten Datenmuster. Dafür nutzt sie das Bayes-Theorem. Damit ist es möglich, die Wahrscheinlichkeit für eine Strategie zu bestimmen, gegeben ein bestimmtes Entscheidungsmuster. Diese sogenannte a-posteriori Wahrscheinlichkeit berechnet sich aus der a-priori Wahrscheinlichkeit der Strategien (Basisraten) und der Wahrscheinlichkeit für das Datenmuster gegeben eine bestimmte Strategie (Likelihoods)⁴.

Welche Strategien gegeneinander getestet werden, wird, wie bei der Regressionsbasierten Klassifikation, im Vorhinein festgelegt. Klassifiziert werden die nichtkompensatorische TTB Heuristik, zwei kompensatorische Strategien: DR und eine gewichtete lineare Entscheidungsregel (WADD-Weighted Additive Rule; Payne, Bettman & Johnson, 1993) und einfaches Raten. Das Entscheidungsmuster eines Teilnehmers wird derjenigen Strategie mit der höchsten Wahrscheinlichkeit zugeschrieben.

³

$$\left(\frac{16!}{2!(16-2)!} \right)$$

⁴ Eine detaillierte Beschreibung der Logik und der Berechnung der Bayesianischen Strategieklassifikation findet sich im Abschnitt 3.6.

Der Vorteil dieser Methode gegenüber der Regressionsbasierten Klassifikation liegt in weniger restriktiven Annahmen. So reichen bereits 10 Objekte und 45 Items⁵ für eine reliable Klassifikation aus (Bröder & Schiffer, 2003b). Das erlaubt einen breiteren Einsatz der Methode. Ein weiterer Vorteil besteht darin, dass unterschiedliche Modelle mit Hilfe des gleichen Optimierungskriteriums (den Likelihoods) an die Daten angepasst werden.

Bröder & Schiffer (2003a) prüften die Validität der Methode in Simulationsstudien. Sie konnten zeigen, dass bei einer niedrigen Fehlerrate unter 20% die Quote der richtigen Klassifikation bei über 90% lag. Die Bayesianische Strategieklassifikation stellt somit die brauchbarste im Moment zur Verfügung stehende Outcome basierte Klassifikationsmethode dar.

2.2.2 Prozessbasierte Methoden

Mit Hilfe der prozessbasierten Verfahren kann die Informationssuche beobachtet und analysiert werden. Von dem Suchverhalten wird auf die zu Grunde liegenden kognitiven Prozesse und damit die verwendeten Entscheidungsstrategien geschlossen (Riedl, Brandstätter & Roithmayr, 2008). Es wird also nicht nur versucht, eine Verbindung zwischen dem Input (Entscheidungsdurchgänge) und dem Output (Ergebnis der Entscheidungsdurchgänge) herzustellen, sondern das Ziel der prozessbasierten Methoden ist es „[to] reveal the train of thought leading to the final decision“ (Svenson, 1979). Die am häufigsten eingesetzten Methoden werden im Folgenden beschrieben.

2.2.2.1 Informationstafeln

Informationstafeln stellen multi-attribute Informationen in Matrixform dar. In den Spalten stehen die Attribute, in den Zeilen die Alternativen und in den Feldern

⁵

$$\left(\frac{10!}{2!(10-2)!} \right)$$

die Cue-Ausprägungen. Die Felder der Matrix sind zu Beginn des Versuchs verdeckt. Die Cue-Ausprägungen befinden sich entweder auf der Rückseite einer Karte oder in einem Briefumschlag und müssen durch den Versuchsteilnehmer aufgedeckt werden. Der Versuchsleiter zeichnet die Reihenfolge auf, in der die Informationen erhoben werden. Daraus lassen sich Rückschlüsse auf die Strategien der Informationssuche und –integration ziehen (Harte & Koele, 2001; Jungermann, Pfister & Fischer, 2005). Unterschieden werden optionsweises von attributweisem Vorgehen (Riedl, Brandstätter & Roithmayr, 2008; Russo & Doscher, 1983; Lohse & Johnson, 1996; Payne, Bettman & Johnson, 1993; Svenson, 1979).

Bei der optionsweisen Suche werden zunächst die Informationen einer Alternative gesammelt und integriert, bevor zu einer anderen Alternative gewechselt wird. Bei der attributweisen Verarbeitung werden zunächst die Alternativen hinsichtlich eines Attributs verglichen, bevor die Ausprägungen eines anderen Attributs betrachtet werden.

Nachteile der Methode der Informationstafeln sind, dass der Versuchsleiter zeitnah das Verhalten des Teilnehmers beobachten muss. Für den Teilnehmer ist der Abruf der Informationen auf Kärtchen oder Briefumschlägen vergleichsweise aufwändig und langsam, denn er kann nur eine Information nach der anderen „entpacken“. Das kann einen Einfluss auf die Auswahl der Verarbeitungsstrategie haben (Renkewitz & Jahn, 2009). Es ist deshalb nicht auszuschließen, dass der eigentliche Prozess der Informationsintegration nicht beobachtet wird. Deshalb werden die Informationstafeln in dieser Form nicht mehr als Methode der deskriptiven Entscheidungsforschung eingesetzt.

2.2.2.2 Mouselab und Eye Tracking

Payne et al. (1993) entwickelten die Methode der Informationstafeln weiter. Die Informationsmatrix wird dabei mit verdeckten Zellen auf einem Bildschirm dargeboten. Die Versuchsteilnehmer führen die Computermaus über die Zellen der Matrix. Wird eine Zelle durch Anklicken ausgewählt, erscheint die Attributinformation. Deshalb wird die Methode auch als Mouselab bezeichnet. Diese

Technik ist bedeutend schneller und einfacher in der Ausführung für die Teilnehmer als die bisher verwendeten Informationstafeln. Weiterhin entfällt die kontinuierliche Beobachtung durch den Versuchsleiter, da der Computer die Suchreihenfolge aufzeichnet.

Russo & Doshier (1983) verwendeten zum ersten Mal Eye Tracking, um die Blickbewegungen der Versuchsteilnehmer am Bildschirm aufzuzeichnen, während diese eine Entscheidungsaufgabe bearbeiteten. Das Verfolgen der Blickbewegungen verursacht für die Probanden die wenigsten kognitiven Kosten für den Informationsabruf und spart am meisten Zeit im Vergleich zum Aufdecken von Kärtchen oder dem Anklicken von Feldern mit der Maus.

Bezüglich der resultierenden Suchmuster zeigen sich selbst zwischen Mouselab und Eye Tracking noch Unterschiede, die durch die Untersuchungsmethode bedingt sind (Lohse & Johnson, 1996). Bei Mouselab wird mehr systematische Suche beobachtet als bei Eye Tracking Untersuchungen. Systematische Suche wird mit der Betrachtung aller Cue-Informationen und kompensatorischer Vorgehensweisen in Verbindung gebracht.

Sowohl in Mouselab als auch in Eye Tracking Untersuchungen werden die Suchmuster der Versuchsteilnehmer aufgezeichnet. Daraus lassen sich Indizes berechnen, welche die Prozesse der Informationssuche und -integration beschreiben (Harte & Koele, 2001; Payne, Bettman & Johnson, 1993; Riedl, Brandstätter & Roithmayr, 2008). Beispielhaft seien hier die Häufigkeit der Betrachtung eines Feldes, die intraindividuelle Variabilität der Suchmuster und die Gesamtdauer der Betrachtung eines Feldes genannt.

Aus der Vielfalt wird ersichtlich, dass es noch keine Einigkeit über die am besten geeigneten Indizes gibt. Diese Problematik sei an dem von Payne (1976) erstellten *Suchindex* skizziert. Der Index berechnet sich aus der Subtraktion der Blickbewegungen zwischen den Alternativen von den Blickbewegungen innerhalb der Alternativen, dividiert durch die Summe der Blickbewegungen zwischen und innerhalb der Alternativen. Ein positiver Wert steht für ein optionsweises Suchmuster und ein negativer Wert für eine attributweise Suche.

Dieser Index weist einige Probleme auf (Harte & Koele, 2001). Jede Abweichung von Null wird entweder als attributweises oder optionsweises Vorgehen interpretiert. Eine mögliche Interpretation für einen Indexwert von Null fehlt. Weiterhin kann der logische Schluss nur in eine Richtung gezogen werden: Bei Verwendung einer Strategie ist ein bestimmtes Suchmuster zu erwarten. Umgekehrt ergibt sich ein ungültiger konditionaler Schluss (*affirmation of the consequent*), denn ein Suchmuster könnte auch andere Ursachen haben. Die kritische Auseinandersetzung mit dem Suchindex soll verdeutlichen, dass trotz der inhaltlichen Plausibilität (Lohse & Johnson, 1996), Indizes präzise aus den theoretischen Annahmen über die Informationssuche der Strategien abgeleitet werden müssen.

2.2.2.3 Lautes Denken

Während Informationstafeln und Eye Tracking Methoden nur die Informationssuche beobachten können, misst die Methode des lauten Denkens (auch als *verbale Protokolle* bezeichnet) sowohl die Informationssuche als auch die zu Grunde liegenden Prozesse direkt (Lohse & Johnson, 1996). Bei dieser Methode werden die Versuchsteilnehmer aufgefordert, ihre Gedanken verbal wiederzugeben, während sie die Entscheidungsaufgabe bearbeiten. Das Gesagte wird aufgezeichnet und im Nachhinein nach einem Kodierungsschema transkribiert (Montgomery & Svenson, 1989). Zum Beispiel können die Häufigkeiten, mit denen bestimmte Attribute genannt werden, über verschiedene Untersuchungsbedingungen hinweg verglichen werden (Svenson & Karlsson, 1986).

Durch direktes Nachfragen ist es möglich, eine genaue Beschreibung des Entscheidungsprozesses zu erhalten. Dennoch muss berücksichtigt werden, dass Menschen nur begrenzt in der Lage sind, die eigenen kognitiven Prozesse introspektiv zu beschreiben (Nisbett & Wilson, 1977). Diese Beschränkung äußert sich darin, dass nicht alle relevanten Informationen durch den Versuchsteilnehmer genannt werden oder das Gesagte nicht dem zu Grunde liegenden kognitiven Prozess entspricht (Russo, Johnson & Stephens, 1989). Letzteres ist zum Beispiel bei

holistischer und paralleler Verarbeitung der Cues zu erwarten (Bröder & Schiffer, 2003b; Jahn, Renkewitz & Kunze, 2007). Außerdem kann lautes Denken eine reaktive Messmethode sein (Lohse & Johnson, 1996). Das heißt, es beeinflusst möglicherweise die Art und Weise der Informationssuche und –verarbeitung. Nicht zuletzt sind die Erfassung des Gesagten und die Transkription extrem zeitaufwändig. Die genannten Punkte stellen den fruchtbaren Einsatz der Methode des lauten Denkens für die deskriptive Entscheidungsforschung in Frage.

2.2.2.4 Entscheidungszeiten

Die Entscheidungszeiten entsprechen den Reaktionszeiten der Versuchsteilnehmer in den Entscheidungsdurchgängen. Diese variieren mit den verwendeten Strategien. Bergert und Nosofsky (2007) zeigten, dass die Entscheidungszeiten mit der Position des ersten diskriminierenden Cues variieren. Dies untersuchten sie an einer Aufgabe, bei der die Cue-Informationen nicht gelernt werden mussten (inferences from givens). Bröder & Gaissmaier (2007), reanalysierten die Untersuchungen zur Gedächtnisabruf-Hypothese und konnten damit auch für inferences from memory zeigen, dass die Entscheidungszeiten der Versuchsteilnehmer zunehmen, je später der erste diskriminierende Cue erscheint.

Die Entscheidungszeiten können einfach über die Reaktionszeiten in den Entscheidungsdurchgängen erhoben werden. Eine Reihe weiterer Untersuchungen setzten sie aufgrund ihrer Genauigkeit und Einfachheit ergänzend zu den Ergebnissen Outcome basierter Klassifikationsmethoden ein (Glöckner, 2006; Glöckner & Betsch, 2008; Jahn, Renkewitz & Kunze, 2007; Renkewitz & Jahn, 2009).

2.2.3 Diskussion der Methoden

Im Folgenden werden Vor- und Nachteile der Outcome und der prozessbasierten Verfahren methodenübergreifend erläutert. Darauf aufbauend wird eine neue Herangehensweise vorgeschlagen.

Outcome basierte Verfahren untersuchen nicht direkt den Entscheidungsprozess (Ford et al., 1989; Payne, Bettman & Johnson, 1993; Svenson, 1979). Dieser steht allerdings im Fokus des Interesses der Untersuchung gedächtnisbasierter Multi-Attribut-Entscheidungen. An sich stellt diese Tatsache keinen Nachteil dar, vorausgesetzt, die Methode ist in der Lage, die verwendeten Entscheidungsstrategien reliabel und valide zu bestimmen. Darin liegt allerdings das Problem der meisten Outcome basierten Verfahren. Sowohl kompensatorische als auch nichtkompensatorische Entscheidungsstrategien können identische Entscheidungsmuster erzeugen. Zum Beispiel entscheiden sich TTB und DR bei einem Paarvergleich mit den Cue-Mustern „++-“ für Alternative 1 und „+--“ für Alternative 2 beide für Alternative 1. Eine Strategieklassifikation, die auf die Entscheidungsmuster aufbaut, wäre in diesem Fall nicht in der Lage, zwischen den Strategien zu differenzieren (Harte & Koele, 2001; Martignon & Hoffrage, 2002; Riedl, Brandstätter & Roithmayr, 2008). Bröder (2002; 2003a) fand eine Lösung für dieses Problem. Er gestaltete die Paarvergleiche in den binären Entscheidungsdurchgängen derart, dass für einen Teil der Items die Strategien unterschiedliche Vorhersagen machen. Bei einem Paarvergleich mit den Cue-Mustern „+--“ für Alternative 1 und „-++“ für Alternative 2, würde sich TTB zum Beispiel für Alternative 1 und DR für Alternative 2 entscheiden. Liegen genug Items vor, die eine Differenzierung der Strategien erlauben, können die von einer Entscheidungsregel vorhergesagten Entscheidungsmuster auf ihren Fit mit den beobachteten Entscheidungsmustern getestet werden. Die Bayesianische Strategieklassifikation baut auf diesen Itemtypen auf. Damit stellt sie ein brauchbares Outcome basiertes Messinstrument dar.

Prozessbasierte Methoden versuchen, die Lücke zwischen dem Input und dem Output zu schließen, indem sie den Entscheidungsprozess beschreiben. Das Spektrum der prozessbasierten Verfahren ist groß. An der Methode der Informationstafeln und den darauf aufbauenden Methoden Mouselab und Eye Tracking wird allerdings bemängelt, dass die gezogenen Schlussfolgerungen eher allgemeiner Natur sind und keine tieferen Einsichten in den zu Grunde liegenden Entscheidungsprozess ermöglichen (Bröder, 2000; Harte & Koele, 2001). Die

zusammenfassenden Indizes könnten demnach maximal Klassen von Entscheidungsstrategien identifizieren (Riedl, Brandstätter & Roithmayr, 2008).

Werden Prozessmaße als einzige Klassifikationsinstrumente eingesetzt, unterläuft ihnen ein logischer Fehler, denn es ist falsch von einem bestimmten Entscheidungsmuster auf eine verwendete Strategie zu schließen (Bröder, 2000a). Ein weiteres Problem besteht darin, dass die prozessbasierten Methoden bisher nur an inferences from givens getestet wurden. Der Grund dafür ist die Annahme, dass der Informationsabruf bei gedächtnisbasierten Entscheidungen nicht beobachtet werden kann (Bröder & Schiffer, 2003b). Um aber die einfachen Heuristiken wie TTB zu testen, die erst bei hohen kognitiven Kosten zum Einsatz kommen, müssen die Informationen aus dem Gedächtnis abgerufen werden.

Bisher hat sich nur ein Ansatz zur Untersuchung gedächtnisbasierter Entscheidungen als fruchtbar erwiesen: Entscheidungszeiten liefern Verhaltensdaten im Laufe des Entscheidungsprozesses. Neben den Entscheidungszeiten lohnt sich die Weiterentwicklung der prozessbasierten Verfahren. Zum einen können diese die Outcome basierten Maße mit direkten Beobachtungsdaten über die Attributverarbeitung ergänzen. Dadurch liefern sie einen Beitrag zur Interpretation der Ergebnisse der Outcome basierten Klassifikation und der Entscheidungszeiten. Die Frage, ob die gesammelten Informationen tatsächlich auch integriert werden, ist zwar noch nicht geklärt (Bröder 2005; Svenson, 1979). Dennoch kann davon ausgegangen werden, dass die resultierenden Entscheidungsmuster durch die Suchreihenfolgen der Informationen beeinflusst werden (Renkewitz & Jahn, 2009).

Beide methodischen Ansätze, Outcome und prozessbasierte Methoden, besitzen neben Schwächen in der praktischen Umsetzung auch große Potenziale. Harte und Koele kommen daher zu dem Schluss: „it is wise not to rely exclusively on one method for studying human behavior in general, or multiattribute evaluation processes in particular“ (2001, S. 44). Ein multimethodales Vorgehen wurde schon eher diskutiert (Einhorn, Kleinmütz & Kleinmütz, 1979; Payne, Braunstein & Carrol, 1978; Svenson, 1979). Bröder (2005) beklagt jedoch einen Mangel an konvergenter Validität zwischen Outcome und prozessbasierten Verfahren. Die Kritik bezieht sich vor allem auf den Befund, dass beim SM überwiegend kompensatorische und bei

prozessbasierten Verfahren überwiegend nichtkompensatorische Strategien klassifiziert werden. Schwachpunkte des SM und des Prozessmaße wurden bereits erläutert. Dennoch sollte eine multimethodale Vorgehensweise, die Outcome und prozessbasierte Methoden ergänzend einsetzt, nicht grundlegend ausgeschlossen werden. Riedl, Brandstätter und Roithmayr (2008) entwarfen zum Beispiel einen Algorithmus, der drei Prozess Indizes mit einem Outcome basierten Maß verknüpfte, und klassifizierten damit erfolgreich 13 verschiedene Strategien an einer Entscheidungsaufgabe, bei der die Informationen vorgegeben wurden.

2.2.4 Anforderungen an ein multimethodales Vorgehen

Für eine neue methodische Herangehensweise schlagen Renkewitz und Jahn (2009) deshalb folgendes Vorgehen vor: Als bereits bewährtes Maß und unabhängiger Indikator wird die Outcome basierte Strategieklassifikation nach Brüder & Schiffer (2003a) eingesetzt. Ein neues auf Eye Tracking beruhendes prozessbasiertes Verfahren soll zusätzliche Informationen über die Attributverarbeitung liefern und die Strategieklassifikation ergänzen. Dieses Verfahren ist ähnlich den bisherigen Untersuchungen zum Eye Tracking. Neu ist die Annahme, dass damit auch gedächtnisbasierte Entscheidungen untersucht werden können. Die Idee wird durch Befunde gestützt, die zeigen, dass Versuchsteilnehmer unter bestimmten experimentellen Bedingungen beim Informationsabruf an die Orte zurückblicken, an denen sie die Informationen zuvor aufgenommen haben (Richardson & Spivey, 2000). Der nächste Abschnitt wird sich deshalb ausführlich mit diesen Befunden und dem Potenzial für die deskriptive Entscheidungsforschung auseinandersetzen.

Die Eye Tracking Methode bietet sich auch deshalb als Prozessmaß an, da sie die geringsten kognitiven Kosten für den Informationsabruf verursacht im Vergleich zum Mouselab oder den Informationstafeln. Außerdem sind die heutigen Verfahren der Blickbewegungsmessung weniger intrusiv als diejenigen der ersten Eye Tracking Untersuchungen (Just & Carpenter, 1976). Aus den Blickdaten werden aufgrund theoretischer Überlegungen zum Informationsabruf bei verschiedenen

Entscheidungsstrategien Indizes abgeleitet. Diese werden zusätzlich mit den Entscheidungszeiten als einem bewährten Maß prozessbasierter Klassifikation verglichen.

2.3 Blickbewegungsmessung als Prozessmaß

Bisher wurde angenommen, dass prozessbasierte Verfahren im Allgemeinen und Eye Tracking im Speziellen nur bei Entscheidungsexperimenten from givens einsetzbar wären. Neue Befunde der Aufmerksamkeitsforschung (Ferreira, Apel & Henderson, 2008; Richardson & Spivey, 2000; Spivey, Richardson & Fitneva, 2004) legen allerdings nahe, dass die bloße Beobachtung der Blickbewegungen Rückschlüsse auf die Informationssuche im Gedächtnis erlaubt. Das Kinobeispiel vom Beginn soll zur Veranschaulichung des Prinzips dienen. Stellen Sie sich vor, ihr Freund oder ihre Freundin hält nun das Kinoheft in der Hand, in dem sie zuvor die Beschreibungen der beiden gesehenen Filme gelesen haben. Während sie gefragt werden, welcher Film mehr einspielen wird, sehen sie mit größerer Wahrscheinlichkeit zu dem Kinoheft, das nun ihr Freund oder ihre Freundin in der Hand halten als zu einem anderen Kinoheft, das sich in einem Regal neben ihnen befindet. Das heißt, obwohl die Informationen nicht mehr sichtbar sind (denn ihr Bekannter/ ihre Bekannte hält das Kinoheft in der Hand), sehen sie zurück an den Ort, an dem sie die Information aufgenommen haben.

Übertragen auf die Forschung zu gedächtnisbasierten Multi-Attribut-Entscheidungen heißt das: Wenn die Blicke zu geleerten Informationstafeln zurückkehren, auf denen zuvor Informationen dargeboten wurden, die aber zum Zeitpunkt des Informationsabrufs nicht mehr da stehen, dann sind die Blickbewegungen diagnostisch für die Informationssuche im Gedächtnis und Eye Tracking liefert ähnliche Daten für gedächtnisbasierte Entscheidungen wie für inferences from givens.

2.3.1 Looking at nothing-Phänomen

Richardson und Spivey (2000) zeigten die Tendenz an Orte zu sehen, an denen Informationen zuvor aufgenommen wurden. Dazu präsentierten sie den Versuchsteilnehmern einen Bildschirm, der in vier Felder unterteilt war. In jedem Feld erschien nacheinander der Kopf einer sprechenden Person. Jede der vier Personen nannte einen Fakt. Die vier Fakten pro Durchgang sollten sich die Teilnehmer einprägen. Unmittelbar danach hörten sie einen Satz der sich inhaltlich auf einen der zuvor gesagten Sätze bezog. Die Versuchsteilnehmer mussten bewerten, ob dieser Satz richtig oder falsch war. Dazu war es erforderlich, das zuvor Gehörte aus dem Gedächtnis abzurufen. Während dieser Verifikationsphase sahen die Teilnehmer auf eine leere Matrix. In dieser Zeit wurden die Blickbewegungen der Teilnehmer aufgezeichnet. Abbildung 1 zeigt beispielhaft einen Durchgang aus dem Experiment von Richardson und Spivey (2000).

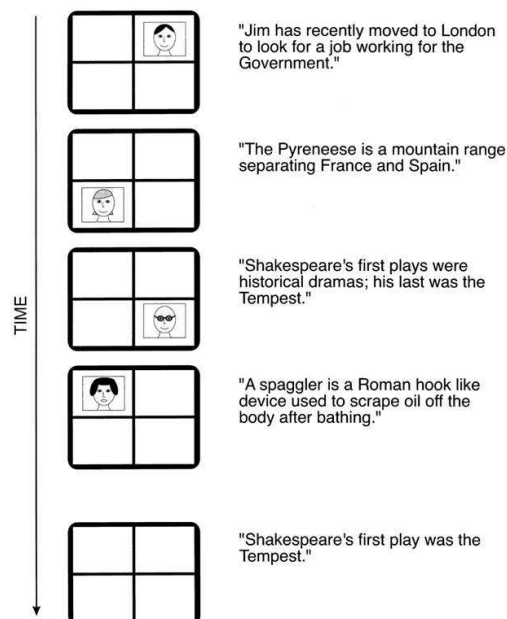


Abbildung 1. Schematische Ansicht eines Durchgangs aus Experiment 1 in Richardson & Spivey (2000). In diesem Beispiel befindet sich das kritische Feld rechts unten.

Die Blickbewegungen wurden als gemittelte Anzahl der Fixationen in eines der vier Felder operationalisiert. Jeder Teilnehmer durchlief acht Durchgänge. Eine ANOVA für Messwiederholungen über den Vergleich des kritischen Feldes, das die zu verifizierende Information enthielt, mit den anderen drei Feldern, ergab signifikant mehr Fixationen pro Durchgang in das kritische Feld (0.73 vs. im Uhrzeigersinn ausgehend vom kritischen Feld 0.44, 0.40, 0.47). Dieser Befund zeigt, dass Versuchsteilnehmer beim Informationsabruf an Orte sehen, die zuvor mit semantischen auditiven Informationen verknüpft wurden.

Spivey und Geng (2001) replizierten diesen Befund an einer visuellen Aufgabe. Dazu präsentierten sie den Versuchsteilnehmern eine 3x3 Matrix, in deren Ecken vier Symbole zu sehen waren. Diese unterschieden sich hinsichtlich der Farbe, der Form und der Orientierung. Zunächst waren alle vier Symbole zu sehen. Anschließend leerte sich der Bildschirm. Beim erneuten Einblenden fehlte eines der Symbole. Die Versuchsteilnehmer waren aufgefordert, entweder die Farbe oder die Orientierung des ausgeblendeten Symbols zu benennen. Dabei wurden ihre Blickbewegungen aufgezeichnet. Die Versuchsteilnehmer zeigten, wie bei Richardson und Spivey (2000) mehr Blicke in das leere Feld als in die anderen Felder (30-50% vs. 7% der Durchgänge). Das Zeigen von Blicken an Stellen, die Informationen enthielten, die aber zum Zeitpunkt des Abrufs nicht mehr vorhanden sind, wird als *Looking at nothing-Phänomen* bezeichnet (Ferreira, Apel & Henderson, 2008; Renkewitz & Jahn, 2009).

Aus den bisherigen Befunden schlussfolgerten die Autoren, dass gemeinsam mit den Fakten okulomotorische Informationen über den Präsentationsort enkodiert werden. Diese werden auch als räumliche Marker bezeichnet (Ballard, Hayhoe, Pook & Rao, 1997; Pylyshyn, 1989). Bei den bisher berichteten Experimenten bleibt allerdings unklar, ob die okulomotorischen Informationen bei der Etablierung der räumlichen Repräsentation eine Funktion übernehmen. Entweder dienen sie der Anreicherung der Gedächtnisrepräsentation mit zusätzlichen Informationen oder die Frage am Beginn der Verifikationsphase führt tatsächlich nur zur Reaktivierung eines okulomotorischen Musters, das während der Enkodierung der Informationen aufgezeichnet wurde und beim Abruf mit ausgeführt wird. Ersteres würde einen

Blick zu dem kritischen Feld auslösen, unabhängig von dessen Ort auf dem Bildschirm. Letzteres würde nur dann einen Blick in das kritische Feld auslösen, wenn die Information während der Enkodierung auch dort präsentiert wurde. In einer Folgeuntersuchung testeten Richardson & Spivey (2000) diese Vorhersage. Dazu führten sie zwei Bedingungen ein. Die erste Bedingung war identisch mit dem ersten bereits geschilderten Experiment. Der einzige Unterschied bestand darin, dass die Felder, in denen keine Informationen präsentiert wurden, maskiert waren. In einer zweiten Bedingung wurden die Felder während der auditiven Präsentation der Fakten systematisch in die Mitte des Bildschirms verschoben, so dass sich zwar die semantischen, nicht aber die okulomotorischen Informationen änderten. Bei der Verifikation blickten die Teilnehmer wieder auf die leere Matrix. Die Ergebnisse zeigen, dass in beiden Bedingungen signifikant mehr Fixationen in das kritische Feld, als in die anderen Felder gezeigt wurden (0.49 vs. 0.34 in Bedingung 1 und 0.43 vs. 0.25 in Bedingung 2). Diesem Ergebnis nach kehren Versuchsteilnehmer beim Erinnern nicht nur an den Ort zurück, an dem sie die Informationen aufgenommen haben (Bedingung 1), sondern auch zu der neuen Position des Feldes (Bedingung 2). Demnach dient die okulomotorische Information der Anreicherung der Gedächtnisrepräsentation.

Hoover & Richardson (2008) untersuchten dieselbe Fragestellung anhand einer anderen Aufgabe. Auf einem Bildschirm wurden zwei Erdhügel gezeigt. Diese befanden sich in zwei Feldern einer 2x2 Matrix. Zu Beginn eines experimentellen Durchgangs kroch ein weißer Hase aus einem der Erdhügel, der andere Hügel diente als Kontrollbedingung. Während das Häschchen zu sehen war, wurde auditiv ein Fakt dargeboten. In Bedingung 1 grub sich derselbe Hase ausgehend von dem ersten Erdhügel zu einem zweiten, der sich ebenfalls in einem der vier Felder der Matrix befand. In Bedingung 2 bewegte sich nach der auditiven Präsentation und dem Auftauchen des einen Hasen ein anderer, identischer Hase vom Rand des Bildschirms aus zu einem Erdhügel in der Matrix. In beiden Bedingungen wurde nach dem erneuten Erscheinen des Hasen (in Bedingung 1 derselbe, in Bedingung 2 ein identischer, aber nicht derselbe Hase) ein Satz eingespielt, der wiederum verifiziert werden musste. Während der Verifikationsphase wurden, wie in den vorhergehenden Experimenten, die Blickbewegungen aufgezeichnet. Die gemittelten

Fixationsdauern zeigten, dass die Versuchsteilnehmer in Bedingung 1 sowohl den ersten als auch den zweiten Erdhügel fixierten. Sie betrachteten demnach alle Orte, die mit dem ersten Hasen assoziiert wurden. Währenddessen wird in Bedingung 2 nur der erste Erdhügel betrachtet, da in dieser Bedingung nur der erste Erdhügel mit dem Fakt und dem ersten Hasen verbunden war. Dieser Befund replizierte die Ergebnisse von Richardson und Spivey (2000). Es wird nicht assoziativ das gleiche Muster betrachtet, sondern nur die Orte erhalten Aufmerksamkeit, die mit dem Gedächtnisinhalt verknüpft wurden.

Ebenso wie die eben beschriebenen Untersuchungen zur Speicherung semantischer auditiver Informationen zeigen Untersuchungen zum Lernen von Objekten (Hollingworth, 2006), gelesenen Inhalten (Altmann, 2004) oder räumlichen Anordnungen visueller Szenen (Johansson, Holsanova, Helmqvist, 2006), dass okulomotorische Informationen über räumliche Positionen mit den zu lernenden Inhalten enkodiert werden. Versuche, Ereignisse, Objekte, Wörter oder Szenen abzurufen, führen zu automatischen Blickbewegungen an die gespeicherten räumlichen Positionen.

Eine Erklärung für das Looking at nothing-Phänomen besteht in einer Externalisierung von Wissen (Ballard, Hayhoe, Pook & Rao, 1997; Richardson & Spivey, 2000; Spivey, Richardson & Fitneva, 2004). Das heißt, Wissen wird nicht nur im Gedächtnis, sondern auch in der Umwelt gespeichert: „people sometimes do not store particular information about the visual world, but instead access it with an eye movement when it is required“ (Richardson, Altmann, Spivey & Hoover, 2009, S.235). Das Looking at nothing-Phänomen tritt als Konsequenz der Aktivierung eines räumlichen Markers auf, der mit der abzurufenden Information assoziiert ist (Altmann, 2004).

Ferreira, Apel und Henderson (2008) gehen von einer integrierten Gedächtnisrepräsentation für Objekte und Ereignisse (episodische Spur) aus. Wird ein Teil der Informationen abgerufen, wird die visuelle Gedächtnisrepräsentation, welche die räumlichen Marker enthält, mit aktiviert. Diese veranlasst die Blickbewegungen zum Ort der Informationsaufnahme. Neu an der Erklärung von Ferreira et al. (2008) ist die Annahme, dass zusätzliche visuelle und linguistische

Details gespeichert werden. Ein externer Speicher wird nur unter bestimmten Umständen als Strategie eingesetzt, um die Arbeitsgedächtnisbelastung zu verringern. Das Looking at nothing-Phänomen ist ihrer Meinung nach nicht nur die Konsequenz, sondern auch die Ursache des Gedächtnisabrufs: „once the look takes place, the location then serves as an additional retrieval cue that further facilitates access of additional aspects of the original trace“ (Ferreira, Apel & Henderson, 2008, S.409).

Welche Rolle dem externen Speicher letztlich zukommt und welche zusätzlichen Informationen gespeichert werden, bleibt Gegenstand weiterer Forschung (Richardson, Altmann, Spivey & Hoover, 2009). Entscheidend für die Entwicklung der prozessbasierten Methode der Strategieklassifikation ist der Befund, dass Blickbewegungen den Gedächtnisabruf induzieren, wenn zuvor visuelle Merkmale der Situation mit den zu lernenden Informationen assoziiert wurden (Renkewitz & Jahn, 2009).

2.4 Fragestellung und experimentelle Umsetzung

Ziel dieser Untersuchung ist die Entwicklung einer Methode, welche Entscheidungsstrategien in einer gedächtnisbasierten Multi-Attribut Entscheidungsaufgabe gemäß des von Renkewitz und Jahn (2009) beschriebenen Vorgehens erfasst. Überprüft wird, ob die bloße Verhaltensbeobachtung mittels Eye Tracking die zu Grunde liegenden Prozesse der Entscheidungsfindung beschreiben kann. Dazu wurden die Befunde zum Looking at nothing-Phänomen und der gedächtnisbasierten Multi-Attribut-Entscheidungen in einer Untersuchung integriert.

Die experimentelle Aufgabe wurde in Anlehnung an das Lernparadigma von Bröder & Schiffer (2003b) gestaltet. Um das Looking at nothing-Phänomen für die Untersuchung gedächtnisbasierter Entscheidungen nutzen zu können, mussten Verbindungen zwischen den Orten der Informationsrepräsentation und den Cue-Informationen geschaffen werden. In der Lernphase standen die Attributausprägungen immer an derselben Position innerhalb eines räumlichen Objekts. In der Entscheidungsphase wurden dann zwei dieser Objekte ohne Cue-

Informationen präsentiert. Um sich für eine der beiden Alternativen entscheiden zu können, mussten die Versuchsteilnehmer die zuvor gelernten Informationen aus dem Gedächtnis abrufen. Währenddessen wurden die Blickbewegungen der Teilnehmer aufgezeichnet.

Gemäß den Befunden zum Looking at nothing-Phänomen sollte sich der Informationsabruf aus dem Gedächtnis während einer gedächtnisbasierten Entscheidungsaufgabe in den Blickbewegungen wiederfinden lassen. Dazu wurde untersucht, ob sich die Suchmuster kompensatorischer und nichtkompensatorischer Strategien unterscheiden. Als Vertreter der kompensatorischen Strategien wurde die DR Heuristik und als Vertreter der nichtkompensatorischen Strategien die TTB Heuristik getestet. Für eine erste Untersuchung der Fragestellung wurden die Strategien instruiert. Die Strategieklassifikation überprüfte anschließend, ob die Teilnehmer die instruierten Strategien auch anwendeten.

Zur Analyse der Blickdaten wurden fünf Hypothesen formuliert, die das Suchmuster der Teilnehmer auf verwendete Strategien beziehen. Die in den Hypothesen formulierten Indizes überprüften die Annahme, dass optionsweise Suche mit kompensatorischen Strategien in Verbindung steht und attributweise Informationssuche beim Einsatz nichtkompensatorische Strategien zum Einsatz kommt (Payne, Bettman & Johnson, 1993; Riedl, Brandstätter & Roithmayr, 2008).

Die Liste der Indizes könnte um alle von Payne et al. (1993) verwendeten Indizes erweitert werden. Für einen ersten Test der Blickbewegungen als Diagnostikum für die Informationssuche in einer gedächtnisbasierten Entscheidungsdomäne sollen allerdings die hier vorgestellten Indizes genügen. Außerdem muss vorweg angenommen werden, dass die Blickbewegungen nicht nur durch die Informationssuche beeinflusst sind, sondern auch von der Qualität der Blickdaten oder der Konzentrationsfähigkeit der Teilnehmer abhängen. Dennoch wird davon ausgegangen, dass das Verfolgen der Blickbewegungen mittels Eye Tracking einen Beitrag zu den im Moment zur Verfügung stehenden Methoden zur Untersuchung des Einsatzes kognitiver Strategien in Entscheidungssituationen aus dem Gedächtnis liefern kann.

Im Folgenden werden die fünf Hypothesen erläutert. Vier beziehen sich auf verschiedene Blickindizes und eine auf die Entscheidungszeiten.

Die **erste Hypothese** nimmt an, dass TTB Verwender mehr Blickwechsel zwischen den Alternativen zeigen als DR Verwender. Dieser Hypothese liegt die Annahme zu Grunde, dass TTB Verwender einzelne Attribute miteinander vergleichen, während DR Verwender die Attributinformationen erst innerhalb einer Alternative integrieren, bevor sie die Alternativen vergleichen. Deshalb sollten DR Verwender weniger Blickwechsel zwischen den Alternativen zeigen. Payne et al. (1993) überprüften diese Vorhersage für Entscheidungen from givens. Da TTB Verwender die Suche nach dem ersten differenzierenden Cue abbrechen, sollte die Anzahl der Blickwechsel für eine bestimmte Zeiteinheit betrachtet werden. Differenziert erst der letzte Cue, so sollten sich die Blickwechselfrequenzen der TTB Verwender nicht von denen der DR Verwender unterscheiden. Deshalb werden die Anzahl der Blickwechsel relativ zur Dauer eines Entscheidungsdurchgangs betrachtet.

Der Anstieg der absoluten Blickwechselfrequenzen für TTB Verwender wird in der **zweiten Hypothese** gesondert betrachtet. Während angenommen wird, dass die Anzahl der Blickwechsel für TTB Verwender größer wird, je später in der Validitätenreihenfolge erstmals ein Cue differenziert, sollte diese dagegen für die DR Verwender konstant bleiben, da letztere unabhängig von der Übereinstimmung von Cue-Werten Informationen auf Alternativebene integrieren.

Bröder und Gaissmaier (2007) sowie Jahn, Renkewitz und Kunze (2007) zeigten einen Anstieg in den Reaktionszeiten der TTB Verwender, je später der erste Cue differenziert. Mit der **dritten Hypothese** wird überprüft, ob sich dieses Ergebnis auch in den Reaktionszeiten der vorliegenden Untersuchung wiederfinden lässt. Sollten diese wie in bisherigen Untersuchung ansteigen und die Ergebnisse der Entscheidungszeiten mit denen der Blickwechselfrequenzen korrespondieren, würde dies die Validität des Index *Anzahl der Blickwechsel in Abhängigkeit vom ersten differenzierenden Cue* erhöhen.

Hypothesen 4 und 5 beziehen sich auf zwei weitere Indizes, die aus den Blickbewegungen der Versuchsteilnehmer ermittelt werden können. Die **vierte Hypothese** prüft, ob DR Verwender mehr Fixationen innerhalb einer Alternative zeigen, bevor sie den Blick auf die andere Alternative richten als TTB Verwender. Es wird angenommen, dass ein mehr an Fixationen mit mehr Informationssuche einhergeht. Integrieren DR Verwender erst die Informationen einer Alternative, bevor sie den Blick wechseln, müssten sie gemäß dem Looking at nothing-Phänomen mehr Blicke innerhalb einer Alternative zeigen als TTB Verwender, die attributweise vorgehen.

Mit der Anzahl der zu integrierenden Informationen sollte auch die Zeit zunehmen, die vergeht, bevor ein Blickwechsel stattfindet. DR integriert alle verfügbaren Informationen, während TTB die Suche nach dem ersten Cue, der eine Entscheidung zulässt, abbricht. Die **fünfte Hypothese** prüft daher, ob die Verweildauern innerhalb einer Alternative für DR Verwender länger sind als für TTB Verwender. Diese Vorhersage wurde in einer ähnlichen Form bereits bei Payne et al. (1993) überprüft. Sie gestalteten eine Entscheidungsaufgabe derart, dass die Versuchsteilnehmer möglichst genau vorgehen sollten. Neben der Anzahl gesammelter Informationen pro Alternative stieg auch die BOXTIME. Das ist die absolute Zeitdauer, die auf die Informationssuche in den Zellen verwandt wurde. Da die fünfte Hypothese zunächst die Dauern auf Alternativebene und nicht auf Ebene einzelner Informationen betrachtet, wird die Verweildauer innerhalb einer Alternative vor einem Blickwechsel zur anderen Alternative analysiert.

3 Methode

3.1 Design

Zur Überprüfung der Fragestellung wurde ein experimentelles Vorgehen gewählt, das aus einer Lernphase und einer Entscheidungsphase bestand. In der Lernphase eigneten sich die Versuchsteilnehmer Eigenschaftsausprägungen von sechs Objekten an. In der anschließenden Entscheidungsphase wurden pro Durchgang jeweils zwei von sechs Objekten ohne die Eigenschaftsausprägungen nebeneinander gezeigt. Die Aufgabe der Teilnehmer bestand darin, sich in 15 aufeinanderfolgenden Durchgängen und zwei vorangestellten Übungsdurchgängen für eine von jeweils zwei präsentierten Alternativen zu entscheiden.

Analysiert wurden das Blickverhalten, die Entscheidungszeiten und die Entscheidungsmuster (abhängige Variablen) unter Berücksichtigung der zuvor instruierten Strategie (unabhängige Variable). Eine Hälfte der Versuchsteilnehmer bekam eine Instruktion, die der nichtkompensatorischen TTB Heuristik entsprach. Die andere Hälfte wurde instruiert, sich nach der kompensatorischen DR Heuristik zu verhalten (between-subjects design). Demnach lag der vorliegenden Untersuchung ein einfaktorielles Design mit dem Faktor „Strategie“ und den Ausprägungen „nach DR“ oder „nach TTB instruiert“ zugrunde.

3.2 Material

Die zu lernenden Stimuli entsprachen sechs fiktiven Pilzen (a-f), die sich hinsichtlich der Cue-Ausprägungen auf vier Cues unterschieden. Die vier Cues waren das dominierende Enzym, die dominierende Aminosäure, der dominierende Mineralstoff und die Verbreitung. Jeder Cue konnte in drei möglichen Ausprägungen vorkommen. Zum Beispiel lag das Enzym in den Ausprägungen: Tyrosinase, Lakkase und Peroxydase vor. Die sechs Pilze und ihre Cue-Ausprägungen sind in Tabelle 2 dargestellt.

Tabelle 2. Pilze mit Cue-Ausprägungen und resultierendem Cue-Muster

Pilz	Cue 1: Enzym	Cue 2: Aminosäure	Cue 3: Mineralstoff	Cue 4: Verbreitung	Cue- Muster
Typischer Giftpilz	Tyrosinase	Ergosterin	Magnesium	selten	++++
a	Tyrosinase	Canavanin	Magnesium	selten	+---
b	Lakkase	Ergosterin	Magnesium	selten	---+
c	Peroxydase	Ergosterin	Magnesium	häufig	---+
d	Tyrosinase	Thiamin	Zink	mittel	+++
e	Tyrosinase	Canavanin	Kalium	selten	+++
f	Tyrosinase	Ergosterin	Zink	mittel	+++

Das Material war wie von Bröder (2005) zur Untersuchung gedächtnisbasierter Entscheidungen gefordert fiktiv gestaltet. Das erzeugte bei allen Teilnehmern die gleiche Wissensbasis und verminderte den Einfluss von Vorwissen. Bezüglich der Menge der Objekte musste ein Kompromiss eingegangen werden. In den bisherigen Untersuchungen, die die Strategieklassifikation nach der von Bröder und Schiffer (2003a) entwickelten Methode einsetzten, wurden 10 alternative Objekte gelernt (Bröder & Gaissmaier, 2007; Kunze, Renkewitz & Jahn, 2007).

Das Erlernen der Objekte erfordert eine aufwändige Lernphase. Die große Anzahl der zu lernenden Cue-Ausprägungen (zum Beispiel 40 bei Jahn, Renkewitz & Kunze, 2007) führt dazu, dass diese in mehreren Wiederholungen erlernt werden müssen. Dem gegenüber werden bei Untersuchungen zum Looking at nothing-Phänomen viel weniger Informationen gelernt (zum Beispiel die vier Fakten in der Untersuchung von Richardson & Spivey, 2000). Die Verifikationsphase (analog zur Entscheidungsphase der vorliegenden Untersuchung) erfolgt unmittelbar nach der Informationspräsentation. Zudem erscheint dort jede Information an einer anderen Stelle auf dem Bildschirm. Das heißt, jede Information wird mit einer anderen räumlichen Position verknüpft. Dies ist bei einer großen Anzahl von Cue-Ausprägungen nicht mehr möglich. Hier werden mehrere Informationen an derselben Bildschirmposition präsentiert. Diese Unterschiede könnten zu einer Abschwächung

des Looking at nothing-Phänomens führen. Um dem Entgegenzuwirken wurden in der vorliegenden Untersuchung anstelle der bisher verwendeten 10 nur 6 Objekte gestaltet. Aus den 6 Objekten, die jeweils durch 4 Cue-Ausprägungen beschrieben wurden, resultierten 24 zu lernende Informationen (Tabelle 2). Die Anzahl von 6 Objekten wurde nicht unterschritten, um eine ausreichende Anzahl diskriminierender Items für die Strategieklassifikation erzeugen zu können.

In der Lernphase wurde den Versuchsteilnehmern jede Alternative einzeln präsentiert. Abbildung A1 im Anhang zeigt eins von sechs Objekten wie sie in der Lernphase gezeigt wurden. In der Mitte des Bildschirms sahen die Versuchsteilnehmer das Foto eines Pilzes. Um das Foto herum waren vier rechteckige Kästen gruppiert. Diese enthielten die Cue-Ausprägungen in verbaler Form. Alle Alternativen waren gemäß dieser Anordnung gestaltet. Das heißt, die Ausprägungen eines Cues befanden sich für alle Alternativen im selben Kasten. Zum Beispiel stand die Information über das Enzym immer im unteren linken Kasten. Durch dieses Präsentationsformat sollte sichergestellt werden, dass die Versuchsteilnehmer jeden Cue mit einer bestimmten Stelle auf dem Bildschirm verknüpften.

In der Entscheidungsphase wurden dann jeweils zwei Alternativen nebeneinander präsentiert (Abbildung B2). Der wichtigste Unterschied zur Lernphase bestand darin, dass in der Entscheidungsphase nur noch die leeren Kästen zu sehen waren. Da die Eigenschaftsausprägungen nicht mehr angezeigt wurden, mussten diese nun von den Versuchsteilnehmern aus dem Gedächtnis abgerufen werden. Die Fotos der Pilze in der Mitte der Kästen dienten dazu, auch ohne sichtbare Cue-Ausprägungen zwischen den Alternativen differenzieren zu können. Zur Vermeidung von Reihenfolgeneffekten, wurden die Paarvergleiche in zwei unterschiedlichen Sequenzen präsentiert. Diese testeten die Items in gegensätzlicher Richtung. Zum Beispiel stand der Paarvergleich b-a in Reihenfolge 1 am Anfang der Entscheidungsphase und in Reihenfolge 2 am Ende. Die beiden Sequenzen wurden vorab so balanciert, dass in jeder Bedingung (TTB, DR instruiert) jede der beiden Reihenfolgen (1, 2) gleich häufig vorkam.

3.3 Prozedur

Die Blickbewegungen der Versuchsteilnehmer wurden nur in der Entscheidungsphase aufgezeichnet. Dennoch kalibrierten die Versuchsleiter den Eye Tracker bereits zu Beginn des Versuchs. Dadurch sollte das Zeitintervall zwischen Lern- und Entscheidungsphase verkürzt werden. Je länger es dauerte, bis die Versuchsteilnehmer das zuvor erworbene Wissen wieder abrufen konnten, desto größer war die Gefahr des Wissensverlusts (Baddeley, 1997). Die durchschnittliche Kalibrationsdauer betrug fünf Minuten. Danach schloss sich die Lernphase an. Nachdem das notwendige Lernziel erreicht wurde, begann die Entscheidungsphase, zu deren Beginn der Eye Tracker erneut kalibriert wurde. Der genaue Ablauf der Lern- und Entscheidungsphasen wird im Folgenden erläutert.

3.3.1 Lernphase

Die Lernphase begann mit der Abfrage demografischer Fakten (Alter, Geschlecht, Studienfach). Danach wurde den Versuchsteilnehmern mitgeteilt, dass ihnen sechs Pilze präsentiert würden, die sich hinsichtlich der Cue-Ausprägungen in vier Merkmalen unterschieden. In der Lernphase sollten sich die Teilnehmer insgesamt 24 Informationen (6 Pilze x 4 Ausprägungen) einprägen. Da alle Eigenschaften für die anschließenden Entscheidungsdurchgänge relevant waren, musste in der Lernphase ein 90% Kriterium erreicht werden. Erst wenn 22 von 24 Cue-Ausprägungen im abschließenden Gedächtnistest einmal korrekt wiedergegeben wurden, war die Lernphase beendet. Die Versuchsteilnehmer wurden darauf hingewiesen, dass die Erreichung des Lernkriteriums ihre aktive Mitarbeit erforderte.

Das Erlernen der Cue-Informationen teilte sich zunächst in zwei Blöcke auf. In jedem Block lernten die Versuchsteilnehmer 3 der 6 Pilze. Der erste Dreierblock begann mit der Präsentation eines Pilzes. Die dazugehörigen Cue-Ausprägungen standen in den Kästen, die um das Foto des Pilzes gruppiert waren. Die Teilnehmer wurden aufgefordert, sich die Informationen in den Kästen einzuprägen. Mittels Tastendruck bestimmten sie selbst, wann der erste Gedächtnistest für diesen Pilz beginnen sollte. Der Gedächtnistest bestand aus einer Multiple-Choice Abfrage

entlang der Cue-Validitäten. Dabei sahen die Versuchsteilnehmer nur noch das Foto des Pilzes mit den leeren Kästen. Die Cue-Ausprägungen waren ausgeblendet. Die valideste Eigenschaft wurde zuerst abgefragt. Drei mögliche Antwortalternativen (z.B. die drei unterschiedlichen Enzyme) standen, über die Durchgänge hinweg randomisiert in einer Zeile unter dem Objekt. Die Versuchsteilnehmer wählten mittels Tastendruck eine der 3 präsentierten Cue-Ausprägungen. Die für die Eingabe der Antworten benötigten Buchstabentasten waren gelb markiert. Die valideste Eigenschaft erschien immer im unteren linken Kasten. Alle weiteren Merkmale kamen im Uhrzeigersinn hinzu. Die Versuchsteilnehmer übten solange die Cue-Ausprägungen eines Pilzes ein, bis dieser von ihnen einmal vollständig und richtig zusammengesetzt wurde. In Abbildung A2 im Anhang ist eine Multiple-Choice Abfragesituation aus der Lernphase dargestellt. Über Kopfhörer erhielten die Teilnehmer zusätzlich ein Richtig- oder Falschsignal als akustisches Feedback. Die vervollständigten Cue-Ausprägungen blieben solange ausgeblendet, bis per Tastendruck zum Lernen des nächsten Pilzes fortgeschritten wurde.

Nachdem die Versuchsteilnehmer die ersten drei Pilze auf diese Weise einzeln gelernt hatten, erfolgte eine Gesamtabfrage der drei Pilze des ersten Blocks. Die Dreierblockabfrage entsprach im Format der Einzelabfrage. Der Test dauerte solange, bis die Versuchspersonen 11 von 12 Ausprägungen korrekt wiedergeben konnten (90% Kriterium). Das Lernen der restlichen drei Pilze erfolgte analog zur Vorgehensweise beim Erlernen der ersten 3 Pilze. Danach wurden alle sechs Pilze in einem abschließenden Gedächtnistest so lange abgefragt, bis 22 der 24 Ausprägungen korrekt wiedergegeben wurden. Die Zuordnung der Pilze zu den zwei Dreierblöcken sowie die Reihenfolge der Abfrage innerhalb der Blöcke waren randomisiert.

3.3.2 Entscheidungsphase

Die Teilnehmer wurden zufällig einer der beiden Instruktionsbedingungen und einer der beiden Reihenfolgen zugewiesen. Zu Beginn der Entscheidungsphase wurde der Eye Tracker erneut kalibriert. Danach lernten die Teilnehmer die Cue-

Validitäten mit den kritischen Cue-Ausprägungen. Diese wurden in Form einer Kurzbeschreibung auf dem Bildschirm präsentiert. Das Enzym Tyrosinase gab den verlässlichsten Hinweis auf die Giftigkeit eines Pilzes. Den zweitverlässlichsten Hinweis gab die Aminosäure Ergosterin. Weniger valide war die Information über den Mineralstoff Magnesium, und der Verbreitungsgrad mit der Ausprägung selten stellte den Cue mit der geringsten Validität dar. Zusätzlich wurden die kritischen Cue-Ausprägungen in Form einer Liste eingeblendet. Abbildung B1 zeigt die Instruktion der Cue-Validitäten und der kritischen Ausprägungen. Die Versuchsteilnehmer wurden zudem aufgefordert, die als „TEMS“ (Tyrosinase, Ergosterin, Magnesium und Selten) bezeichnete Liste kritischer Ausprägungen dem Versuchsleiter zu nennen. Dabei blieb der Bildschirm leer.

Die nächste Instruktion beschrieb die anzuwendende Entscheidungsstrategie und war experimentell variiert. Die eine Hälfte der Versuchsteilnehmer wurde gemäß der nichtkompensatorischen TTB Heuristik aufgefordert, sich aufgrund der ersten Cue-Ausprägung zu entscheiden, die einen Pilz als wahrscheinlich giftiger auswies. Alle übrigen Eigenschaftsausprägungen sollten ignoriert werden. Die andere Hälfte bekam die kompensatorische DR instruiert. Die Versuchsteilnehmer sollten zunächst für jede der beiden Alternativen die Anzahl der kritischen Eigenschaftsausprägungen bestimmen und danach diejenige mit der höheren Anzahl an Übereinstimmungen auswählen. Besaßen beide Pilze die gleiche Anzahl positiver Cue-Ausprägungen, sollten die Versuchsteilnehmer raten.

Nach der Instruktion der Entscheidungsstrategien begannen die Entscheidungsdurchgänge. Die Teilnehmer wurden aufgefordert, sich zu entscheiden, welcher von zwei eingeblendeten Pilzen im Hinblick auf das Kriterium giftiger war. In welchem Ausmaß die Versuchsteilnehmer die instruierte Strategie tatsächlich anwendeten, wurde im Nachhinein mit Hilfe der Strategieklassifikation (Bröder & Schiffer, 2003a) berechnet. Eine Beschreibung dieser Methode erfolgt im Abschnitt 3.6. Die Versuchsteilnehmer antworteten auf die 15 Entscheidungen und zwei Übungsdurchgänge mittels Eingabe der Zahl 1, für den linken Pilz und Eingabe von 2 für den rechten Pilz.

3.4 Apparatur

Die Blickbewegungen wurden mit Hilfe des berührungsfreien Eye Trackers iView X™ RED P/T der Firma SMI (SensoMotoric Instruments, Teltow/Berlin) aufgezeichnet. Die Kamera stand an der rechten Seite des Bildschirms. Der Abstand zwischen Auge und Kamera betrug ca. 80 cm. Um diesen Abstand konstant zu halten und Kopfbewegungen, die zu einer Verschlechterung der Tracking Qualität führen konnten, zu vermeiden, wurde in der Entscheidungsphase eine Kinnstütze benutzt. Die Blickbewegungen konnte der Versuchsleiter auf einem zweiten Bildschirm kontinuierlich verfolgen.

Die 9 aufeinanderfolgend gezeigten Kalibrationspunkte und die experimentellen Stimuli wurden auf einem 19“ Zoll LCD Monitor mit einer Auflösung von 800 x 600 Pixeln präsentiert. Sowohl die Objektpräsentation als auch die Datenaufzeichnung erfolgten mittels E-Prime 1.1. Die Aufzeichnung der Blickbewegungen durch den Eye Tracker wurden über Steuersignale von E-Prime gestartet und beendet. Zur Synchronisation der Verhaltensdaten mit den Blickdaten sendete E-Prime „Event Codes“ zum Eye Tracker. Das heißt, zu Beginn eines jeden Durchgangs wurde in das vom Eye Tracker erstellte Datenfile eingetragen, welchen der Durchgänge der Versuchsteilnehmer gerade bearbeitete. Eine Erläuterung zur Funktionsweise des Eye Trackers findet sich in Anhang C. Das akustische Feedback während der Lernphase wurde mittels Kopfhörer präsentiert. Die Eingaben tätigten die Versuchsteilnehmer auf einer handelsüblichen Tastatur. Die Analyse der gewonnenen Daten erfolgte mit Hilfe der Auswertungssoftware BeGaze 1.2 sowie dem Statistikpaket SPSS 14.0. Die Teststärken wurden mit Hilfe der frei verfügbaren Software G-Power 3.0.5 berechnet.

3.5 Versuchsteilnehmer

Am Experiment nahmen 32 Studierende der Technischen Universität Chemnitz teil. Die Teilnehmer erhielten entweder Versuchspersonenstunden oder partizipierten freiwillig. Um eine ausreichende Qualität der Blickdaten im Voraus zu gewährleisten, wurden nur Personen ohne Brille oder Kontaktlinsen und ohne

Schminke im Augenbereich getestet. 2 Studierende mussten von der Datenanalyse ausgeschlossen werden. Ursachen dafür waren zum einen, dass ein Teilnehmer das erforderliche Lernkriterium nicht erreichte. Bei diesem wurde der Versuch nach 60 min abgebrochen. Bei einem weiteren Teilnehmer konnten die Blickbewegungen nicht in ausreichend guter Qualität aufgezeichnet werden. Gründe dafür waren technische Schwierigkeiten des Eye Trackers bei der Erfassung der Korneareflexion aufgrund der Färbung der Iris. Die Ausfälle blieben insgesamt in einem vertretbaren Rahmen. Von 30 Versuchsteilnehmern lag ein vollständiger Datensatz zur weiteren Analyse bereit. Von den Teilnehmern waren 24 weiblich und 6 männlich. Das Durchschnittsalter betrug 24,2 Jahre ($SD = 5,9$). Die Stichprobe setzte sich, bis auf 2 Studierende der Medienkommunikation, aus Studierenden der Psychologie zusammen.

3.6 Bayesianische Strategieklassifikation

Als Grundlage für die Analyse der Blickdaten dient die Strategieklassifikation nach der von Bröder und Schiffer (2003) entwickelten Methode. Deshalb wird deren Berechnung im Folgenden genauer beschrieben.

Ziel dieser Klassifikationsmethode ist es, jeder Person diejenige Entscheidungsstrategie (TTB, DR, Raten) zuzuordnen, die für ihr Entscheidungsmuster am wahrscheinlichsten ist. Anders gesagt: Gegeben ein bestimmtes Datenmuster, welche Strategie hat dieses Datenmuster am wahrscheinlichsten erzeugt? Diese Fragestellung kann mit Hilfe des Bayes-Theorems (Gleichung 1) gelöst werden.

$$\frac{p(M/D)}{p(\bar{M}/D)} = \frac{p(M)}{p(\bar{M})} \times \frac{p(D/M)}{p(D/\bar{M})} \quad (1)$$

Auf der linken Seite steht der Quotient der a-posteriori Wahrscheinlichkeiten, das heißt der Wahrscheinlichkeiten für die zu testenden Modelle M und \bar{M} gegeben ein bestimmtes Datenmuster D . Die Modelle entsprechen den zu testenden

Entscheidungsstrategien. Die a-posteriori Wahrscheinlichkeit kann berechnet werden aus dem Verhältnis der a-priori Wahrscheinlichkeiten (Basisraten) der Modelle multipliziert mit dem Verhältnis der Likelihoods, das heißt der Wahrscheinlichkeiten für das Datenmuster gegeben ein bestimmtes Modell (bedingte Wahrscheinlichkeit).

Des Weiteren müssen drei grundlegende Annahmen getroffen werden. Erstens ist es nicht möglich, ein vollständiges Set möglicher Strategien zu spezifizieren. Das heißt, es wird nicht explizit berücksichtigt, dass Teilnehmer andere als die zuvor in Betracht gezogenen Strategien (DR, TTB und Raten) verwendet haben könnten oder die Strategie im Verlauf der Experimentalphase gewechselt haben könnten. Ein Beispiel für diesen Fall wäre ein Versuchsteilnehmer, der im ersten Teil der Experimentalphase nach TTB und im zweiten Teil nach DR entschied. Das Problem unvollständig spezifizierter Vorannahmen über das Verhalten von Menschen findet sich bei jeder Methode, die versucht, Modelle gegeneinander zu testen. Die Aussagen bleiben immer auf die zuvor in Betracht gezogenen Möglichkeiten begrenzt (Robert & Pashler, 2000, zit. nach Bröder & Schiffer, 2003a). Demnach sind auch die Ergebnisse relativ zu bewerten und stellen nur Annäherungen an das reale Geschehen dar (Riefer & Batchelder, 1988, zit. nach Bröder & Schiffer, 2003a). Modellfreie Analysemethoden für diesen Bereich gibt es noch nicht.

Zweitens kann nicht von einem fehlerfreien Antwortverhalten der Versuchsteilnehmer ausgegangen werden. Deshalb wird eine uniforme Verteilung der Fehlerwahrscheinlichkeit angenommen. Allerdings stellt ein konstanter Fehlerterm über alle Durchgänge hinweg ebenfalls eine Vereinfachung dar. Eine Lösung könnte ein empirischer Vergleich der Konsistenz der Entscheidungen bei verschiedenen Itemtypen über verschiedene Situationen hinweg bieten. Die Notwendigkeit dafür stellen die Autoren allerdings in Frage. Ein solcher Vergleich wird daher im Rahmen dieser Klassifikation nicht durchgeführt. Drittens sind zur Berechnung der a-posteriori-Wahrscheinlichkeiten Annahmen über die Höhe der Basisraten erforderlich. Aufgrund mangelnden Vorwissens über die a-priori-Wahrscheinlichkeiten wird auch hier eine uniforme Verteilung angenommen.

Das Set möglicher Strategien wird also im Vorhinein festgelegt. In dieser Untersuchung werden die Strategien TTB, DR und eine weitere „Strategie“ Raten gegeneinander getestet. Dass Versuchsteilnehmer nur geraten haben, ist in solchen Fällen plausibel, in denen sie sich nicht mehr an die zuvor gelernten Cue-Ausprägungen erinnern konnten. Weiterhin werden uniforme Verteilungen für die Fehlerrate und die a-priori Wahrscheinlichkeiten angenommen. Bei Betrachtung der Gleichung 1 zeigt sich nun, dass die Berechnung der a-posteriori Wahrscheinlichkeit nur noch von der Bestimmung der Likelihoods abhängt. Zur Berechnung der bedingten Wahrscheinlichkeiten müssen in einem nächsten Schritt Annahmen über die Strategien derart getroffen werden, dass zwischen ihnen differenziert werden kann. Umgesetzt wird dies mit Hilfe von drei Itemtypen, für die unterschiedliche Vorhersagen bezüglich der Entscheidungen zwischen den Alternativen getroffen werden. Diese sind in Tabelle 3 exemplarisch dargestellt.

Tabelle 3. Beispiele für Itemtypen und Vorhersagen der Strategien

	<u>Itemtyp 1</u>		<u>Itemtyp 2</u>		<u>Itemtyp 3</u>	
	A	B	A	B	A	B
Cue 1	+	-	+	+	+	-
Cue 2	-	+	+	-	+	+
Cue 3	+	+	-	+	-	+
Cue 4	+	-	-	+	-	-
TTB	A		A		A	
DR	A		B		A oder B (50%)	
Raten	A oder B (50%)		A oder B (50%)		A oder B (50%)	
Items	a-c, d-a, a-e, c-b, e-d, d-f		f-a, b-d, e-b, b-f, d-c		b-a, c-e, f-c, f-e	

Für die sechs in Tabelle 2 präsentierten Objekte ergeben sich bei einer vollständigen Kombination 15 Paarvergleiche⁶ (Binomialkoeffizient). Bei Itemtyp 1 entscheiden sich sowohl TTB als auch DR für dieselbe Alternative. Im Beispiel in Tabelle 3 differenziert bereits der erste Cue zu Gunsten von Option A. Deshalb entscheidet TTB für A. DR entscheidet sich ebenfalls für A, das drei positive Cue-Ausprägungen besitzt. B stimmt nur in zwei Merkmalen mit der Cue-Hierarchie überein. Dieser Itemtyp ist wenig informativ, kommt aber mit sechs Items am häufigsten vor. Itemtyp 2 differenziert zwischen den Strategien. Im Beispiel entscheidet sich TTB für A, da der zweite Cue zu Gunsten von A differenziert. Objekt B besitzt mehr positive Cue-Ausprägungen, weshalb sich die kompensatorische DR Strategie für diese Alternative entscheidet. Die Güte der Klassifikation wird durch die Anzahl der Items des Typs 2 bestimmt. Dieser ist in der vorliegenden Untersuchung fünf Mal vertreten. Bei Items des Typ 3 trifft TTB eine Entscheidung. DR muss raten, da beide Alternativen die gleiche Anzahl positiver Cue-Ausprägungen besitzen. 4 von 15 Items fallen in diese Kategorie.

Nachdem aus den theoretischen Annahmen über die Entscheidungsstrategien gegeneinander testbare Itemtypen spezifiziert wurden, muss die Verteilung unsystematischer Antwortfehler spezifiziert werden. Verwendet eine Person eine Entscheidungsstrategie k mit $k \in \{TTB, DR, Raten\}$, dann entscheidet sie sich mit einer konstanten Fehlerwahrscheinlichkeit ε_k in jedem Durchgang falsch. Diese Wahrscheinlichkeiten stellen modellinkompatible Wahlen dar. Eine Person kann demnach modellkompatible und modellinkompatible Wahlen treffen. Auf dieser Grundlage können die Likelihoods L_k über eine verbundene Binomialverteilung (Gleichung 2) wie folgt berechnet werden.

$$L_k = p(n_{jk}|k, \varepsilon_k) = \prod_{j=1}^3 \binom{n_j}{n_{jk}} \times (1 - \varepsilon^k)^{n_{jk}} \times \varepsilon^{(n_j - n_{jk})} \quad (2)$$

6

$$\left(\frac{6!}{2! \times (6-2)!} \right) = 15$$

Für jede Strategie k wird das Produkt der drei Einzelwahrscheinlichkeiten $p(n_{jk})$ für jeden Itemtyp j mit $j \in \{1,2,3\}$ berechnet. Das Entscheidungsmuster einer Person wird derjenigen Strategie mit der maximalen Likelihood zugeordnet. Eine Einzelwahrscheinlichkeit $p(n_{jk})$ ergibt sich aus der Binomialgleichung als Wahrscheinlichkeit dafür, dass eine Person bei n_j Vergleichen eines Itemtyps j , n_{jk} mit einer der drei Entscheidungsstrategien k modellkompatible Wahlen trifft. Sie berechnet sich aus der Anzahl aller möglichen Reihenfolgen vom Umfang n_j , in denen n_{jk} modellkompatible Wahlen vorkommen (Binomialkoeffizient) multipliziert mit der Wahrscheinlichkeit für n_{jk} modellkompatible Wahlen und der Wahrscheinlichkeit für $n_j - n_{jk}$ modellinkompatible Wahlen (Fehlerwahrscheinlichkeit). Die Wahrscheinlichkeit für eine modellkompatible Wahl berechnet sich als $1 - \varepsilon^k$ und die Wahrscheinlichkeit für eine modellinkompatible Wahl entspricht ε^k (Gleichung 3).

$$\varepsilon_k = \left[\sum_{j=1}^J (n_j - n_{jk}) \right] \div \left[\sum_{j=1}^J n_j \right] \quad (3)$$

Die Schätzung des Fehlers für eine Person und eine bestimmte Strategie k ist der Anteil der modellinkompatiblen Wahlen an allen Wahlen. Ersteres berechnet sich als Differenz der Anzahl der Items eines Typs und der Anzahl der mit diesem Typ kompatiblen Wahlen, aufsummiert über die drei Itemtypen. Für die Strategie Raten und bei Items des Typ $j = 3$ der Strategie DR ist $\varepsilon_k = 0,5$. In diesem Fall rät DR, und es gehen nur die ersten beiden Itemtypen in die Berechnung der Likelihoods ein. Rechnerisch wird dies berücksichtigt, indem für DR $J = 2$ eingesetzt wird. Ist $n_j = n_{jk}$ ergibt sich eine Fehlerwahrscheinlichkeit $\varepsilon_k = 0$. Tritt dieses Ergebnis ein, dann sind alle Entscheidungen kompatibel mit den durch das Modell vorhergesagten. Die Likelihood für eine Strategie, bei der für $\varepsilon_k = 0$ zutrifft, nimmt den Wert Null an. Da in diesem Fall ein Teilnehmer alle Items im Sinne der Strategie k beantwortet hat, kann er gemäß dieser Strategie klassifiziert werden. Wie bei Bröder & Schiffer (2003a, 2003b) werden Entscheidungsmuster mit einer Fehlerwahrscheinlichkeit von

$\hat{\varepsilon} > 0.4$ als Raten klassifiziert. Der einzige Fall, in dem keine Klassifikation erfolgen kann, stellt sich ein, wenn die Likelihoods für mehrere Strategien denselben Wert annehmen.

4 Ergebnisse

Zunächst werden die Ergebnisse zu den Lernzeiten und benötigten Lerndurchgängen, sowie zur Dauer der Entscheidungsphase berichtet. Danach folgt die Betrachtung der Strategieklassifikation, die die Voraussetzung für die im Anschluss berichtete Analyse der abhängigen Variablen gemäß den Hypothesen 1-5 bildet.

4.1 Zeiten und Durchgänge der Lernphase

Zur Erreichung des 90% Kriteriums in der Lernphase benötigten die Versuchsteilnehmer im Mittel 26 Minuten und im Durchschnitt 52,4 Abfragen. Dabei gab es große interindividuelle Unterschiede ($SD = 24,3$; Spannweite = 22 – 108 Durchgänge).

4.2 Zeiten der Entscheidungsphase

Zur Bearbeitung der 17 Entscheidungsdurchgänge wurden im Mittel 4,5 Minuten ($SD = 2,1$ min; Spannweite = 1,9 – 11,2 min) gebraucht. TTB Verwender benötigten durchschnittlich 2,9 Minuten und DR Verwender 3,4 Minuten (der Berechnung dieser Zeiten wurden die Ergebnisse der Strategieklassifikation zu Grunde gelegt). Die benötigte Zeit zur Bearbeitung der Entscheidungsphase unterscheidet sich nicht signifikant. Den Entscheidungsdurchgängen voraus gingen eine erneute Kalibrierungsphase und die Instruktion der Entscheidungsphase. Beides zusammen beanspruchte im Mittel ca. 10 Minuten.

4.3 Bayesianische Strategieklassifikation

In Tabelle 4 sind die Häufigkeiten der klassifizierten Strategien abgebildet. Auffällig ist, dass sich nur 18 von 30 Versuchsteilnehmern gemäß der instruierten Strategie verhielten. Das entspricht einem Anteil von 60%. Demnach haben 8 von 15

Teilnehmern in der TTB Bedingung und 10 von 15 in der DR Bedingung die Manipulation der unabhängigen Variablen verstanden und sich gemäß der Instruktion verhalten. Den anderen Teilnehmern (7 vs. 5) fiel es offensichtlich schwer, die instruierte Entscheidungsregel auf die Entscheidungsaufgabe anzuwenden. Dafür spricht auch der relativ hohe Anteil von 20% der Versuchsteilnehmer, deren Vorgehen als Raten klassifiziert wurde. In bisherigen Untersuchungen in denen Strategien sowohl instruiert als auch post hoc berechnet wurden, liegt der Anteil der als Raten klassifizierten Versuchsteilnehmer zwischen 4 und 20,9% (Bröder & Schiffer, 2003b; Jahn, Renkewitz & Kunze, 2007; Renkewitz & Jahn, 2009). Die als Rater klassifizierten Versuchsteilnehmer wurden aus der Datenanalyse ausgeschlossen, da für ihr Blickverhalten keine Hypothesen vorlagen.

Bei 2 Teilnehmern mit TTB Instruktion und 4 DR Instruierten waren die Likelihoods für die jeweils andere Strategie höher. Zur Erhöhung der statistischen Power wurden diese gemäß ihrem Entscheidungsverhalten und nicht nach der experimentellen Bedingung gruppiert⁷. Das heißt, Teilnehmer, die eine TTB Instruktion erhalten haben, aber als DR Verwender klassifiziert wurden, wurden in der weiteren Analyse gemeinsam mit den DR Instruierten und Klassifizierten betrachtet. Dies betraf die Teilnehmer mit den Nummern 4, 12, 16 und 27. Ebenso wurden die Daten der DR Klassifizierten mit TTB-Instruktion mit denen der DR Instruierten und Klassifizierten verrechnet (Versuchspersonen 14 und 29). Insgesamt verwendeten so 40% der Teilnehmer die nichtkompensatorische TTB Heuristik und 40% die kompensatorische DR Heuristik. Alle Versuchsteilnehmer konnten einer der Strategien (TTB, DR und Raten) zugeordnet werden.

Als Indikator für die Konsistenz des Entscheidungsverhaltens können die durchschnittlichen Fehlerwahrscheinlichkeiten der als TTB oder DR Verwender klassifizierten Teilnehmer herangezogen werden. Diese indizieren, in wie vielen Durchgängen ein Teilnehmer gegen die Vorhersage einer Strategie verstoßen hat. Die durchschnittliche Fehlerrate der TTB Verwender lag bei 22% und die der DR

⁷ Werden die inkonsistenten Datensätze von der Analyse ausgeschlossen, so bleibt das Befundmuster im Hinblick auf die Hypothesen das gleiche. Durch die Umkodierung wird allerdings die statistische Power und damit die Wahrscheinlichkeit vorhandene Effekte aufzudecken erhöht.

Verwender bei 13%. DR Verwender entschieden demnach konsistenter als TTB Verwender. Da DR Verwender die Cue-Hierarchie nicht berücksichtigen müssen und nur die Cue-Ausprägungen abrufen, ist ihnen die Bearbeitung der Aufgabe eventuell leichter gefallen und sie haben deshalb im Mittel weniger Fehler gemacht.

Tabelle 4. Häufigkeiten und Prozentanteile (in Klammern) der Strategieklassifikation nach Bayes

	Strategieklassifikation			<i>N</i>
	TTB	DR	Raten	
TTB	8 (26,6)	2 (6,7)	5 (16,6)	15
DR	4 (13,4)	10 (33,3)	1 (3,4)	15
<i>N</i>	12 (40,0)	12 (40,0)	6 (20,0)	30

4.4 Blickwechsel zwischen Alternativen pro Sekunde

Die erste Hypothese bezieht sich auf die Anzahl der Blickwechsel zwischen den Alternativen relativ zur Zeitdauer des Durchgangs. TTB Verwender sollten mehr Blickwechsel zwischen den Alternativen zeigen als DR Verwender, da erstere einzelne Attribute miteinander vergleichen. DR Verwender gehen dagegen eher optionsweise vor, was sich in geringeren Anzahlen von Blickwechseln pro Zeiteinheit zeigen sollte.

Zunächst mussten die Blickwechselfrequenzen zwischen den Alternativen bestimmt werden. Dazu wurden mit Hilfe der Auswertungssoftware BeGaze quadratische Areas of Interest (AOIs) definiert. Jedes überdeckte eine Alternative vollständig und ca. die Hälfte des Bildschirms. Der Abstand zu den Stimuli war an allen Stellen konstant. In der Bildschirmmitte grenzten die AOIs aneinander (Abbildung D1). Ein Blickwechsel wurde definiert als zwei aufeinander folgende Fixationen in verschiedene AOIs. Eine Fixation wurde in BeGaze durch eine minimale Fixationsdauer von 80 ms und eine Ausdehnung von 2° Sehwinkel definiert. Auf dieser Grundlage errechnete die Software die Anzahl der Blickwechsel

von dem linken ins rechte und von dem rechten ins linke AOI. Diese wurden anschließend in SPSS eingelesen.

Danach erfolgte eine explorative Betrachtung der *Blickwechsel gesamt*, der Reaktionszeiten und der Qualität der Blickdaten. Für die Beantwortung der Fragestellung sind nur die Gesamtanzahl der Blickwechsel von links nach rechts und von rechts nach links pro Teilnehmer und Durchgang relevant. Deshalb wurden diese zu einem Wert Blickwechsel gesamt summiert. Im Anschluss daran erfolgte eine explorative Betrachtung der Daten. Dazu wurden die Daten zunächst mit Hilfe von Streudiagrammen betrachtet. Anschließend wurden extreme Ausreißer in Box Plots identifiziert. Werte über 18 Blickwechsel gesamt pro Durchgang, sowie Blickwechselhäufigkeiten, die 3 Interquartilsbereiche über dem 75% Quartil lagen, mussten von der Analyse ausgeschlossen werden. Dies betraf 3% der Werte. Ein Wert ist definiert als das Ergebnis eines Teilnehmers in einem der getesteten Indizes pro Entscheidungsdurchgang. Bei den Reaktionszeiten blieben Werte über 45 Sekunden bis zum Tastendruck (2% der Reaktionszeiten) unberücksichtigt.

Zur Beurteilung der Qualität der Blickdaten wurden die Zeiten eines Durchgangs, in denen ein Versuchsteilnehmer in eines der beiden AOIs sah, relativ zur Gesamtzeit eines Durchgangs betrachtet. Durchgänge in denen dieser Wert unter 60% der Zeit lag, konnten in der Datenanalyse nicht berücksichtigt werden. Gründe für einen niedrigen Wert können sein, dass die Versuchsteilnehmer in dieser Zeit nicht in eines der beiden AOIs sahen oder der Eye Tracker aufgrund eines Verlusts der Korneareflexion keine Blickdaten aufzeichnen konnte. Schließlich können die Versuchsteilnehmer auch durch häufiges Blinzeln die Aufzeichnung unterbrochen haben. In 20.8% aller Durchgänge über alle Versuchsteilnehmer hinweg lag dieses Kriterium unter dem kritischen Wert von 60% (62 von 360 Werten). Versuchsteilnehmer 12 und 27 fielen ganz aus der Analyse heraus.

Im Anschluss wurde der Index *Blickwechsel gesamt pro Sekunde* berechnet. Die Anzahl der Blickwechsel soll für TTB-Verwender abhängig sein von der Position des ersten differenzierenden Cues in der Cue-Hierarchie. Je tiefer der erste differenzierende Cue, desto mehr Blickwechsel sollten auftreten und desto länger sollte ein Durchgang dauern. Um ein von der Dauer unabhängiges Maß zu erhalten,

wurde die Gesamtanzahl der Blickwechsel eines Durchgangs durch seine Dauer geteilt. Für den anschließenden Vergleich der Blickwechsel pro Sekunde zwischen den beiden instruierten Strategien wurden die Daten der 15 Entscheidungsdurchgänge pro Person aggregiert.

Der im Anschluss durchgeführte t-Test für unabhängige Stichproben zum Vergleich der mittleren Anzahl der Blickwechsel pro Sekunde zwischen den Bedingungen TTB und DR wurde nicht signifikant. ($t(20) = 1.04$, $p = .156$, $d = 0.47$). Abbildung 2 veranschaulicht dennoch die Tendenz für eine erhöhte Anzahl von Blickwechseln pro Sekunde bei TTB Verwendern. Eine post hoc durchgeführte Power-Analyse weist darauf hin, dass die Teststärke zu gering war, um einen vorhandenen Effekt statistisch aufzudecken ($1-\beta = .28$).

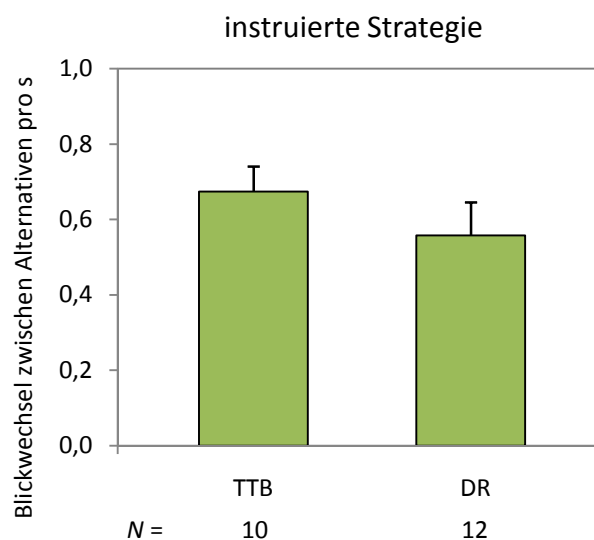


Abbildung 2. Mittlere Anzahl der Blickwechsel zwischen den Alternativen pro Sekunde für die nach der Strategieklassifikation klassifizierten TTB bzw. DR Verwender. Die Fehlerbalken entsprechen dem Standardfehler des Mittelwertes.

4.5 Blickwechsel in Abhängigkeit der Position des ersten differenzierenden Cues

Aufgrund der attributweisen, nicht kompensatorischen Vorgehensweise bei der Verwendung der TTB Heuristik sollte die absolute Häufigkeit der Blickwechsel bei

den TTB Verwendern linear mit der Position des ersten differenzierenden Cues steigen. Das heißt, je später der erste differenzierende Cue, desto mehr Blickwechsel zeigen die TTB Verwender. Bei DR Verwendern sollte dieser Trend nicht auftreten. Zur Überprüfung der Fragestellung wurden die Blickwechsel gesamt pro Person zu vier Werten aggregiert. Jeder Wert entspricht den Antworten einer Person auf den Items einer Gruppe. Eine Gruppe wird definiert durch die Items, bei denen entweder der erste, zweite, dritte oder vierte Cue differenziert. Die vier Gruppen und die Vorhersage der Hypothese sind in Tabelle 5 dargestellt. Bei der folgenden Berechnung wurden die gemittelten Blickwechsel gesamt für die Gruppe 1 mit denen der Gruppe 2 verglichen. Dieser Test wurde bis Gruppe 4 fortgesetzt.

Tabelle 5. Paarvergleiche geordnet nach Gruppen bei denen TTB differenziert (Entscheidungszeit = Zeit, Blickwechsel = BW)

Gruppe 1	Gruppe 2	Gruppe 3	Gruppe 4
Cue 1 diff.	Cue 2 diff.	Cue 3 diff.	Cue 4 diff.
b-a, a-c, b-d, e-b, b-f, d-c, c-e, f-c,	f-a, d-f, f-e	d-a, a-e,	c-b, e-d
Vorhersage TTB:	BW Gruppe 1 < BW Gruppe 2 < BW Gruppe 3 < BW Gruppe 4 Zeit Gruppe 1 < Zeit Gruppe 2 < Zeit Gruppe 3 < Zeit Gruppe 4		

Abbildung 3 zeigt einen Anstieg in den Blickwechseln gesamt, je später der erste Cue in der Reihenfolge der Validitäten erscheint. Bei DR Verwendern findet sich dieser Anstieg nicht. Mit Hilfe von Kontrastanalysen ist es möglich, präzise Hypothesen wie die hier vorliegende zu testen. Dazu werden die Vorhersagen durch Kontraste spezifiziert. Dieses Vorgehen ermöglicht zudem die Berechnung teststärkerer Ergebnisse im Vergleich zu herkömmlichen Varianzanalysen (Sedlmeier & Renkewitz, 2008). Sowohl für die Blickwechsel gesamt als auch für die Reaktionszeiten (Hypothese 3) wurde ein linearer Anstieg über die 4 Gruppen differenzierender Cues vermutet. Auf der Grundlage bisheriger Befunde zu Reaktionszeiten (Bergert & Nosofsky, 2007; Bröder & Gaissmaier, 2007; Renkewitz

& Jahn, 2009) wurde ein linearer Kontrast wie folgt spezifiziert: $\lambda_{Cue1} = -3$, $\lambda_{Cue2} = -1$, $\lambda_{Cue3} = 1$, $\lambda_{Cue4} = 3$. Bei der Kontrastanalyse für abhängige Stichproben wurde im Anschluss ein Maß für die Passung der Vorhersage (Kontrast) mit dem Ergebnis berechnet. Dieses wurde dann wie bei dem t-Test im Einstichprobenfall gegen Null getestet. Bei einem einseitigen $\alpha = 0.05$ wurde der vorliegende Kontrast für die TTB Verwender signifikant ($t_{Kontrast}(9) = 2.58$, $p = .015$, $g = 0.82$). Das Effektstärkenmaß g ist den üblichen Konventionen nach (Cohen, 1992, zit. nach Sedlmeier & Renkewitz, 2008) hoch und spricht damit für ein konsistentes Auftreten des Effekts über die Versuchsteilnehmer hinweg. Bei den DR Verwendern zeigte sich dieser Anstieg nicht ($t_{Kontrast}(10) = 1.39$, $p = .097$, $g = 0.42$). Mit Hilfe der Kontrastanalyse war es möglich den linearen Anstieg der Blickwechselhäufigkeiten mit der Position des ersten differenzierenden Cues statistisch aufzudecken.

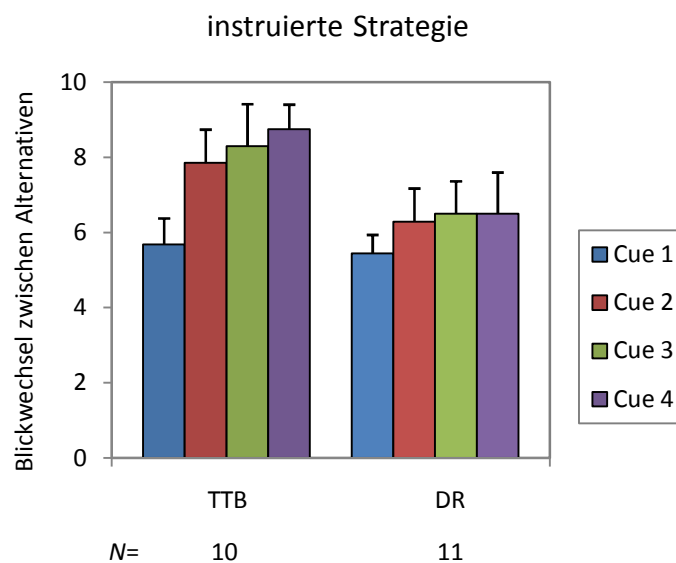


Abbildung 3. Mittlere Anzahl der Blickwechsel zwischen den Alternativen in Abhängigkeit des ersten differenzierenden Cues für die nach der Strategieklassifikation klassifizierten TTB bzw. DR Verwender. Die Fehlerbalken entsprechen dem Standardfehler des Mittelwertes.

Anhand der Blickwechsel gesamt scheint es demnach möglich, das serielle Vorgehen bei der Verwendung der TTB Heuristik abzubilden. Bei kompensatorischer Vorgehensweise nach der DR Heuristik findet sich dieser Anstieg in den

Blickwechselln gesamt über die definierten Gruppen differenzierender Cues wie erwartet nicht.

4.6 Entscheidungszeiten

Interessant ist nun, ob sich der beobachtete Anstieg der Blickwechsellhäufigkeiten auch in den Entscheidungszeiten der vorliegenden Untersuchung finden lässt. Dazu wurden dieselben Gruppeneinteilungen wie beim Test der zweiten Hypothese verwandt (Tabelle 5). Ebenso wurden die Daten auf dieselbe Art und Weise aggregiert. Es wird angenommen, dass die Entscheidungszeiten der TTB Verwender zunehmen, je später der erste differenzierende Cue erscheint. Bei DR Verwendern sollte kein Anstieg der Zeiten zu finden sein.

Abbildung 4 veranschaulicht den Trend über einen Anstieg der Reaktionszeiten der TTB Verwender über die 4 Gruppen differenzierender Cues hinweg. Bei DR Verwendern findet sich dieses Muster wie erwartet nicht. Auch für die Reaktionszeiten wurde eine Kontrastanalyse für abhängige Stichproben gerechnet. Da der relative Abstand zwischen den Lambdagewichten und nicht die absolute Größe für die Berechnung der Kontrastanalyse entscheidend ist, wurden dieselben Kontraste wie beim Test der Hypothese 2 verwandt ($\lambda_{Cue1} = -3, \lambda_{Cue2} = -1, \lambda_{Cue3} = 1, \lambda_{Cue4} = 3$). Abbildung 4 zeigt den linearen Anstieg der Reaktionszeiten für die TTB Verwender ($t_{Kontrast}(9) = 2.65, p=.014, g = 0.83$). Bei den DR Verwendern wird der gleiche lineare Kontrast nicht signifikant ($t_{Kontrast}(10) = 1.81, p=.051, g = 0.54$).

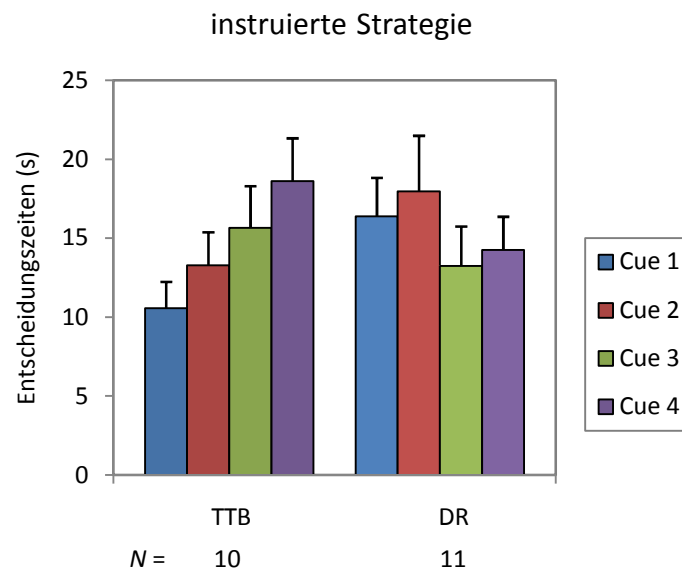


Abbildung 4. Mittlere Entscheidungszeiten in Sekunden in Abhängigkeit des ersten differenzierenden Cues für die nach der Strategieklassifikation klassifizierten TTB bzw. DR Verwender. Die Fehlerbalken entsprechen dem Standardfehler des Mittelwertes.

Wie in den bisherigen Untersuchungen zu Reaktionszeiten, in denen diese als Prozessmaß zum Abruf der Informationen bei gedächtnisbasierten Multi-Attribut-Entscheidungen verwendet wurden (Bröder & Gaissmaier, 2007; Jahn, Renkewitz & Kunze, 2007; Renkewitz & Jahn, 2009), konnte auch für die vorliegende Untersuchung der Anstieg der Reaktionszeiten für TTB Verwender gezeigt werden. In diesem bewährten Maß zeigt sich demnach dasselbe Muster, wie in den Blickwechseln gesamt als Parameter der Blickbewegungsmessung.

4.7 Anzahl der Fixationen vor einem Blickwechsel

Die vierte Hypothese bezog sich auf die Anzahl der Fixationen in ein AOI bevor ein Blickwechsel in das andere AOI stattfand. Gemäß der Vorhersage sollten hier DR Verwender mehr Fixationen innerhalb einer Alternative vor einem Blickwechsel zeigen als TTB Verwender. Zur Überprüfung der Hypothese wurde ein weiterer Index *Anzahl der Fixationen bevor Blickwechsel* berechnet. Dazu mussten,

wie bei den Blickwechseln gesamt, die Anzahlen der Fixationen in die AOIs mit Hilfe von BeGaze berechnet und in SPSS eingelesen werden. Die anschließende explorative Datenanalyse zeigte, dass 2% der Werte über 80 Fixationen gesamt lagen. Diese wurden von der Datenanalyse ausgeschlossen.

Der Index Fixationen gesamt vor Blickwechsel berechnete sich als Summe der Fixationen ins linke und rechte AOI dividiert durch die Anzahl der Blickwechsel gesamt für jeden Durchgang. Die 15 Werte einer Person für diesen Index wurden zu einem zusammengefasst. Der im Anschluss gerechnete t-Test für unabhängige Stichproben wurde nicht signifikant ($t(20) = 1.58$, $p = .065$, $d = 0.71$). Die post hoc durchgeführte Power Analyse zeigte, dass die Wahrscheinlichkeit diesen angenommenen Effekt mit inferenzstatistischen Verfahren aufzudecken klein war ($1 - \beta = .48$). Abbildung 5 veranschaulicht dennoch die Tendenz für eine erhöhte Anzahl von Fixationen in eine Alternative vor einem Blickwechsel bei den DR Verwendern, im Gegensatz zu den TTB Verwendern.

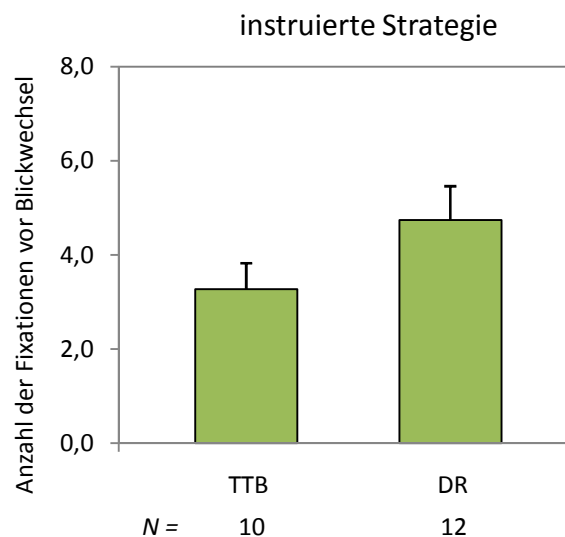


Abbildung 5. Mittlere Anzahl der Fixationen in ein AOI vor einem Blickwechsel in das andere AOI für die nach der Strategieklassifikation klassifizierten TTB bzw. DR Verwender. Die Fehlerbalken entsprechen dem Standardfehler des Mittelwertes.

4.8 Verweildauern vor einem Blickwechsel

Ähnlich wie die Anzahl der Fixationen sollte auch die Gesamtdauer aller Fixationen in ein AOI, bevor ein Blickwechsel stattfand bei DR Verwendern länger sein als bei TTB Verwendern. Bevor der Index *Dauer der Fixation vor Blickwechsel* berechnet wurde, mussten die Dauern der Fixationen in die AOIs mit Hilfe von BeGaze berechnet und in SPSS eingelesen werden. 7 Werte (2.4%) wurden aus der Analyse ausgeschlossen, da diese einen Dauer von 40 sec pro Durchgang überschritten oder 3 Interquartilsbereiche über dem 75% Quartil lagen. Die Dauer der Fixationen gesamt vor einem Blickwechsel berechnet sich als Summe der Dauern ins linke und rechte AOI geteilt durch die Blickwechsel gesamt. Der Index wurde über die 15 Werte pro Person, entsprechend den 15 Entscheidungsdurchgängen, gemittelt. Abbildung 6 zeigt, dass DR Verwender länger in ein AOI sahen, bevor sie in das andere AOI wechselten, als TTB Verwender.

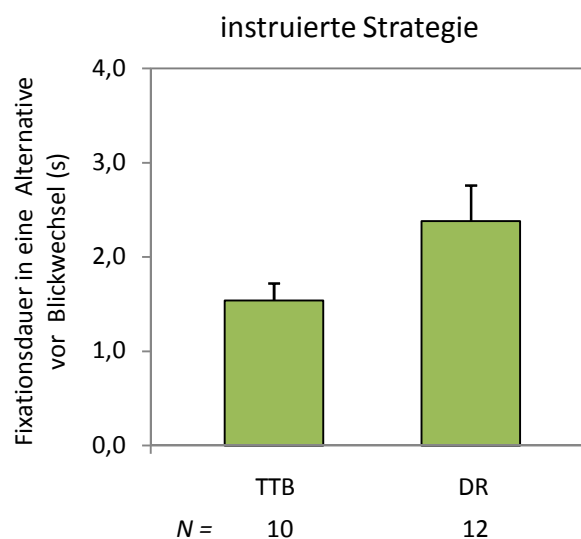


Abbildung 6. Mittlere Anzahl der Fixationsdauern in ein AOI vor einem Blickwechsel in das andere AOI für die nach der Strategieklassifikation klassifizierten TTB bzw. DR Verwender. Die Fehlerbalken entsprechen dem Standardfehler des Mittelwertes.

Der entsprechende t-Test für unabhängige Stichproben bei einem einseitigen α von 0.05 wird signifikant ($t(20) = 1.89, p = .037, d = 0.85$). Der Effekt ist groß (Cohen,

1992, zit. nach Sedlmeier & Renkewitz, 2008). Dieses Ergebnis ist inhaltlich plausibel insofern, dass DR Verwender zunächst versuchen, alle Cue-Ausprägungen einer Alternative abzurufen, bevor sie dies für die andere Option tun. Dieses Vorgehen beansprucht mehr Zeit als nach dem ersten differenzierenden Cue zu suchen, wie es für TTB Verwender vorhergesagt wurde. Dies traf zumindest für die Items zu, bei denen spätestens der dritte Cue differenzierte.

5 Diskussion

5.1 Ziel der Untersuchung

Ziel der vorliegenden Untersuchung war es zu testen, ob die Gedächtnissuche nach Attributinformationen bei verschiedenen Entscheidungsstrategien zu unterschiedlichen Suchmustern führte. Zur Überprüfung dessen wurde eine neue multimethodale Herangehensweise entwickelt, deren zentraler Bestandteil die Einführung der Eye Tracking Methode als Prozessmaß in einer gedächtnisbasierten Multi-Attribut Entscheidungssituation war.

5.2 Zusammenfassung und Interpretation der Ergebnisse

Die aus der Entscheidungsaufgabe resultierenden Suchmuster wurden mit Hilfe von vier verschiedenen Blickbewegungsindizes betrachtet. Diese leiteten sich aus Überlegungen zur Informationssuche bei gedächtnisbasierten Entscheidungen ab. Zusätzlich wurden die Entscheidungszeiten analysiert. Zunächst werden die Ergebnisse der Blickwechsel pro Sekunde (Hypothese 1), die Anzahl der Fixationen vor einem Blickwechsel (Hypothese 4) und die Dauer der Fixationen vor einem Blickwechsel (Hypothese 5) diskutiert. Die Blickwechsel in Abhängigkeit des ersten diskriminierenden Cues (Hypothese 2) werden gemeinsam mit den Ergebnissen der Entscheidungszeiten (Hypothese 3) betrachtet.

Alle Indizes weisen deskriptiv in hypothesenkonforme Richtung. Das heißt, die Blickbewegungen auf den geleerten Informationstafeln unterscheiden sich in erwarteter Weise zwischen den verschiedenen Entscheidungsstrategien. Konkret zeigten TTB Verwender mehr Blickwechsel pro Sekunde zwischen den beiden Alternativen als DR Verwender. Wie angenommen, zeigten DR Verwender mehr Fixationen innerhalb einer Alternative, bevor sie den Blick auf die andere Alternative richteten als TTB Verwender. Allerdings zeigt sich erst bei der Dauer der Fixationen vor einem Blickwechsel, dass DR Verwender signifikant länger innerhalb einer Alternative verweilten, bevor sie den Blick wechselten als TTB Verwender.

Die Befunde deuten darauf hin, dass für die nichtkompensatorische TTB Heuristik mehr attributweise Suche auftritt. Das zeigt sich in der erhöhten Anzahl von Blicken zwischen den Alternativen, in der geringeren Anzahl von Fixationen in einer Alternative und damit auch in einer geringeren Verweildauer innerhalb einer Alternative. Für die kompensatorische DR findet sich ein Suchmuster, das eher für optionsweise Suche charakteristisch ist. DR Verwender zeigen weniger Blickwechsel zwischen den Alternativen pro Sekunde und mehr Fixationen innerhalb einer Alternative, die mit einer höheren Verweildauer einhergehen.

Im Folgenden werden drei mögliche Erklärungen für das Ausbleiben signifikanter Ergebnisse der ersten und vierten Hypothesen erläutert. Erstens resultierte die geringe Anzahl von Versuchsteilnehmern, die pro Bedingung analysiert wurden, in einer zu geringen statistischen Power (.28 und .48). Würde die Wahrscheinlichkeit, den Effekt zu finden, wenn er zu finden ist, in einer Folgeuntersuchung erhöht werden, könnten auch die bisher nur deskriptiv beobachteten Effekte signifikant werden.

Zweitens muss auch in Betracht gezogen werden, dass die Blicke in der Entscheidungsphase nicht indikativ für den Informationsabruf aus dem Gedächtnis gewesen sein könnten. Dafür verantwortlich kann eine zu schwache Verknüpfung des Gedächtnisabrufs mit Blickbewegungen sein. Das heißt, der Gedächtnisabruf führte möglicherweise nicht in dem erwarteten Ausmaß zu Blickbewegungen an die Orte der Informationspräsentation. Unterschiede zwischen dem für diese Untersuchung gestalteten Lernparadigma und den Untersuchungsmaterialien, die bisher zur Überprüfung des Looking at Nothing-Phänomens eingesetzt wurden (Hoover & Richardson, 2008; Richardson & Spivey, 2000; Spivey & Geng, 2001), können für diese schwache Verknüpfung ursächlich sein.

Ein wesentlicher Unterschied zu den früheren Studien bestand darin, dass in der Mitte der Anordnung von vier Kästen Bilder von sechs verschiedenen Pilzen präsentiert wurden. Das war notwendig, damit die Alternativen in der Entscheidungsphase auch ohne Cue-Informationen unterscheidbar waren. Bei der Datenanalyse zeigte sich allerdings, dass ein Großteil der Versuchsteilnehmer stärker die Bilder in der Mitte der Stimuli fokussierten als die leeren Kästen, die in der

Lernphase die Cue-Informationen enthielten. Weiter wurden im Unterschied zu den bisherigen Untersuchungen zum Looking at nothing-Phänomen, in dem hier verwendeten Lernparadigma mehr als eine Information mit einem Ort auf dem Bildschirm verknüpft. Insgesamt lernten die Versuchsteilnehmer sechs verschiedene Pilze, die sich hinsichtlich der Ausprägungen in vier Cues unterschieden. Jeder Cue lag in drei verschiedenen Ausprägungen vor, die über die sechs Alternativen hinweg variierten. Jeder Kasten eines Objekts war demnach mit sechs Informationen belegt.

Die ausgedehnte Lernprozedur kann schließlich bei den Versuchsteilnehmern zur Anwendung mnemonischer Strategien geführt haben, die weniger Blicke zu den jeweiligen Ausprägungen der Alternativen erforderten. Einige Versuchsteilnehmer berichteten im Nachhinein, die Merkmale der Pilze mit Merkmalen der Pilzbilder assoziiert zu haben (zum Beispiel, Farbe oder Form des Pilzes). Oder sie bildeten Akronyme aus den Anfangsbuchstaben der Attributausprägungen. Die verwendeten mnemonischen Strategien können die Speicherung und den Abruf visueller Indizes überlagert haben (Ostendorf, Finke & Ploner, 2004). Sowohl die erhöhte Salienz der Bilder als auch die Anzahl zu lernender Informationen können den Einsatz mnemonischer Strategien gefördert und dazu geführt haben, dass die Blickwechsel nicht in dem Maß indikativ für den Gedächtnisabruf waren wie in anderen Untersuchungen des Looking at nothing-Phänomens.

Eine dritte Erklärung stellt die Annahme in Frage, dass nur diejenigen Informationen abgerufen und damit angesehen werden, die für die Entscheidungsfindung gemäß einer bestimmten Strategie erforderlich sind (Abelson & Levi, 1985; Bröder, 2005). Stattdessen kann es sein, dass alle Informationen angesehen werden, aber nur der Teil der Informationen integriert wird, der gemäß einer bestimmten Strategie für die Entscheidungsfindung relevant ist. Die Befunde zu den Blickwechseln je später der erste differenzierende Cue erscheint als auch die der Entscheidungszeiten sprechen eher dagegen. Denn sowohl die Anzahl der Blickwechsel zwischen den Alternativen, als auch die Reaktionszeiten steigen bei den TTB Verwendern, je später in der Validitätenreihenfolge erstmals ein Cue differenziert. Bei den DR Verwendern findet sich dieser Anstieg wie erwartet nicht. Das Ergebnis der Blickwechsel in Abhängigkeit des ersten differenzierenden Cues

spricht für einen attributweisen und nicht vollständigen Informationsabruf, der für nichtkompensatorische Strategien angenommen wird (Payne, Bettman & Johnson, 1993; Riedl, Brandstätter & Roithmayr, 2008).

Dieser Befund wird zusätzlich von den Ergebnissen der Entscheidungszeiten unterstützt. Diese replizieren frühere Befunde (Bröder & Gaissmaier, 2007; Jahn, Renkewitz & Kunze, 2007; Renkewitz & Jahn, 2009). Auffällig bei den Entscheidungszeiten der DR Verwender ist ein nicht überzufälliger Anstieg der Entscheidungszeiten bis zum zweiten Cue. Eine Erklärung dafür könnte sein, dass die Teilnehmer, die als DR Verwender klassifiziert wurden, die Strategie nicht konsistent anwendeten.

Bevor allerdings die Ergebnisse der Strategieklassifikation noch einmal genauer betrachtet werden, lässt sich für die Ergebnisse zu den Hypothesen festhalten, dass diese zumindest teilweise mit den Vorhersagen übereinstimmen, die aus den theoretischen Annahmen über die strategiespezifische Informationssuche abgeleitet wurden. Das heißt, es ist prinzipiell möglich, das Looking at nothing-Phänomen zu nutzen, um damit distinkte Suchmuster in den Blickbewegungen für verschiedene Entscheidungsstrategien in einer gedächtnisbasierten Entscheidungsaufgabe abzuleiten. Ebenso scheint das Verfolgen der Blickbewegungen hilfreich bei der Unterscheidung kognitiver Strategien, die bei gedächtnisbasierten Multi-Attribut-Entscheidungen auftreten.

Die Strategieklassifikation diene als unabhängiger Indikator für die Strategiewahl. Indem die Klassifikation vor der Analyse der Suchmuster durchgeführt wurde, konnte die Güte der Strategieinstruktion erfasst werden. Dabei zeigte sich, dass sechs Teilnehmer entgegen der instruierten Strategie entschieden. Dies konnte mit Hilfe der Klassifikation aufgedeckt und die entsprechenden Teilnehmer bei der Analyse der jeweils anderen Gruppe zugeordnet werden. Nicht zuletzt erlaubte die Strategieklassifikation, die Suchmuster unter der Voraussetzung des Einsatzes einer kompensatorischen oder nichtkompensatorischen Entscheidungsstrategie zu vergleichen.

Zwei Ergebnisse der Strategieklassifikation werden im Folgenden noch einmal näher betrachtet. TTB Verwender antworteten inkonsistenter als DR Verwender (22% vs. 13%). Es gibt einen relativ hohen Anteil (20%) von Teilnehmern, die als Rater klassifiziert wurden. Diese Ergebnisse könnten dadurch erklärt werden, dass Teilnehmer ganz andere als die hier getesteten Strategien (TTB, DR und Raten) verwendeten. Es kann sein, dass Versuchsteilnehmer, die als TTB klassifiziert wurden, eine linear gewichtete Regel wie WADD anwendeten oder dass sie die Strategie im Verlauf der Entscheidungsphase wechselten. Die Strategieklassifikation berücksichtigt zwar Fehler im Antwortverhalten der Teilnehmer durch die Berechnung eines Fehlermodells. Dennoch muss der Fehler weder, wie vom Modell angenommen, uniform verteilt sein, noch müssen Teilnehmer nur die in der Klassifikation berücksichtigten Strategien verwendet haben.

Obwohl die Instruktion an einem Beispiel durchgesprochen wurde, kann eine mögliche Erklärung für den hohen Anteil an Ratern auch durch Nichtverständnis der Instruktion entstanden sein. Plausibler scheint aber die Erklärung, dass aufgrund der Zeit, die zwischen Lern- und Entscheidungsphase verstrichen ist, die Cue-Informationen nicht mehr ausreichend gut abgerufen werden konnten. Insgesamt war die Güte der vorliegenden Klassifikation ausreichend, um die zu testenden Strategien zuverlässig zu klassifizieren.

5.3 Forschungspraktische Konsequenzen und Ausblick

Der mit dieser Untersuchung vorgelegte erste Test der Blickbewegungsmessung ergab, dass sich die Suchmuster kompensatorischer und nichtkompensatorischer Entscheidungsstrategien in einer gedächtnisbasierten Entscheidungssituation zumindest auf Alternativebene unterscheiden. Wünschenswert wäre eine Unterscheidung auf Attribut- oder Individuumsebene und idealerweise auf Itemebene. Dies würde Analysen interindividueller Unterschiede erlauben. Damit könnte gezeigt werden, dass Strategien keine distinkten Kategorien kognitiver Prozesse darstellen müssen, sondern Ergebnis eines Mechanismus sind, der mit einem individuell verschiedenen Bedürfnis nach Informiertheit einhergeht

(Newell, 2005; Lee & Cummins, 2004). Dazu ist allerdings die Kenntnis der genauen Bedingungen erforderlich, unter denen das Looking at nothing-Phänomen in gedächtnisbasierten Multi-Attribut Entscheidungsumwelten gezeigt wird.

Einen offensichtlichen Einfluss hat wie in anderen Untersuchungen zu gedächtnisbasierten Multi-Attribut-Entscheidungen (Bröder, 2003b; Jahn, Renkewitz & Kunze, 2007) die Gestaltung des Materials. Die verwendeten Bilder zogen offensichtlich zu viel Aufmerksamkeit auf sich. Deshalb ersetzten Renkewitz und Jahn (2009) in einer Folgeuntersuchung die Bilder durch Symbole (zum Beispiel Kreise oder Rechtecke), die sich nicht mehr in der Mitte, sondern außerhalb der vier Kästen befanden.

Außerdem erleichterten sie die Speicherung der Informationen durch einen Gedächtnistest zu Beginn der Entscheidungsphase. Dabei sahen die Versuchsteilnehmer die Stimuli einzeln ohne Cue-Informationen und riefen diese aus dem Gedächtnis ab. Anschließend sahen sie den Stimulus noch einmal mit Cue-Informationen. Auf diese Weise konnten vergessene Cue-Ausprägungen noch einmal gelernt werden. Eine Verkürzung der Zeitspanne zwischen Lern- und Entscheidungsphase wirkte dem Vergessen zusätzlich entgegen. Es waren offensichtlich diese Veränderungen im Lernparadigma, die dazu führten, dass die Blickwechsel pro Sekunde in der Folgeuntersuchung statistische Signifikanz erreichten. Weiterhin analysierten die Autoren die Blickdaten auf Attributebene. Überraschend zeigte sich, dass TTB Verwender alle Cues länger betrachteten, je später der erste diskriminierende Cue erschien. Differenzierte zum Beispiel der dritte Cue, so stieg die Verweildauer im Feld, das den dritten und vierten Cue beinhaltete, stärker an als im ersten und zweiten Feld. Bei DR Verwendern fand sich dieser Effekt nicht. Dieser Befund könnte auf die bereits erwähnte Betrachtung aller Cues hindeuten und deren Integration in Abhängigkeit der Position des ersten differenzierenden Cues. Die Optimierungen im Lernparadigma führten insgesamt zu klareren Ergebnissen und ermöglichten eine Analyse auf Attributebene. Weitere Untersuchungen sollten genauer prüfen, ob die Suchmuster indikativ für die Cue-Integration sind oder ob die Versuchsteilnehmer alle Informationen betrachten und nur die von einer Strategie (TTB oder DR) „geforderten“ Informationen integrieren.

Neben den bisher genannten Punkten des Untersuchungsmaterials und der Speicherung sollte auch eine Erleichterung der Enkodierung zu einem stärkeren Looking at nothing-Effekt führen. Eine Möglichkeit könnte sein, die Anzahl zu lernender Informationen zu reduzieren oder zumindest weniger Informationen mit einer räumlichen Position zu verknüpfen. Um eine reliable Klassifikation zu gewährleisten, wurde allerdings sowohl in der vorliegenden als auch in der Folgeuntersuchung, die das hier entwickelte Lernparadigma verwendete, die Anzahl von 6 zu lernenden Objekten nicht unterschritten. Renkewitz und Jahn (2009) erhöhten die Reliabilität der Klassifikation zusätzlich, indem sie die Items des Typs 2 doppelt vorgaben, denn diese führen zu unterschiedlichen Vorhersagen für TTB und DR Verwender.

Um den Ergebnissen mehr Gewicht zu geben, ist es unerlässlich, neben der Instruktion der Strategien eine Bedingung zu gestalten, in der die kognitiven Prozesse zur Entscheidungsfindung nicht vorgegeben sind. In der Folgeuntersuchung von Renkewitz und Jahn (2009) ergaben sich unter freier Strategiewahl ähnliche Suchmuster wie in der instruierten Bedingung.

Mögliche Anwendungsfelder der neuen methodischen Vorgehensweise bieten Untersuchungssituationen, die versuchen Faktoren zu identifizieren, die einen Einfluss auf die Strategiewahl in einer gedächtnisbasierten Multi-Attribut Entscheidungssituation haben könnten. Beispielhaft seien hier Zeitdruck (Christensen-Szalanski, 1980; Rieskamp & Hoffrage, 2008), externe monetäre Kosten (Bröder, 2000b) oder der Einfluss von Persönlichkeitsvariablen wie Intelligenz (Bröder, 2005) genannt. Für diese Art von Untersuchungen kann mit Hilfe der Aufzeichnung und Analyse der Blickbewegungen gezeigt werden, wie sich die zu Grunde liegenden kognitiven Prozesse verändern.

Ein weiterer Einsatzbereich könnte sich aus den Befunden von Newell, Rakow, Weston & Shanks (2004) ergeben. Sie zeigten, dass für die Bewertung des Nutzens eines Cues in einer Entscheidungssituation nicht nur die Validität herangezogen wird wie es die TTB Heuristik annimmt, sondern auch die Diskriminationsrate. Das ist die „Basisrate von Cue-Konstellationen, in denen ein Cue zwischen Objekten diskriminiert“ (Bröder, 2005, S. 68). Das Produkt aus Validität und

Diskriminationsrate ist der Success eines Cues. In einer fiktiven Entscheidungssituation bei der die Informationen vorgegeben waren, gestalteten sie die Alternativen derart, dass Entscheidungsstrategien basierend auf den Validitäten, den Diskriminationsraten und dem Success zu unterschiedlichen Suchreihenfolgen führten. Diese entsprachen eindeutig dem Success-Kriterium. Allerdings setzten die Autoren hohe externe Kosten mit hohen kognitiven Kosten gleich. Mit Hilfe der hier vorgestellten methodischen Vorgehensweise ließe sich überprüfen, ob der Success auch kognitive Prozesse in gedächtnisbasierten Multi-Attribut Entscheidungssituationen dominiert.

Zusammenfassend lässt sich festhalten, dass Outcome basierte Verfahren zwar versuchen, die Lücke zwischen Input und Output durch die Analyse resultierender Entscheidungsmuster zu schließen (Bröder, 2005). Können die Alternativen allerdings nicht derart gestaltet werden, dass die Entscheidungsstrategien unterschiedliche Alternativen präferieren, kommen Outcome basierte Klassifikationsmethoden zu keinen informativen Ergebnissen (Harte & Koele, 2001). An dieser Stelle helfen Prozessmaße, denn diese können zumindest auf aggregierter Gruppenebene zeigen, dass sich die Anzahl der Fixationen und Blickwechsel zwischen verschiedenen Entscheidungsstrategien unterscheiden. Für den Bereich der gedächtnisbasierten Entscheidungen sind die Prozessmaße bisher rar gewesen. Nur die Entscheidungszeiten haben sich als ökonomisches und reliables Instrument erwiesen. Die Ergebnisse dieser Untersuchung stützen die Idee, dass Blickbewegungen wertvolle Informationen über die kognitiven Prozesse während der Gedächtnissuche liefern können. Damit steht der gedächtnisbasierten Entscheidungsforschung ein weiteres Prozessmaß zur Verfügung, das ergänzend zur bewährten Outcome basierten Klassifikation nach Bröder & Schiffer (2003a) eingesetzt werden kann.

Literatur

- Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition*, *93*, B79-B87.
- Baddeley, A. D. (1997). *Human memory: Theory and practice* (überarb. und erw. Aufl.). London: Psychology Press.
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*, 723–767.
- Baron, J. (2000). *Thinking and deciding* (3. Aufl.). Cambridge, MA: Cambridge University 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*, 107–129.
- Bröder, A. (2000a). A methodological comment on behavioral decision research. *Psychologische Beiträge*, *42*, 645–662.
- Bröder, A. (2000b). Assessing the empirical validity of the “Take-The-Best” heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26*, 1332–1346.
- Bröder, A. (2002). Take The Best, Dawes’ Rule, and compensatory decision strategies: A regression-based classification method. *Quality & Quantity*, *36*, 219–238.
- Bröder, A. (2005). *Entscheiden mit der "adaptiven Werkzeugkiste": Ein empirisches Forschungsprogramm*. Lengerich: Pabst.
- Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multiattribute decisions. *Psychonomic Bulletin & Review*, *14*, 895–900.
- Bröder, A., & Newell, B. R. (2008). Challenging some common beliefs: Empirical work within the adaptive toolbox metaphor. *Judgment and Decision Making*, *3*, 205–214.

- Bröder, A., & Schiffer, S. (2003a). Bayesian strategy assessment in multi-attribute decision making. *Journal of Behavioral Decision Making*, *16*, 193–213.
- Bröder, A., & Schiffer, S. (2003b). Take the best versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, *132*, 277–293.
- Bröder, A., & Schiffer, S. (2006). Stimulus format and working memory in fast and frugal strategy selection. *Journal of Behavioral Decision Making*, *19*, 361–380.
- Christensen-Szalanski, J. J. J. (1980). A further examination for the selection of problem solving strategies: The effects of deadlines and analytic aptitudes. *Organizational Behavior and Human Performance*, *25*, 107–122.
- Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571–582.
- Dawes, R. M. (1988). *Rational choice in an uncertain world*. San Diego: Harcourt Brace Jovanovich.
- Dawes, R. M., & Coorigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*, 95–106.
- Dougherty, M. R., Thomas, R., & Franco-Watkins, A. M. (2008). Psychological plausibility of the theory of probabilistic mental models and the fast and frugal heuristics. *Psychological Review*, *115*, 199–213.
- Duchowski, A. (2007). *Eye tracking methodology*. London: Springer.
- Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, *51*, 380–417.
- Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology*, *32*, 53–88.
- Einhorn, H. J., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. *Psychological Review*, *86*, 465–485.

- Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Science*, *12*, 405–410.
- Fiedler, K. (1978). Multiple Regression - Ein Modell der Urteilsbildung? *Zeitschrift für Sozialpsychologie*, *9*, 117–128.
- Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M., & Doherty, M. L. (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes*, *43*, 75–117.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669.
- Gigerenzer, G., Goldstein, D. G., & Hoffrage, U. (2008). Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty, Franco-Watkins, and Thomas (2008). *Psychological Review*, *115*, 230–239.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506–528.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Glöckner, A. (2006). *Die Bewältigung komplexer Entscheidungssituationen mittels automatischer Informationsverarbeitungsprozesse*. Unveröffentlichte Dissertation, Universität Erfurt.
- Glöckner, A., & Betsch, T. (2008). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgment and Decision Making*, *3*, 215–228.
- Goldstein, W. M., & Hogarth, R. M. (1997). *Research on judgment and decision making*. New York: Cambridge University Press.
- Harte, J. M., & Koele, P. (2001). Modeling and describing human judgement processes: The multiattribute evaluation case. *Thinking and Reasoning*, *7*, 29–49.

- Hastie, R., & Dawes, R. M. (2001). *Rational choice in an uncertain world: The psychology of judgment and decision making*. Thousand Oaks, CA: Sage.
- Hausmann, D. (2004). *Informationssuche im Entscheidungsprozess. Die Nützlichkeit von Hinweis-Cues und der Anspruch an Urteilssicherheit*. Dissertationsschrift. Zürich: Zentralstelle der Studentendruckerei.
- Hoffman, P. J. (1960). The paramorphic representation of clinical judgment. *Psychological Bulletin*, 57, 116–131.
- Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 32, 58–69.
- Hoover, M. A., & Richardson, D. C. (2008). When facts go down the rabbit hole: contrasting features and objecthood as indexes to memory. *Cognition*, 108, 533–542.
- Jahn, G., Renkewitz, F., & Kunze, S. (2007). Heuristics in multi-attribute decision making: Effects of representation format. In D. S. McNamara & G. Trafton (Hrsg.), *Proceedings of the Twenty-Ninth Annual Conference on the Cognitive Science Society* (S. 383-388). Mahwah, NJ: Erlbaum.
- Johansson, R., Holsanova, J., & Helmqvist, K. (2006). Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science*, 30, 1053-1079.
- Joos, M., Rötting, M., & Velichowski, B. M. (2003). Die Bewegungen des menschlichen Auges: Fakten, Methoden, innovative Anwendungen. In G. Rickheit, T. Herrmann, & W. Deutsch (Hrsg.), *Psycholinguistik/ Psycholinguistics. Ein internationales Handbuch/ An International Handbook* (S. 142–168). Berlin & NY: de Gruyter.
- Jungermann, H., & Lütge, C. (2009). *Entscheidung und Urteil. Philosophie und Psychologie im Dialog: Vol. 8*. Göttingen: Vandenhoeck & Ruprecht.

- Jungermann, H., Pfister, H.-R., & Fischer, K. (2005). *Die Psychologie der Entscheidung: Eine Einführung* (2. Aufl.). Heidelberg: Spektrum Akademischer Verlag.
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8, 441–480.
- Kahnemann, D., Slovik, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, MA: Cambridge University Press.
- Kinnear, P. R., & Colin, D. G. (2007). *SPSS 14 made simple*. New York: Psychology Press.
- Laplace, P. S. (1814/1986). *Philosophischer Versuch über die Wahrscheinlichkeit*. Leipzig: Akademische Verlagsgesellschaft Geest & Portig.
- Lee, M. D., & Cummins, T. D. R. (2004). Evidence accumulation in decision making: Unifying the “take the best” and the “rational” models. *Psychonomic Bulletin & Review*, 11, 343–352.
- Lohse, G. L., & Johnson, E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes*, 68, 28–43.
- Martignon, L., & Hoffrage, U. (2002). Fast, frugal and fit: Simple heuristics for paired comparison. *Theory and Decision*, 52, 29–71.
- Montgomery, H., & Svenson, O. (1989). A think-aloud study of dominance structuring in decision processes. In H. Montgomery & O. Svenson (Hrsg.), *Process and structure in human decision making* (S. 135–150). Chichester, UK: Wiley.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.
- Newell, B. R., & Bröder, A. (2008). Cognitive processes, models and metaphors in decision research. *Judgment and Decision Making*, 3, 195–204.

- Newell, B. R., Rakow, T., Weston, N. J., & Shanks, D. R. (2004). Search strategies in decision making: The success of “success”. *Journal of Behavioral Decision Making, 17*, 117–137.
- Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing “one-reason” decision making. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*, 53–65.
- Newell, B. R., Weston, N. J., & Shanks, D. R. (2003). Empirical tests of a fast-and-frugal heuristic: Not everyone “takes-the-best”. *Organizational Behavior and Human Decision Processes, 91*, 82–96.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231–259.
- Pachur, T. (2009). Diskussion: Komplexe Modelle des Entscheidens: Eine konstruktive Erweiterung der Bounded Rationality-Perspektive. In: A. B. Eder, K. Rothermund, S. R. Schweinberger, M. C. Steffens, & H. Wiese (Hrsg.), *51. Tagung experimentell arbeitender Psychologen und Psychologinnen in Jena* (S. 77-78). Lengerich: Pabst.
- Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance, 16*, 366–387.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.
- Payne, J. W., Braunstein, M. L., & Carroll, J. S. (1978). Exploring predecisional behavior: An alternative approach to decision research. *Organizational Behavior and Human Performance, 22*, 17–44.
- Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32*, 65–97.

- Renkewitz, F., & Jahn, G. (im Druck). Tracking memory search for cue information. In A. Glöckner & C. Wittemann (Hrsg.), *Foundations for tracing intuitions: Challenges, findings and categorizations*. New York: Psychology Press.
- Richardson, D. C., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: looking at things that aren't there anymore. *Cognition*, *76*, 269–295.
- Riedl, R., Brandstätter, E., & Roithmayr, F. (2008). Identifying decision strategies: A process and outcome-based classification method. *Behavior Research Methods*, *40*, 795–807.
- Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, *127*, 258–276.
- Russo, J. E., & Doshier, B. A. (1983). Strategies for multiattribute binary choices. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *9*, 676–696.
- Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory and Cognition*, *17*, 759–769.
- Sedlmeier, P., & Renkewitz, F. (2008). *Forschungsmethoden und Statistik in der Psychologie*. München: Pearson Studium.
- Shafir, E., & LeBoeuf, R. A. (2002). Rationality. *Annual Review of Psychology*, *53*, 491–517.
- Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin*, *134*, 207–222.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *63*, 129–138.
- Simon, H. A. (1957). *Models of man: Social and rational*. New York: Wiley.
- Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. *Psychological Research*, *65*, 235–241.

- Spivey, M. J., Richardson, D. C., & Fitneva, S. A. (2004). Thinking outside the brain: Spatial indices to visual and linguistic information. In J. M. Henderson & F. Ferreira (Hrsg.), *The interface of vision, language and action* (S. 161-189). New York: Psychology Press.
- Steyer, R., & Eid, M. (2001). *Messen und Testen*. Berlin: Springer.
- Svenson, O. (1979). Process descriptions of decision making. *Organizational Behavior and Human Performance*, 36, 86–112.
- Svenson, O., & Karlsson, G. (1986). Attractiveness of decision alternatives characterized by numerical and non-numerical information. *Scandinavian Journal of Psychology*, 27, 74–84.
- von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.
- Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology*, 60, 53–85.

Anhang

Anhang A: Untersuchungsmaterial der Lernphase

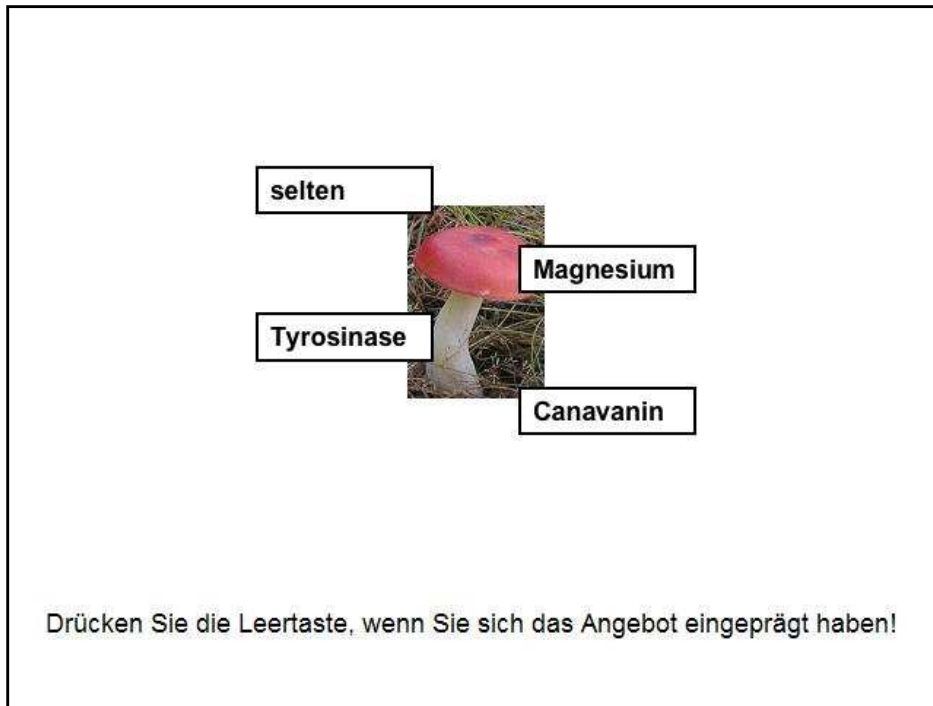


Abbildung A1. Objektpräsentation in der Lernphase

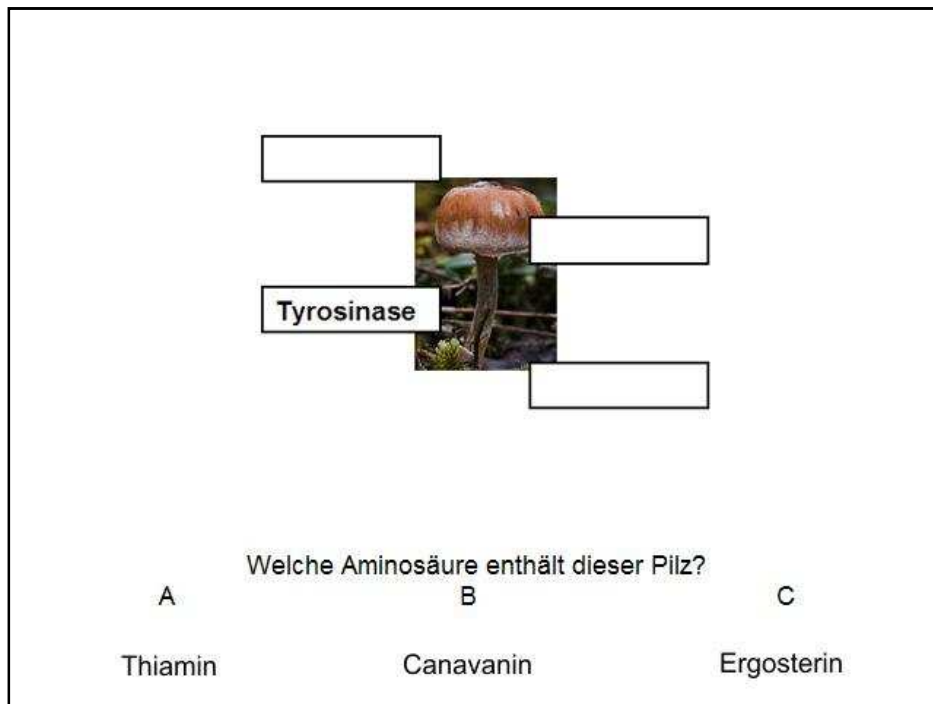


Abbildung A2. Abfrage der Cue-Ausprägungen in der Lernphase

Anhang B: Untersuchungsmaterial der Entscheidungsphase

Der verlässlichste Hinweis auf die Giftigkeit eines Pilzes ist das dominierende Enzym. Tyrosinase-haltige Pilze sind wahrscheinlich giftig, Lakkase- und Peroxydase-haltige Pilze eher nicht.

Die dominierende Aminosäure ist der zweitverlässlichste Hinweis. Ergosterin-haltige Pilze sind wahrscheinlich giftig, Thiamin- und Canavanin-haltige Pilze eher nicht.

Der dominierende Mineralstoff ist der drittverlässlichste Hinweis. Magnesium-haltige Pilze sind wahrscheinlich giftig, Kalium- und Zink-haltige Pilze eher nicht.

Der Verbreitungsgrad ist der viertverlässlichste Hinweis. Seltene Pilze sind wahrscheinlich giftig, mittel verbreitete und häufige Pilze eher nicht.

Der typische Giftpilz sieht also folgendermaßen aus:

Typischer Giftpilz	Tyrosinase
	Ergosterin
	Magnesium
	seltene

Bitte prägen Sie sich diese Information ein!

- Drücken Sie bitte die Leertaste -

Abbildung B1. Instruktion der Entscheidungsrelevanz

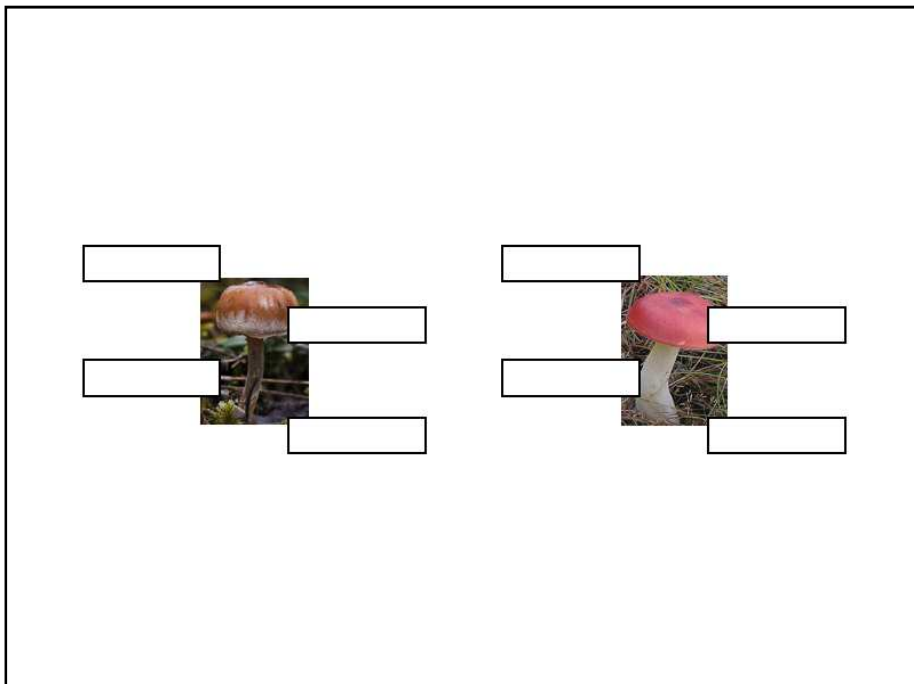


Abbildung B2. Abfragesituation in der Entscheidungsphase

Anhang C: Funktionsweise des Eye Trackers

Zur Erfassung der Augenbewegungen müssen mehrere visuelle Bewegungsmerkmale aufgezeichnet werden (Duchowski, 2007). Der in dieser Untersuchung verwendete Eye Tracker verrechnete die Position der Pupillenmitte mit der Korneareflexion einer in den Eye Tracker eingebauten Infrarotlichtquelle (*Korneareflex-Methode*).

Zum einen werden die in die Retina eintretenden Infrarotlichtstrahlen zu einem Großteil reflektiert. Die Pupille erscheint in der mit einer separaten Software erstellten Videoaufzeichnung als helle, scharf umgrenzte Fläche (*bright pupil Effekt*, Duchowski, 2007). Zum anderen wird das Licht von der Kornea reflektiert (*Korneareflexion* oder *erstes Purkinje-Bild* genannt, Joos, Rötting & Velichkowsky, 2002). Diese Eye Tracking Methode verwendet Infrarotlicht um Blendungen der Versuchsteilnehmer durch „sichtbares“ Licht zu vermeiden. Die Kamera zeichnet die erzeugten Reflexionen auf und verrechnet anschließend die Positionen der Pupillenmitte und der Korneareflexion miteinander. Eine ausführliche Beschreibung der verwendeten Berechnungsmethoden findet sich bei Duchowski (2007). Die dahinterstehende Logik ist wie folgt: Bewegt eine Person ihre Augen, verändert sich die Position der beiden Parameter zueinander. Bewegt sie dagegen nur den Kopf, bleibt die Differenz konstant. Durch diesen Mechanismus werden zum einen Kopfbewegungen von Bewegungen des Auges unterschieden. Zum anderen wird durch die relative Position der beiden Blickparameter zueinander der Blickfokus bestimmt. Dazu muss am Beginn der Aufzeichnung die Korneareflexion relativ zur Pupillenmitte an kritischen Punkten festgehalten werden. Dieser Vorgang wird als *Kalibration* bezeichnet. Die Stellung der Korneareflexion zur Pupillenmitte an den neun Kalibrationspunkten ist in Abbildung C1 dargestellt.

Das in dieser Untersuchung verwendete Eye Tracking System zeichnete die Blickbewegungen mit einer Genauigkeit von $0,5^\circ$ Sehwinkel auf. Die zeitliche Auflösung der Eye Tracking Kamera betrug 50 Hz pro Minute.

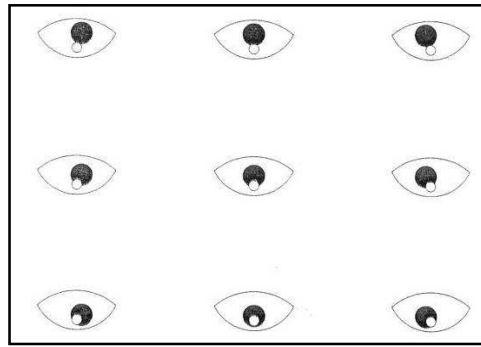


Abbildung C1. Position der Pupille in Relation zur Korneareflexion (Duchowski, 2007, S.58)

Anhang D: Areas of Interest

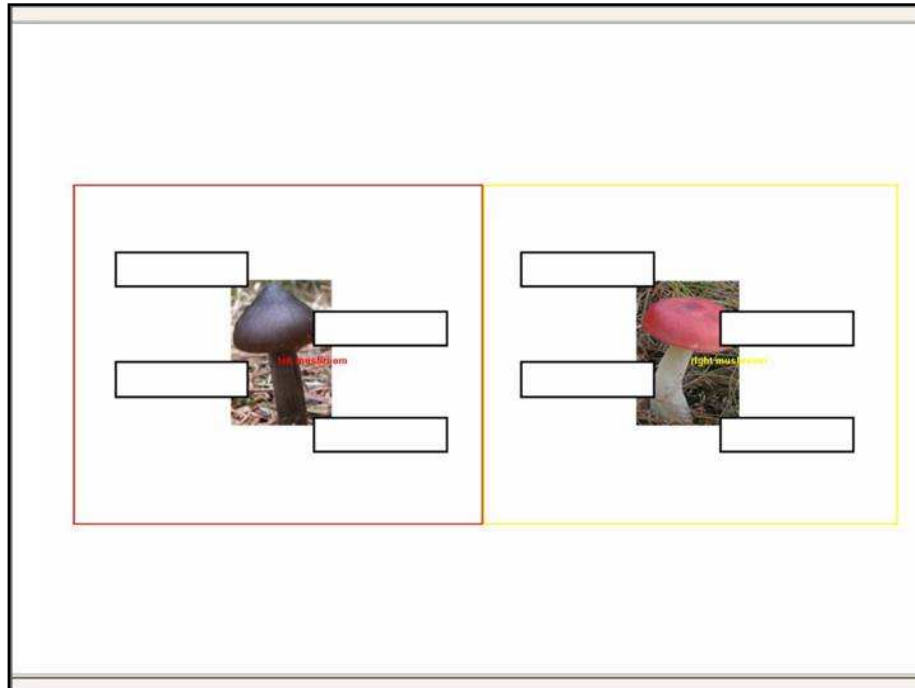


Abbildung D1. Areas of Interest (AOIs) über dem rechten und linken Stimuli eines Entscheidungsdurchgangs im Auswertungsmodus