Das Streben nach Konsistenz im Entscheidungsprozess

Eine Untersuchung möglicher Einflussfaktoren

Dissertation

zur Erlangung des akademischen Grades
eines Doktors der Philosophie (Dr. phil.)

der
Erziehungswissenschaftlichen Fakultät
der Universität Erfurt

vorgelegt von
Tanja Ostermann

Erfurt 2010

Erstes Gutachten: Prof. Dr. Tilmann Betsch (Universität Erfurt)
Zweites Gutachten: Prof. Dr. Ralf Rummer (Universität Erfurt)
Drittes Gutachten: Prof. Dr. Arndt Bröder (Universität Bonn)

Tag der Disputation: 02.07.2010
Datum der Promotion: 02.07.2010

urn:nbn:de:gbv:547-201000592

Zusammenfassung

Das Ziel der vorliegenden Arbeit bestand darin, das Verständnis des Konsistenzmaximierungsprozesses in probabilistischen Inferenzentscheidungen durch die gezielte Untersuchung potentieller Einflussfaktoren zu vertiefen. Das Streben nach Konsistenz und seine Untersuchung wurden im theoretischen Teil der Arbeit zunächst allgemein betrachtet. Hierbei wurde die historische Entwicklung von der Gestaltpsychologie zu den *Parallel Constraint Satisfaction* (PCS)-Ansätzen aufgezeigt. Daraufhin wurde das Thema Konsistenz spezifisch für den Kontext der Entscheidungspsychologie, auf den in dieser Arbeit fokussiert wurde, betrachtet. Es wurden Theorien und Modelle vorgestellt, in denen der Aspekt der Erhöhung von Konsistenz zentral ist. Besondere Aufmerksamkeit wurde dem PCS-Modell für probabilistische Inferenzentscheidungen von Glöckner und Betsch (2008b) gewidmet, in dem der Prozess der Konsistenzmaximierung den Kernmechanismus eines jeden Entscheidungsprozesses darstellt. Anschließend wurden Befunde empirischer Arbeiten, in denen die Maximierung von Konsistenz als systematische Reevaluation von Informationen im Entscheidungsprozess erfasst wurde, vorgestellt. Anhand dieser Befunde wurde dargelegt, dass die Erhöhung von Konsistenz als wichtiger Bestandteil des Informationsverarbeitungsprozesses verstanden werden kann und die Suche nach Faktoren, die einen Einfluss auf die Stärke prädezisionaler Informationsumwertungen haben, erforderlich ist. Im Anschluss an diese theoretischen Ausführungen wurde zum empirischen Teil der Arbeit übergeleitet: Auf Basis des PCS-Modells für probabilistische Inferenzentscheidungen wurden Charakteristika der Situation und der Person, die einen Einfluss auf die Stärke der prädezisionalen Umwertung von Cue-Validitäten haben können, untersucht. Hierfür wurden insgesamt sieben Experimente und drei Simulationsstudien durchgeführt. Im ersten Experiment konnte unter Verwendung eines Solomon-Viergruppenplans der wichtige Nachweis erbracht werden, dass prädezisionale Reevaluationen in probabilistischen Inferenzentscheidungen auf die Reflexion über eine Entscheidungsaufgabe zurückgeführt werden können. Im folgenden zweiten Experiment wurde untersucht, ob systematische Reevaluationen von Cue-Validitäten auch bei einer ausgeglichenen Informationslage auftreten und welchen Einfluss die wiederholte Darbietung einer Entscheidungsaufgabe hat. Das erwartungskonträre Ergebnis, dass keine Umwertungen bei einer ausgeglichenen Informationslage aufgetreten sind, gab Anlass zu einer systematischen Untersuchung des Einflusses der Informationslage.

In den Experimenten 3 und 4 konnte übereinstimmend gefunden werden, dass systematische Reevaluationen lediglich bei einer unausgeglichenen Informationslage auftreten. Neben dem Einfluss der Informationslage wurden auch personale Faktoren hinsichtlich ihres Einflusses auf die Stärke prädezisionaler Informationsumwertungen untersucht. In den Experimenten 4 und 5 wurden verschiedene personale Faktoren berücksichtigt. Allerdings fand sich für keinen der Faktoren ein überzeugender Zusammenhang. Drei Simulationsstudien sollten zur Klärung der Frage, wie gut das PCS-Modell die empirischen Ergebnisse vorhersagen kann, beitragen. Hierfür wurden Simulationen auf individueller Ebene durchgeführt. Ein wichtiger Befund war, dass bei ausgeglichenen im Gegensatz zu unausgeglichenen Informationslagen die Simulationsergebnisse keinen signifikanten Beitrag zur Vorhersage der Posttest-Validitäten machten, wenn für den Einfluss der Pretest-Validitäten kontrolliert wurde. In zwei weiteren Experimenten wurde erneut der Einfluss der Wiederholung untersucht. Während eine reine Wiederholung einer identisch konzipierten Entscheidungssituation in einer überwiegenden Abwertung von Cues resultierte, führte der Erhalt einer zusätzlichen entscheidenden Information eher zu einer systematischen Umwertung von Cues. Zusammenfassend war festzuhalten, dass die Annahme bidirektionaler Prozesse in der Entscheidungsfindung gestützt werden konnte. Allerdings waren weder die größtenteils moderate Evidenz für systematische Umwertungen noch das Ausbleiben selbiger bei ausgeglichenen Informationslagen nach dem PCS-Modell zu erwarten gewesen. Vor dem Hintergrund bereits bestehender sowie in dieser Arbeit erzielter Befunde wurde die Vermutung aufgestellt, dass es gewisser Mindestvoraussetzungen für das erfolgreiche Wirken des Konsistenzmaximierungsprozesses bedarf.

Schlagworte:

Konsistenz
Informationsverzerrung
Entscheidungsforschung
Parallel constraint satisfaction

Abstract

The aim of this thesis was to gain a deeper understanding of the consistency maximizing process in probabilistic inference tasks by analyzing potential influencing factors. Concerning the theoretical background the striving for consistency and its analysis were first considered in general. The development from Gestalt Psychology to parallel constraint satisfaction (pcs) models was demonstrated. After that the question of consistency was considered specifically for the domain of decision making on which a focus was put in this thesis. Theories and models were presented in which the increase of consistency is central. Special attention was paid to the pcs model of Glöckner and Betsch (2008b) in which the process of consistency maximizing is regarded as the core mechanism of decision making. Subsequently results of empirical studies in which the maximization of consistency was measured as a systematic reevaluation of information during decision making were presented. On the basis of these results it was argued that the increase of consistency can be regarded as a central part of information processing and that the search for factors which could influence the strength of the predecisional reevaluation is necessary. Following the theoretical considerations the empirical studies were presented: On the basis of the pcs model for probabilistic inferences characteristics of the situation and of the person which might influence the strength of predecisional reevaluation of cue validities were analysed. Seven experiments and three simulation studies were therefore conducted. Using a Solomon four group design it was shown in the first experiment that predecisional information distortion is due to the reflection on a decision task. In the following second experiment it was analysed if the systematic reevaluation of cue validities also occurs using a balanced information pattern and which influence the repetition of a decision task has. The unexpected result that no reevaluation occured using a balanced information pattern gave reason to a systematic analysis of the influence of the information pattern. In experiments 3 and 4 it was shown that systematic reevaluation only occurs using an unbalanced information pattern. Besides the influence of the information pattern also personal factors were examined concerning their influence on the strength of predecision information distortion. In experiments 3 and 4 several personal factors were considered but for none a convincing relation was found. Three simulation studies were conducted to clarify whether the pcs model can predict the empirical results. Therefore simulation studies on an individual level

were conducted. An important result was that with balanced in contrast to unbalanced information patterns the simulation results don't contribute significantly to the prediction of the posttest validities when controlling for the influence of pretest validities. In two further experiments the influence of repetition was again examined. While the mere repetition of an identically designed decision task resulted in a predominant devaluation of cues, the receipt of a further important information lead to a more systematic reevaluation of cues. In summary, the assumption of bidirectional processing in decision making was supported. However, neither the mostly weak evidence for systematic reevaluation nor their absence in balanced information patterns was expected according to the pcs model. Against the background of existing and of the results achieved in the presented studies it was argued that some minimum conditions are necessary for the consistency maximizing process to be successful.

Keywords:

Consistency
Information distortion
Decision making
Parallel constraint satisfaction

Danksagung

Diese Dissertation habe ich im Rahmen eines Promotionsstipendiums am Max-Planck-Institut zur Erforschung von Gemeinschaftsgütern als Mitglied der von Andreas Glöckner geleiteten Arbeitsgruppe *Intuitive Experts* angefertigt. Zahlreiche Personen haben mich während der Arbeit an dieser Dissertation begleitet und unterstützt. Bei ihnen möchte ich mich an dieser Stelle von ganzem Herzen bedanken.

Andreas Glöckner möchte ich besonders herzlich dafür danken, dass er die Realisierung meines eigenen Forschungsvorhabens im Rahmen des Promotionsstipendiums erst ermöglichte und mich in dieser Zeit unterstützte. Seine Hinweise und Anregungen waren für mich stets hilfreich. Bei Tilmann Betsch möchte ich mich insbesondere dafür bedanken, dass er sich gerne zur Betreuung dieser Arbeit bereit erklärt hat und durch sein Interesse an dieser Arbeit zu meiner Motivation beitrug. Mein ausdrücklicher Dank gilt zudem Arndt Bröder, mit dem ich anregende Gespräche führen durfte und der mir die Möglichkeit zum wissenschaftlichen Austausch in seinem Diplomanden- und Doktorandenseminar gab.
Des Weiteren gilt mein Dank den studentischen Hilfskräften: Sebastian Schier und Darius Wieczorek danke ich für die zuverlässige Übernahme der Programmierungsarbeiten. Bei Susann Fiedler, Maria Gladigau, Lena Heuel, Janet Kleber, Esther Kordecki, Marie Landsberg, Maria Rische und Kristina Schönfeldt möchte ich mich herzlich für ihre Unterstützung bei der Datenerhebung und zum Teil Dateneingabe bedanken.
Ausdrücklich möchte ich auch den Mitarbeitern des Max-Planck-Instituts zur Erforschung von Gemeinschaftsgütern für ihre freundliche Unterstützung in organisatorischen Fragen danken.
David Hoffmann und Christian Werner danke ich für das umsichtige Korrekturlesen dieser Arbeit.

Mein besonderer Dank gilt meinen Eltern, Angelika und Heinz Ostermann, die mich bei allen Entscheidungen, die ich treffe, stets mit ganzen Kräften unterstützen.
Milena Hartmann bin ich besonders für ihre moralische Unterstützung dankbar.
Aus tiefem Herzen danke ich Fabian Kurz, der mich mit seiner Geduld, seinem Humor und seinem Glauben an mich während der Arbeit an dieser Dissertation unterstützt hat.

Inhaltsverzeichnis

1. EINLEITUNG 12

2. THEORETISCHER TEIL 15

2.1 KONSISTENZ – ZENTRALE CHARAKTERISTIKA 15

2.2 DIE ERFORSCHUNG DER KONSISTENZ 17

2.2.1 Die Gestaltpsychologie als Ursprung der Konsistenzforschung 17

2.2.2 Kognitive Konsistenztheorien 19

2.2.3 Parallel constraint satisfaction-Netzwerke 32

2.3 DAS STREBEN NACH KONSISTENZ: ENTSCHEIDUNGSTHEORIEN UND -MODELLE 44

2.3.1 Bisherige Ansätze 44

2.3.2 Das parallel constraint satisfaction-Modell von Glöckner und Betsch (2008) 54

2.4 BEFUNDE ZUM STREBEN NACH KONSISTENZ IN DER ENTSCHEIDUNGSFORSCHUNG . 61

2.4.1 Prädezisionale Informationsverzerrung 61

2.4.2 Charakteristika der Aufgabe 67

2.4.3 Charakteristika der Person 73

2.5 ZUSAMMENFASSUNG UND SCHLUSSFOLGERUNG 77

3. EMPIRISCHER TEIL 82

3.1 ÜBERBLICK ÜBER DEN EMPIRISCHEN TEIL 82

3.2 METHODISCHE VORÜBERLEGUNGEN 83

3.2.1 Die Messung subjektiver Cue-Validitäten 83

3.2.2 Die Auswertung der Veränderung subjektiver Cue-Validitäten 87

3.3 EXPERIMENT 1: DER EINFLUSS DES PRETESTS 89

3.4 EXPERIMENT 2: DER EINFLUSS DER WIEDERHOLUNG BEI EINER AUSGEGLICHENEN INFORMATIONSLAGE 97

3.5 DER EINFLUSS DES INFORMATIONSMUSTERS UND CHARAKTERISTIKA DER PERSON AUF DIE VERÄNDERUNG SUBJEKTIVER CUE-VALIDITÄTEN 104

3.5.1 Experiment 3: Die Veränderung subjektiver Cue-Validitäten bei ausgeglichenen und unausgeglichenen Informationslagen 104

3.5.2 Experiment 4: Die Veränderung subjektiver Cue-Validitäten in Abhängigkeit von dem Informationsmuster und Charakteristika der Person 113

3.5.3 Experiment 5: Charakteristika der Person und die Veränderung subjektiver Cue-Validitäten 124

3.5.4 Zusammenfassung und Diskussion 133

3.6 SIMULATIONSSTUDIEN 138

3.6.1 Methodische Vorüberlegungen 138

3.6.2 Simulationsstudie 1: Die Vorhersage der Posttest-Valditätseinschätzungen aus Experiment 3 143

3.6.3 Simulationsstudie 2: Die Vorhersage der Posttest-Validitätseinschätzungen aus Experiment 4 148

3.6.4 Simulationsstudie 3: Die Vorhersage der Posttest-Validitätseinschätzungen aus Experiment 5 154

3.6.5 Zusammenfassung und Diskussion 158

3.7 DER EINFLUSS DER WIEDERHOLUNG DER ENTSCHEIDUNGSSITUATION AUF DIE VERÄNDERUNG SUBJEKTIVER CUE-VALIDITÄTEN 160

3.7.1 Experiment 6: Die Veränderung von Cue-Validitäten bei wiederholter im Vergleich zu einmaliger Darbietung der Entscheidungssituation 161

3.7.2 Experiment 7: Der Einfluss entscheidender Informationen für die Veränderung von Cue-Validitäten nach wiederholten Entscheidungssituationen 175

3.7.3 Zusammenfassung und Diskussion 190

4. DISKUSSION 194

4.1 ZUSAMMENFASSENDE DARSTELLUNG DER EMPIRISCHEN BEFUNDE 195

4.1.1 Der Einfluss des Pretests 196

4.1.2 Charakteristika der Situation 197

4.1.3 Charakteristika der Person 201

4.1.4 Vorhersage der Ergebnisse durch das PCS-Modell 203

4.1.5 Kritische Aspekte 207

4.2 ALLGEMEINE DISKUSSIONSPUNKTE UND OFFENE FRAGEN 208

4.2.1 Die Grenzen der Konsistenzmaximierung? 208

4.2.2 Vor- und Nachteile der Konsistenzmaximierung 211

4.2.3 Methoden des (De-)Biasing 213

4.3 IMPLIKATIONEN UND PERSPEKTIVEN 214

4.3.1 Für die Entscheidungsforschung 214

4.3.2 Für die Konsistenzforschung 219

4.3.3 Für praktische Kontexte 220

4.4 SCHLUSSBETRACHTUNG 222

LITERATURVERZEICHNIS 225

ABBILDUNGSVERZEICHNIS 243

TABELLENVERZEICHNIS 244

ANHANG 246

ANHANG A: MATERIAL UND INSTRUKTIONEN 246

ANHANG B: DESKRIPTIVER ÜBERBLICK ÜBER DIE VALIDITÄTSEINSCHÄTZUNGEN 286

ANHANG C: WEITERE ERGEBNISSE 295

1. EINLEITUNG

Stellen Sie sich vor, Sie möchten in eine neue Wohnung umziehen. Sie stehen vor der Wahl zwischen zwei potenziellen Wohnungen A und B, deren Miete etwa gleich hoch ist. Neben der Höhe der Miete sind Ihnen folgende Charakteristika besonders wichtig: der Zustand der Wohnung, die Wohnlage, die Verkehrsanbindung und die Entfernung zu Ihrem Arbeitsplatz. Während sich Wohnung A durch eine attraktive Wohnlage und eine gute Verkehrsanbindung auszeichnet, sind der renovierungsbedürftige Zustand der Wohnung sowie die große Entfernung zu Ihrem Arbeitsplatz eher ungünstig. Wohnung B hingegen ist zwar in einem hervorragenden Zustand und nur wenige Meter von Ihrem Arbeitsplatz entfernt, liegt jedoch in einem weniger attraktiven Viertel, das mit öffentlichen Verkehrsmitteln zudem vergleichsweise schwer zu erreichen ist. Nehmen wir nun an, dass Sie sich für Wohnung A entscheiden. Wie sind Sie zu dieser Entscheidung gekommen? Vermutlich können Sie nachträglich Gründe aufführen, die Ihre Entscheidung als richtig rechtfertigen: Gerade die Wohnlage und die Verkehrsanbindung erscheinen Ihnen nun besonders wichtig, da Sie Möglichkeiten der Freizeitgestaltung eröffnen. Die Renovierungsarbeiten betrachten Sie jetzt als Chance, die Wohnung nach Ihren Vorstellungen zu gestalten. In dem längerem Weg zum Arbeitsplatz und zurück sehen Sie nun die Möglichkeit, zwischen Freizeit und Arbeit zu trennen und abzuschalten. Auf diesem Wege rechtfertigen Sie Ihre Entscheidung vor sich und vor anderen. Doch wie insbesondere die Forschung der letzten 20 Jahre gezeigt hat, können diese Umwertungen von Informationen auch bereits vor der eigentlichen Entscheidung stattfinden. Dadurch wird eine *konsistente* Entscheidungssituation konstruiert, die es Ihnen ermöglicht, trotz widersprüchlicher Informationen eine konfidente Entscheidung zu treffen. Dieses Prozesses sind Sie sich nicht bewusst, er ist aber maßgeblich für Ihre Entscheidungsfindung.

Das in diesem Beispiel angesprochene *Streben nach Konsistenz* kam bereits in der Gestaltpsychologie, hier in der *Tendenz zur guten Gestalt* (Wertheimer, 1923), zum Ausdruck. Die nachfolgenden kognitiven Konsistenztheorien (z.B. Abelson und Rosenberg, 1958; Heider, 1946; Heider, 1958; Osgood und Tannenbaum, 1955) erhoben dieses Phänomen zum zentralen Untersuchungsgegenstand, methodische und theoretische Unzulänglichkeiten führten allerdings zu einem Stillstand der Konsistenzforschung. Erst die sogenannten *parallel constraint satisfaction* (PCS)-Ansätze (z.B. McClelland und

Rumelhart, 1981; Read und Miller, 1998a; Thagard, 1989) überwanden diese Schwächen und konnten der Dynamik und Komplexität der menschlichen Informationsverarbeitung gerecht werden. In der Entscheidungsforschung wurde das Thema der Konsistenz zunächst lediglich in Form der postdezisionalen Dissonanzreduktion (Festinger, 1957; Festinger, 1964) betrachtet. Die Erkenntnis, dass derartige Umwertungen auch in der prädezisionalen Phase auftreten, bedurfte hingegen einer gewissen Zeit, ehe sie sich durchsetzte. Brownstein (2003) trug schließlich in einem Review überzeugende Evidenz dafür zusammen, dass Informationsumwertungen prädezisional auftreten. Damit wurde der Fokus auch von der Vermeidung von Dissonanz auf das Streben nach Konsistenz gelegt. In der Folgezeit konnte dann das Verständnis für dieses Phänomen, ausgedrückt in prädezisionalen Informationsverzerrungen, vertieft werden. Hierzu trugen auf der einen Seite gezielte Forschungsarbeiten bei, die ergaben, dass diese Umwertungen sowohl nahezu universell als auch in ihrer Stärke durch verschiedene Faktoren, situationaler wie auch personaler Art, moderierbar sind. Auf der anderen Seite konnten auch theoretische Fortschritte erzielt werden. So haben sich auch in der Entscheidungsforschung PCS-Ansätze als sinnvoll erwiesen, da sie die konsistenzgeleitete Dynamik der menschlichen Informationsverarbeitung abbilden können. Ein PCS-Modell, das den Prozess der Konsistenzmaximierung als zentralen Mechanismus eines jeden Entscheidungsprozesses versteht, haben Glöckner und Betsch (2008b) aufgestellt. Sie spezifizieren ihr Modell für sogenannte probabilistische Inferenzentscheidungen. Glöckner und Betsch grenzen sich nicht nur von prominenten Entscheidungsansätzen, die eine adaptive Wahl zwischen mehreren Strategien postulieren (z.B. Gigerenzer, Todd & the ABC Research Group, 1999; Payne, Bettman & Johnson, 1988), ab, sondern sie erheben das Streben nach Konsistenz zum zentralen Prinzip und können den zugrunde liegenden Prozess formal darstellen.

Die Zielsetzung der vorliegenden Arbeit besteht darin, das Verständnis des Prozesses der Konsistenzmaximierung zu vertiefen und damit einen Beitrag zu der Erforschung des Strebens nach Konsistenz in der Entscheidungsforschung zu leisten. Vor dem Hintergrund des darzustellenden Forschungsstandes und auf Basis des PCS-Modells von Glöckner und Betsch (2008b) werden dabei sowohl Charakteristika der Situation als auch der Person hinsichtlich ihres Einflusses auf die Stärke prädezisionaler Informationsverzerrungen in probabilistischen Inferenzentscheidungen untersucht.

Nach einer ersten Klärung des Begriffs der Konsistenz (Kapitel 2.1) wird ein historischer Überblick über die Konsistenzforschung gegeben (Kapitel 2.2). Hierbei wird

aufgezeigt, wie in der Entwicklung von der Gestaltpsychologie über die kognitiven Konsistenztheorien zu den PCS-Ansätzen nicht nur theoretische und methodische Unzulänglichkeiten überwunden werden konnten, sondern wie sich dadurch auch das Verständnis des Strebens nach Konsistenz erweitert hat. Im anschließenden Kapitel 2.3 wird eine Fokussierung auf den Bereich der Entscheidungsforschung vorgenommen. Es werden Ansätze beleuchtet, die den konstruktiven Charakter des Entscheidungsprozesses hervorheben. Auch in diesem Kapitel wird der Mehrwert der PCS-Ansätze unterstrichen. Das PCS-Modell von Glöckner und Betsch (2008b) wird ausführlich dargestellt und es wird begründet, weshalb es die Grundlage der vorliegenden Arbeit bildet. Die Vorstellung des Forschungsstandes in Kapitel 2.4 vermittelt eine aktuelle Charakterisierung prädezisionaler Informationsverzerrungen als Ausdruck des Konsistenzstrebens. Eine Zusammenfassung bildet den Abschluss des theoretischen Teils der Arbeit (Kapitel 2.5). Auf Basis der Kernaspekte werden eine Arbeitsdefinition für Konsistenz vorgeschlagen und Schlussfolgerungen für die Forschung gezogen, die zum empirischen Teil dieser Arbeit überleiten. Nach einem Überblick über diesen Teil (Kapitel 3.1) sowie methodische Vorüberlegungen, in denen die bei der Messung und Auswertung der Daten vorgenommene Vorgehensweise erläutert wird (Kapitel 3.2), werden die einzelnen Experimente vorgestellt. Ein Solomon-Viergruppenplan bildet den Anfang, da er durch die kritische Prüfung des Untersuchungsdesigns zentral für die folgenden Experimente ist (Kapitel 3.3). Im folgenden Experiment wird unter Verwendung einer ausgeglichenen Informationslage der Einfluss der Wiederholung untersucht (Kapitel 3.4). In Kapitel 3.5 wird die Bedeutung der Ausgeglichenheit der Informationslage gezielt erforscht. Des Weiteren wird dem Einfluss personaler Faktoren nachgegangen. Ergänzt und erweitert werden die Erkenntnisse durch Simulationsstudien (Kapitel 3.6). Schließlich wird der Einfluss der Wiederholung einer Entscheidungssituation gezielt untersucht (Kapitel 3.7). In einer abschließenden Diskussion in Kapitel 4 werden die Befunde zunächst zusammengefasst und vor dem Hintergrund der theoretischen Grundlagen interpretiert (Kapitel 4.1). Des Weiteren werden allgemeine Diskussionspunkte und offene Fragen erörtert (Kapitel 4.2) sowie Implikationen und Perspektiven für Forschung und Praxis aufgezeigt (Kapitel 4.3). Eine Schlussbetrachtung rundet diese Arbeit ab (Kapitel 4.4).

2. THEORETISCHER TEIL

Im folgenden Kapitel wird die theoretische Basis dieser Forschungsarbeit vorgestellt: Die allgemeine Grundlage bildet die Konsistenzforschung, die dann spezifisch für den Bereich der Entscheidungspsychologie betrachtet wird. Zunächst wird in Kapitel 2.1 das dieser Arbeit zugrunde liegende Verständnis von Konsistenz erläutert, indem wichtige Charakteristika herausgearbeitet werden. Anschließend werden theoretische Ansätze, in denen das Bedürfnis nach Konsistenz thematisiert wurde, vorgestellt (Kapitel 2.2). Im folgenden Kapitel 2.3 findet eine Fokussierung auf den für die vorliegende Arbeit zentralen Bereich statt: die Entscheidungsforschung. Es werden Theorien und Modelle vorgestellt, in denen der Aspekt der Konsistenz eine zentrale Stellung einnimmt. Empirische Befunde, die das Streben nach Konsistenz im Bereich der Entscheidungen aufzeigen (Kapitel 2.4), runden die Darstellung ab. Abschließend werden in Kapitel 2.5 die Kernaspekte zusammengefasst. Basierend auf den vorherigen Ausführungen werden eine Arbeitsdefinition für Konsistenz vorgeschlagen und die Fragestellung dieser Arbeit präzisiert, die zum empirischen Teil überleitet.

2.1 KONSISTENZ – ZENTRALE CHARAKTERISTIKA

Zentrale Kennzeichen von *Konsistenz* sind „Einheitlichkeit, Zusammenhang und Widerspruchsfreiheit“ (Fremdwörterbuch, o.J.). Diese Charakterisierung bildet auch die gemeinsame Basis der Arbeiten der psychologischen Konsistenzforschung. Allerdings werden die Merkmale nicht durchgängig unter dem Begriff der Konsistenz subsumiert, sondern ebenfalls unter Begriffen wie Balance (Heider, 1946; Heider, 1958), Harmonie (Rumelhart, Smolensky, McClelland & Hinton, 1986), Kohärenz (z.B. Holyoak und Simon, 1999), Konsonanz (Festinger, 1957; Festinger, 1964; Shultz und Lepper, 1996) und Organisation (z.B. Read, Vanman & Miller, 1997). Da der Begriff *Konsistenz* die Nähe zu den kognitiven Konsistenztheorien und damit auf der einen Seite auch zu der Gestaltpsychologie und auf der anderen Seite zu den PCS-Ansätzen am besten ausdrückt, wird er in der vorliegenden Arbeit präferiert. Angesichts der verbreiteten Verwendung dieses Begriffes ist es notwendig, neben den Grundelementen weitere Charakteristika herauszuarbeiten, um das dieser Arbeit zugrunde liegende Verständnis von Konsistenz zu verdeutlichen. Neben einer

begrifflichen Eingrenzung werden hierbei bereits zentrale Merkmale, die sich aus der Konsistenzforschung ableiten lassen (vgl. Kapitel 2.2) und im Folgenden näher ausgeführt werden, zusammenfassend vorgestellt.

In der vorliegenden Arbeit wird stets die kognitive Konsistenz verschiedener Elemente in einer bestimmten Situation betrachtet. Folglich wird Konsistenz nicht etwa hinsichtlich der Frage der Kontinuität, z.B. hinsichtlich der Konsistenz des Verhaltens, analysiert. Es ist davon auszugehen, dass Konsistenz im Gegensatz zu Inkonsistenz angenehm und damit anziehend ist. Folglich kann von einem *Bedürfnis* bzw. einem *Streben*[1] nach Konsistenz gesprochen werden, was wiederum bereits impliziert, dass der zu diesem Zustand führende Prozess dynamisch ist. Angesichts dessen ist die Annahme, dass sich die Struktur eines Systems sowie die Gewichtung und Interpretation einzelner Elemente in Richtung Konsistenz verändern können, nahe liegend. Hierbei ist als wichtiges Charakteristikum zu nennen, dass Konsistenz sowohl durch die Restrukturierung und Reevaluation widersprechender als auch stützender Informationen erzielt werden kann. Die Tatsache, dass Konsistenz als erstrebenswerter Zustand aufgefasst werden kann, schließt nicht aus, dass sie noch weiteren Zielen dienen kann. Da in dieser Arbeit aber auf den Prozess der Konstruktion von Konsistenz fokussiert wird, werden weitere Ziele nicht berücksichtigt.

Diese Charakterisierung erlaubt eine erste, notwendige begriffliche Klärung und grenzt die Betrachtung von Konsistenz und des zu diesem Zustand führenden Prozesses ein. Die Ausführungen der folgenden Kapitel 2.2, 2.3 und 2.4 werden das Verständnis von Konsistenz weiter vertiefen, so dass auf Basis dieser Erkenntnisse eine Arbeitsdefinition (Kapitel 2.5) herausgearbeitet werden kann. Als Ausgangsbasis lässt sich folgende Umschreibung von Konsistenz festhalten:

Konsistenz bezeichnet einen angenehmen Zustand der Einheitlichkeit, des Zusammenhangs und der Widerspruchsfreiheit von Elementen, der durch Restrukturierungen und Reevaluationen dieser Elemente angestrebt wird.

[1] *Streben* wird hier nicht in einem motivationalen Sinne verstanden, sondern als generelle Tendenz eines Systems, einen Zustand höherer Konsistenz zu erreichen.

2.2 DIE ERFORSCHUNG DER KONSISTENZ

Im nachfolgenden Kapitel wird ein Überblick über die psychologische Konsistenzforschung gegeben. Es werden die theoretischen Ansätze, die wesentliche Beiträge zu der Frage der Konsistenz geleistet haben und somit für die vorliegende Arbeit von besonderer Relevanz sind, vorgestellt. Zunächst wird auf die Bedeutung der Gestaltpsychologie eingegangen (Kapitel 2.2.1), woran sich eine kurze Erörterung der wichtigsten Konsistenztheorien anschließt (Kapitel 2.2.2). Aus den Stärken und Schwächen des gestaltpsychologischen Ansatzes sowie aus der weiterhin bestehenden Relevanz des Themas Konsistenz resultierte die Entwicklung von PCS-Netzwerken, die einen neuen Zugang zur Frage der Konsistenz ermöglichten und Gegenstand von Kapitel 2.2.3 sind.

2.2.1 Die Gestaltpsychologie als Ursprung der Konsistenzforschung

Zwei bedeutende Elemente des gestaltpsychologischen Ansatzes qualifizieren diesen besonders als Ursprung der Konsistenzforschung: Erstens rückt das Ganze mit seinen emergenten Eigenschaften in den Mittelpunkt der Betrachtung und zweitens wird eine Tendenz zu guten Gestalten angenommen (Wertheimer, 1924). Die Tatsache, dass die kognitiven Konsistenztheorien sowie insbesondere die PCS-Ansätze auf der Gestaltpsychologie aufbauen (Read et al., 1997; Simon, Snow & Read, 2004), unterstreicht deren Bedeutung für die Konsistenzforschung im Allgemeinen. Während eine Thematisierung der Konsistenztheorien und PCS-Netzwerke in den beiden folgenden Kapiteln 2.2.2 und 2.2.3 stattfindet, werden im vorliegenden Kapitel die Grundzüge der Gestaltpsychologie[2] nachgezeichnet. Angesichts der Breite dieser Richtung ist eine Fokussierung auf die für die vorliegende Arbeit relevanten Aspekte hierbei unerlässlich.

Die Gestaltpsychologie nahm ihren Anfang bei Wertheimers Arbeiten über das Sehen und damit im Bereich der Wahrnehmung (Metzger, 1966). Wertheimer (1924) verweist selbst auf die Bedeutung des Philosophen Christian von Ehrenfels bei der Begründung dieser psychologischen Richtung. Von Ehrenfels (1890) legte unter Bezugnahme auf eine Arbeit von Mach dar, dass eine bestimmte Summe an Elementen durch ein weiteres Element, nämlich durch die sogenannte Gestaltqualität, ergänzt wird. Diese sei die Voraussetzung für

[2] Entsprechend der allgemeinen Auffassung wird unter Gestaltpsychologie die Richtung der Berliner Schule verstanden. Neben dem Begründer Max Wertheimer sind vor allem Wolfgang Köhler, Kurt Koffka und Kurt Lewin als namhafte Vertreter zu nennen (Häcker und Stapf, 2009).

das Wiedererkennen von z.B. Melodien, Figuren und Bewegungen. Die Gestaltpsychologie grenzte sich von der Annahme des Hinzutretens eines neuen Elements – der Gestaltqualität – ab und nahm stattdessen an, dass für die Wahrnehmung eines Elementes das Ganze entscheidend ist (Wertheimer, 1924). Das einzelne Element sei stets als Teil des Ganzen zu betrachten. Dass dieses Ganze eine über die einzelnen Elemente hinausgehende Qualität hat, wird eindrücklich bei Koffka (1935) deutlich:

> It has been said: The whole is more than the sum of its parts. It is more correct to say that the whole is something else than the sum of its parts, because summing is a meaningless procedure, whereas the whole-part relationship is meaningful. (S. 176)

In diesem Zusammenhang ist zudem die Betonung der Wirksamkeit dynamischer Kräfte wichtig (z.B. Köhler, 1947). Die wichtigsten Prinzipien bzw. Faktoren der Gestaltpsychologie benannte Wertheimer (1923). Hierzu zählen die Faktoren der Nähe, der Gleichheit, des gemeinsamen Schicksals, der objektiven Einstellung, der Geschlossenheit, der guten Kurve sowie der Faktor der Gewohnheit bzw. der Erfahrung. Gemäß dem Faktor der Nähe werden Gruppen von Elementen mit kleinem Abstand eher zusammengefasst. Der Faktor der Gleichheit bezeichnet die Tendenz, gleichartige Elemente gemeinsam zu gruppieren. Werden Elemente gemeinsam verändert, so werden diese als Gruppe wahrgenommen. Der dahinter stehende Faktor ist derjenige des gemeinsamen Schicksals. Wenn Konstellationen von Elementen fortwährend variiert werden, so bestimmt nach dem Faktor der objektiven Einstellung die Reihenfolge der Darbietung dieser Konstellationen, wie eine bestimmte Konstellation wahrgenommen wird. Der Faktor der Geschlossenheit bezieht sich auf die Tendenz, eher geschlossene Figuren wahrzunehmen. Zudem werden Konstellationen bevorzugt, in denen die Elemente, beispielsweise eine Kurve, ihrem natürlichen Verlauf folgen. Dies wird mit dem Faktor der guten Kurve benannt. Schließlich bezieht sich der Faktor der Erfahrung darauf, dass die Wahrnehmung einer bestimmten Konstellation von der Gewohnheit beeinflusst ist. Das übergreifende Prinzip ist die *Tendenz zur guten Gestalt*, die etwa durch „Einfachheit, Regelmäßigkeit, inneres Gleichgewicht, Symmetrie, Geschlossenheit, einheitlicher Gesamtcharakter, Aufbau nach einem durchweg herrschenden Gesamtprinzip“ (Metzger, 1966, S. 710) charakterisiert ist. Besonders relevant für die nachfolgenden Ausführungen ist gerade dieses Prinzip. Wenn nämlich Wertheimer (1923) bemerkt, dass für die Wahrnehmung einer guten Gestalt keine exakte Anordnung der Punkte erforderlich ist, impliziert dies, dass gute Gestalten *konstruiert* werden. Besonders deutlich

wird dies in dem von Wertheimer (1924) formulierten „Grundproblem der Gestalttheorie“, in dem die Geschehnisse an einem Teil des Ganzen auf die Wirksamkeit der Strukturgesetze des Ganzen zurückgeführt werden. Knapp und präzise formuliert es auch Koffka (1935): “A gestalt is therefore a product of organization, organization the process that leads to gestalt” (S. 682). Gerade dieser Punkt, der die Möglichkeit der bereits angesprochenen Restrukturierungen und Reinterpretationen impliziert, ist für kognitive Konsistenztheorien (z.B. Festinger, 1957; Heider, 1946) und für PCS-Ansätze (z.B. Glöckner und Betsch, 2008b; Simon, Pham, Le & Holyoak, 2001) entscheidend.

Die Breite des Geltungsbereiches gestaltpsychologischer Prinzipien wird besonders durch die Anwendung auf Bereiche außerhalb der Wahrnehmung (Koffka, 1935; Köhler, 1947) verdeutlicht. Es sei hervorgehoben, dass die Gestaltpsychologie nicht nur bestimmte Phänomene auf andere Weise zu erklären suchte, sondern auch generell einen neuen Zugang zu wissenschaftlichen Problemen propagierte. Anstatt die einzelnen Elemente zu trennen und ihre Zusammenhänge zu analysieren, sollten eher das Ganze und seine Organisationsprinzipien studiert werden (Wertheimer, 1924).

Obgleich die gestaltpsychologischen Prinzipien die Grundlage für eine Vielzahl sozialpsychologischer Arbeiten darstellten, nahm ihr Einfluss im Laufe der Jahrzehnte ab (Read et al., 1997). Als Hauptgründe führen die Autoren die mangelnde Konkretisierung von Begriffen wie holistische Verarbeitung sowie methodische Unzulänglichkeiten hinsichtlich der Untersuchung komplexer dynamischer Prozesse an. Somit konnte die Tendenz zur guten Gestalt als „abstraktes und übergeordnetes Organisationsprinzip“ (Müsseler und Prinz, 2002, S. 39) nicht konkretisiert werden.

Bevor die für diese Problematik vorgeschlagenen Lösungsansätze, nämlich PCS-Netzwerke, thematisiert werden, wird im folgenden Kapitel 2.1.2 zunächst eine Klasse von Theorien vorgestellt, die – wie bereits erwähnt – maßgeblich von der Gestaltpsychologie beeinflusst wurden: die Konsistenztheorien.

2.2.2 Kognitive Konsistenztheorien

Newcomb attestierte der kognitiven Konsistenzforschung bereits 1968 eine lange Vergangenheit, aber eine kurze Geschichte (Newcomb, 1968b). Diese setzt für ihn mit den Arbeiten Heiders ein, wobei er durchaus die Einflüsse weiterer Arbeiten anerkennt, u.a. diejenigen des Gestaltpsychologen Lewin. Andere Forscher stellen einen deutlicheren und überzeugenden Bezug zwischen der Gestaltpsychologie und den kognitiven Konsistenztheo-

rien her. So konstatieren Read et al. (1997), dass der gestaltpsychologische Ansatz von zentraler Bedeutung für die Entwicklung kognitiver Konsistenztheorien war. Bei der Begründung betonen sie die Rezeption der Annahme struktureller Dynamik. Simon, Snow und Read (2004) führen diese Auffassung weiter aus, indem sie vier Prinzipien struktureller Dynamik anführen, die entscheidend für kognitive Konsistenztheorien waren: (1) Kognitive Zustände werden nicht nur durch einzelne Elemente bestimmt, sondern auch durch die holistisch zu sehende Struktur. (2) Die Eigenschaften einer Struktur sind dynamisch. (3) Im Rahmen dieser Dynamik besteht eine Tendenz zur Prägnanz. (4) Die dynamischen Veränderungen beinhalten Rekonstruktionen der kognitiven Elemente. Eine Vielzahl von Theorien wurde entwickelt, die sich der Frage der kognitiven Konsistenz in verschiedener Hinsicht angenommen haben (z.B. Abelson und Rosenberg, 1958; Festinger, 1957; Heider, 1958; Osgood und Tannenbaum, 1955). Ein Blick auf die Publikationsdaten macht bereits deutlich, was vielfach konstatiert wurde (z.B. Shultz und Lepper, 1996; Simon, Snow et al., 2004): Die kognitive Konsistenzforschung hatte Mitte des 20. Jahrhunderts ihre Blütezeit, in der Folgezeit nahm das Interesse an diesbezüglicher Forschungsarbeit aus Gründen, die später erläutert werden, ab. Der von Abelson et al. (1968) herausgegebene Review kognitiver Konsistenztheorien umfasst folglich deren wichtigste Ansätze. Als gemeinsamer Nenner dieser Theorien wird die Annahme, dass eine Person durch ihr Verhalten nach der Maximierung der internalen Konsistenz ihres kognitiven Systems strebt, gesehen (Newcomb, 1968b). Im Folgenden werden nun die wichtigsten Theorien im Überblick vorgestellt. Da eine erschöpfende Darstellung nicht das Ziel des Kapitels ist, wird auf die wichtigsten und für die weiteren Ausführungen relevantesten Aspekte der Theorien in ihrer ursprünglichen Fassung fokussiert. Dadurch wird verdeutlicht, wie Konsistenz in diesen Ansätzen interpretiert und wie das Streben nach Konsistenz erforscht wurde. Der im Review von Abelson et al. (1968) vertretenen Auffassung folgend, dass die Dissonanztheorie, die Balancetheorie, die Kongruitätstheorie, die Theorie affektiver-kognitiver Konsistenz, die Psycho-Logik und die Theorie McGuires das Spektrum an Konsistenztheorien abdecken, werden diese Theorien referiert. Eingeleitet wird diese Darstellung durch die Arbeit von Asch (1946), der Read et al. (1997) nicht nur hohe Bedeutung zumessen, sondern auch einen deutlichen Bezug zur Gestaltpsychologie attestieren.

Asch (1946) widmet sich der Frage, wie der Prozess der Eindrucksbildung zu beschreiben ist. Insbesondere möchte er klären, wie eine Vielzahl an Charakteristika in ein einheitliches Bild integriert wird, und zwar mit einer hohen Schnelligkeit und Leichtigkeit.

Ausgehend von vorherigen Befunden, nach denen Versuchspersonen zur Konstruktion eines konsistenten Gesamtbildes anstelle der Auflistung separater Eigenschaften tendieren, untersuchte Asch in insgesamt zehn Experimenten, welche Faktoren die Eindrucksbildung beeinflussen. Hierbei wurde den Versuchspersonen stets eine Liste an Adjektiven vorgelesen, auf deren Basis sie den Charakter der Person beschreiben sollten. Die Kernbefunde Aschs waren, dass zentrale Eigenschaften einen größeren Einfluss auf das Gesamtbild haben als periphere Eigenschaften, dass die Interpretation einer Eigenschaft im Kontext anderer Eigenschaften erfolgt, dass die Reihenfolge der Eigenschaften für das resultierende Gesamtbild von Bedeutung ist und dass eine Eigenschaft auch weitere Eigenschaften impliziert. An mehreren Stellen wird die von Read et al. (1997) thematisierte Nähe zur Gestaltpsychologie deutlich. So umschreibt Asch (1946) etwa mit folgender Aussage die Konstruktion einer guten Gestalt: “In some manner he shapes the separate qualitities into a single, consistent view” (S. 261). Dass von einem Streben nach Konsistenz gesprochen werden kann, wird etwa in der folgenden Äußerung deutlich: “The subject aims at a clear view (…)” (S. 278). Des Weiteren spricht Asch die strukturelle Dynamik an, wenn er auf Basis seiner experimentellen Ergebnisse folgert, dass Eigenschaften in einem dynamischen Verhältnis zueinander stehen: So kann etwa Eigenschaft A (schnell) aus Eigenschaft B (fachkundig) folgen. Bezogen auf die gesamten Eigenschaften und den daraus resultierenden Gesamteindruck spricht Asch treffend von einer “relational determination” (S. 283). Auch Asch selbst expliziert die Nähe zur Gestaltpsychologie, und zwar hinsichtlich des Prinzips, dass ein einzelnes Teil je nach Kontext einen unterschiedlichen Stellenwert und eine unterschiedliche Bedeutung haben kann. Angesichts dieser Äußerungen und vor dem Hintergrund der Ausführungen des Kapitels 2.1.2 erscheint die von Asch gezogene Parallele zu den Gestaltqualitäten nach Ehrenfels eher überraschend bzw. unpassend. Insgesamt betrachtet verdeutlicht Aschs Arbeit aber durchaus die Nähe zur Gestaltpsychologie und zudem die Bedeutung von Konsistenz bei der Eindrucksbildung.

Eine besonders bedeutende Stellung innerhalb der Konsistenzforschung nimmt die Balancetheorie Heiders (1946; 1958) ein. Newcomb (1968b) hebt hervor, welch maßgeblichen Einfluss Heiders Theorie auf nachfolgende Theorien genommen hat, insbesondere die zentrale Annahme “that an imbalanced set of cognitions is associated with ‘tension’ and the arousal of forces which tend to restore or to attain states of balance” (S. XVI). Ausgehend von der Feststellung, dass Strukturen balanciert oder unbalanciert sein können, erläutert Heider (1946; 1958), wann ein Balancezustand vorherrscht und wie dieser erreicht

werden kann. Er konzentriert sich dabei auf Strukturen, die aus zwei oder drei Elementen bestehen und zwischen denen Beziehungen existieren. Hinsichtlich der Elemente differenziert Heider zwischen der Referenzperson *p*, einer weiteren Person *o* und einer unpersönlichen Entität *x*. Hinsichtlich der Beziehungen unterscheidet er zwischen solchen, die die Art der Beziehung zwischen zwei Einheiten oder Elementen beschreiben, und solchen, die die Valenz einer Beziehung beschreiben. Erstere umfasst z.B. Ähnlichkeit, Mitgliedschaft und Besitz, letztere z.B. Liebe, Wertschätzung und Ablehnung. In seiner umfassenderen Arbeit von 1958 bezeichnet Heider diese Beziehungen als *unit relations* und als *sentiment relations*. In Dyaden liegt ein balancierter Zustand laut Heider dann vor, wenn alle Beziehungen zwischen ihnen positiv oder negativ sind. In Triaden ist Balance erzielt, wenn entweder alle Beziehungen positiv sind oder wenn zwei Beziehungen negativ sind und eine positiv ist. Wie bereits angesprochen, besteht in einer unbalancierten Struktur die Tendenz zur Balance. Hierbei sind Veränderungen sowohl in den *unit relations* als auch in den *sentiment relations* möglich. Heider ist sich durchaus einiger Unzulänglichkeiten seiner Theorie bewusst, wenn er ihr einen vorläufigen Status attestiert (1958) oder auf einige Ausnahmen seiner Balance-Definition hinweist, z.B. bei Neid, Eifersucht oder niedriger Selbstachtung (1946). Read et al. (1997) führen die mangelnde Erklärungskraft der Balancetheorie für Strukturen, in denen Eifersucht besteht, auf den formalen Charakter derselben zurück. Die Berücksichtigung der zugrunde liegenden Konzepte (Freundschaft vs. Partnerschaft) halten sie für zielführender. Weitere Kritikpunkte an der Theorie Heiders sind die fehlende Berücksichtigung der Intensität einer *sentiment relation* sowie die Reduktion der Theorie auf Dyaden und Triaden (z.B. Klauer, 2006; Read et al., 1997). Außerdem kann die Richtung der Veränderung ebenso wenig vorhergesagt werden wie der Einfluss auf das Verhalten (Herkner, 1991). Während im Rahmen dieser Darstellung die ursprüngliche Theorie Heiders als Zwischenglied zwischen der Gestaltpsychologie und der Konsistenzforschung im Vordergrund steht, sei darauf hingewiesen, dass die Theorie in der Folgezeit Erweiterungen und Modifikationen unterzogen wurde, in denen ein Teil der Kritik berücksichtigt wurde (Klauer, 2006). Der interessierte Leser sei insbesondere auf die Arbeiten von Cartwright und Harary (1956) und Newcomb (1968a) hingewiesen. Auch die im Folgenden darzustellenden Ansätze sind – wie bereits angesprochen – von der Balancetheorie Heiders beeinflusst, sind aber eher als eigenständige Konsistenztheorien zu betrachten.

Eine Theorie von besonderer Prominenz und beachtlichem Einfluss ist die Dissonanztheorie von Leon Festinger (1957; Festinger, 1964). Klassische Experimente, etwa zur Bedeutung der Vergütungshöhe für die Bewertung einer langweiligen Aufgabe (Festinger und Carlsmith, 1959), zum postdezisionalen *spreading apart* von Alternativen (Brehm, 1956) sowie zur Notwendigkeit, ein gezeigtes Verhalten zu rechtfertigen (Aronson und Mills, 1959) oder ein unterlassenes Verhalten zu begründen (Aronson und Carlsmith, 1963) sind über die Grenzen der Psychologie bzw. der Wissenschaft hinaus allgemein bekannt. Harmon-Jones und Harmon-Jones (2007) konstatieren nicht nur eine Überlegenheit der Dissonanztheorie in der Sozialpsychologie von 1950 bis 1970, sondern attestieren ihr sogar, die psychologische Betrachtungsweise kognitiver Prozesse revolutioniert zu haben. Obgleich dieser Theorie ein Mangel an Präzision vorgeworfen wird, ist ihr Einfluss unbestritten (Aronson, 1968). An dieser Stelle ist die Einschätzung von Festinger (1957), dass die Reduktion von Dissonanz ein basaler Prozess ist, wichtig: Hierdurch wird nicht nur die Vielzahl an Anwendungsmöglichkeiten begründet, sondern zugleich die Bedeutung der Theorie unterstrichen. Festinger (1957) gründet seine Theorie auf zwei zentrale Hypothesen: (1) Dissonanz ist ein psychologisch unangenehmer Zustand. Folglich besteht die Motivation, Dissonanz zu reduzieren und so Konsonanz zu erreichen. (2) Eine Erhöhung bestehender Dissonanz wird durch die Vermeidung entsprechender Situationen und Informationen verhindert. Mit *Dissonanz* bezeichnet Festinger eine bestimmte Art von Relation zwischen zwei Elementen: "*These two elements are in a dissonant relation if, considering these two alone, the obverse of one element would follow from the other*" (S. 13, Hervorhebung im Original). Dementsprechend ist eine Relation konsonant, wenn ein Element aus dem anderen folgt. Ist beides nicht der Fall, so ist die jeweilige Relation irrelevant. Für die Dissonanztheorie sind lediglich die relevanten, das heißt die dissonanten und konsonanten Relationen bedeutsam. Die Stärke der Dissonanz ergibt sich aus dem Verhältnis dissonanter zu allen relevanten Kognitionen. Diese Kognitionen sind hierbei entsprechend ihrer Bedeutung gewichtet. Die Reduktion von Dissonanz kann auf unterschiedlichen Wegen erfolgen. Festinger nennt zum einen die Änderung eines behavioralen oder eines umweltbezogenen kognitiven Elements. Hierbei ist zu beachten, dass diese Elemente mehr oder weniger resistent gegenüber Veränderungen sein können. Die Höhe der Resistenz ist unter anderem davon abhängig, ob ein kognitives Element mit Verlustängsten oder mit Befriedungsgefühlen verbunden ist. Vor allem ist stets der Realitätsbezug zu wahren. Festinger verdeutlicht dies an dem Beispiel eines Fußgängers, der nicht leugnen kann, dass er gerade die Straße

entlangläuft, ohne diesen Bezug aufzugeben. Es wird stets das Element verändert, das am wenigsten veränderungsresistent ist. Eine weitere Möglichkeit besteht im Hinzufügen neuer kognitiver Elemente, wodurch das Gewicht der Dissonanz reduziert wird. Festinger bezieht seine Theorie auf verschiedene Anwendungsbereiche, nämlich Entscheidungen, erzwungene Einwilligung (*forced compliance*), Informationssuche und soziale Unterstützung, und stellt entsprechende empirische Daten vor. An dieser Stelle sei aufgrund des direkten Bezugs zur Fragestellung dieser Forschungsarbeit der Anwendungsbereich der Entscheidungen kurz skizziert. Festinger nimmt an, dass Dissonanz nach einer Entscheidung zwischen zwei Optionen mit positiven und negativen Aspekten auftreten kann, da die getroffene Wahl dissonant ist mit den negativen Aspekten der gewählten Option und mit den positiven Aspekten der nicht gewählten Option. Festinger (1964) ordnet die Dissonanzreduktion explizit der postdezisionalen Phase zu. Als wichtige Voraussetzung für das Auftreten von Dissonanz nennt er das *Commitment*, das heißt die Bindung an eine Entscheidung. Diese Bindung impliziert für Festinger, dass das folgende Verhalten durch die Entscheidung beeinflusst wird. Die Größe der Dissonanz ist nach Festinger (1957) von der Wichtigkeit der Entscheidung, der Attraktivität der nicht gewählten Alternative und der kognitiven Überlappung (*cognitive overlap*) der beiden Alternativen abhängig. Letzteres bezeichnet die Ähnlichkeit zweier Alternativen. Als Beispiel vergleicht Festinger die Wahl zwischen zwei Büchern mit der Wahl zwischen einem Buch und einem Konzertticket – im ersten Fall ist die Ähnlichkeit zwischen den Optionen deutlich höher als im zweiten Fall. Für die Reduktion der Dissonanz führt Festinger drei Strategien an, und zwar erstens die (eher seltene) psychologische Rücknahme der Entscheidung, indem etwa die Entscheidung als falsch anerkannt wird oder jegliche Verantwortung geleugnet wird, zweitens die Änderung von Kognitionen über die Alternativen und drittens die Bildung eines kognitiven *Overlaps*. Nicht nur für den Anwendungsbereich der Entscheidungen ist die Frage interessant, in welcher Situation welche Strategie verwendet wird. Diesbezüglich sieht Aronson (1968) Mängel in der Theorie und verweist auf individuelle Unterschiede als Erklärungsmöglichkeit. Personfaktoren könnten seiner Meinung nach auch für weitere Punkte relevant sein, z.B. für die Höhe der Toleranz von Dissonanz und das Empfinden von Dissonanz. Lediglich angemerkt werden soll, dass Aronson selbst bezüglich des letzten und seiner Meinung nach wichtigsten Punktes die Bedeutung des Selbstkonzeptes hervorhebt. Ergänzend ist anzuführen, dass die Dissonanz nie so hoch sein darf, dass sie nicht die Konsequenzen der Entscheidung rechtfertigt, da ansonsten keine Entscheidung getroffen würde (Zajonc, 1960).

Auch Aronsons (1968) Feststellung, dass manchmal Personen eine bestehende Dissonanz nicht reduzieren und nicht vermeiden, da auf diesem Wege ein Lernen aus Fehlern ermöglicht wird, verdient Beachtung. Die Dissonanztheorie ist in verschiedener Hinsicht kritisiert worden. Eine für die vorliegende Betrachtung passende Zusammenfassung geben Simon und Holyoak (2002): Sie nennen insbesondere die unzureichende Berücksichtigung von Komplexität, die Fokussierung auf die Vermeidung von Dissonanz statt auf das Streben nach Konsistenz sowie die fehlenden bzw. mangelhaften Erklärungen auf prozessualer Ebene. Zudem hätte die Forschung zur kognitiven Dissonanz, inklusive der Weiterentwicklungen, die die Dissonanztheorie erfahren hat (z.B. Aronson, 1968; Cooper und Fazio, 1984), eine Abkehr von den gestaltpsychologischen Prinzipien zur Folge gehabt. Zwar kommen Harmon-Jones und Harmon-Jones (2007) in einem aktuellen Review von Forschungsarbeiten und theoretischen Weiterentwicklungen der Dissonanztheorie zu dem Schluss, dass die ursprüngliche Konzeption der Dissonanztheorie die meiste Stützung erfährt, dennoch bleiben die genannten Kritikpunkte bestehen und wiegen schwer.

Ein Modell von weitaus geringerer Breite ist das Kongruitätsprinzip von Osgood und Tannenbaum (1955). Dieses Modell ist verwandt mit Heiders Balancetheorie und kann aufgrund seiner Fokussierung auf den Einstellungsbereich als deren Spezialfall betrachtet werden (Stahlberg und Frey, 1997). Tannenbaum (1968) bezeichnet sie als eine der Konsistenztheorien, deren Anwendungsbereich am stärksten begrenzt ist. Osgood und Tannenbaum (1955) weisen selbst in der ursprünglichen Konzeption des Kongruitätsprinzips explizit darauf hin, dass ihr Modell keinen Bezug zu einem übergreifenden Ansatz hat, und sehen zu diesem Zeitpunkt auch keine Notwendigkeit, diesen Anspruch erheben zu wollen. Tannenbaum (1968) macht in seinem Review des Kongruitätsprinzips darauf aufmerksam, dass die Forschung zu der Beurteilung komplexer Stimuli schließlich zu einem übergreifenden Modell, dem Kongruitätsmodell kognitiver Interaktion, führte. Er selbst befürwortet aber eine separate Betrachtung des Kongruitätsprinzips für Einstellungsänderung, was er mit den unterschiedlichen theoretischen Annahmen der beiden Modelle sowie mit empirisch ermittelten Differenzen zwischen diesen begründet. Daher soll im Folgenden lediglich die ursprüngliche Konzeption des Modells dargelegt werden. Osgood und Tannenbaum (1955) schlagen ein Modell vor, mit dem sie die Veränderung von Einstellungen vorhersagen können. Diese Forschungsarbeit gründet – wie Tannenbaum (1968) festhält – auf der damals neuen Methode des semantischen Differentials (Osgood, Suci & Tannenbaum, 1957). Bei dieser Methode werden Begriffe auf mehreren siebenstufigen Skalen eingeschätzt, wobei die

Endpunkte der Skalen mit -3 und +3 bzw. mit bipolaren Adjektivpaaren, wie z.B. aktiv – passiv, bezeichnet sind. Ergebnisse von Faktorenanalysen zeigten, dass sich die Bedeutung eines Konzepts auf drei Dimensionen abbilden lässt, nämlich Aktivierung, Potenz und Evaluation. Die Einstellung zu einem Konzept lässt sich laut Osgood und Tannenbaum (1955) auf der evaluativen Dimension abbilden und kann damit mehr oder weniger polarisiert sein. Diese Einstellungen können sich je nach Konzept durchaus unterscheiden, ohne dass Veränderungsdruck bestünde. Werden allerdings zwei Objekte durch eine Aussage verbunden, so wird die Frage nach der Kongruität relevant. Kongruität besteht, wenn beide Objekte gleich hoch eingeschätzt werden, und zwar in der gleichen evaluativen Richtung bei einer positiven Aussage und in der entgegengesetzten Richtung bei einer negativen Aussage. Hier zeigt sich deutlich die Verwandtschaft mit Heiders Balancetheorie, insbesondere aber ihre Weiterentwicklung: Bei der Kongruität ist nicht nur das Vorzeichen, sondern auch die Stärke der Einstellung relevant (Stahlberg und Frey, 1997). Ist keine Kongruität vorhanden, besteht laut Osgood und Tannenbaum (1955) Druck, diese durch Einstellungsänderung zu erzielen. Hierbei werden sowohl die Quelle der Äußerung als auch das Konzept verändert, wobei die Änderung invers proportional zur Stärke der Polarisierung erfolgt. Das heißt, dass die wenig polarisierten Einstellungen eher verändert werden als die stark polarisierten. Zudem ist zu berücksichtigen, dass aufgrund des Richtungseinflusses einer Aussage das Konzept im Vergleich zur Quelle stärkere Änderungen erfährt, und dass die Unglaubwürdigkeit einer Nachricht dem Änderungsdruck entgegenwirkt. Diese Annahmen wurden von den Autoren mathematisch formalisiert und empirisch gestützt. Ein weiterer interessanter Befund, den Tannenbaum (1968) darlegt, sei an dieser Stelle noch erwähnt: So konnte Tannenbaum zeigen, dass auch die Einstellung zu einem zweiten Objekt, das zwar nicht direkt mit dem ersten Objekt verbunden ist und zu dem keine Aussage gemacht wird, das aber mit der Quelle der Äußerung verbunden ist, ebenfalls verändert wird, wenn aufgrund einer Inkongruität zwischen der Quelle und dem ersten Konzept die Einstellung zur Quelle verändert worden ist. Dies verdeutlicht eindrucksvoll, dass auch nach diesem Ansatz Konsistenz stets ganze Strukturen betrifft und die wechselseitigen Verflechtungen sowie deren Dynamik zu beachten sind. Abschließend ist festzuhalten, dass auch wenn der Anwendungsbereich der Theorie begrenzt ist, insbesondere ihre Präzision sowie die Quantifizierung der Einstellungen ihren Status rechtfertigen. Die Möglichkeit, die Höhe der Reduktion von Inkonsistenz vorherzusagen, wertet etwa Herkner (1991) als deutlichen Vorzug dieser Theorie.

Eine weitere bedeutende Konsistenztheorie, die sich auf den Einstellungsbereich bezieht, ist die sogenannte Theorie der affektiv-kognitiven Konsistenz von Rosenberg (1956; Rosenberg, 1960). Rosenberg (1968) selbst schreibt seinem Ansatz eher den Status eines Modells als den einer Theorie zu und hält die Charakterisierung als "radial structure model of intraattitudinal balancing" (S. 73) für passender als die geläufige Bezeichnung als Theorie der affektiv-kognitiven Konsistenz. Hiermit wird ein wichtiges Merkmal dieses Ansatzes angesprochen, nämlich die Einschätzung als strukturelles Modell. So geht es weniger um die Untersuchung von Bedingungen und Einflussfaktoren, die für das Lernen oder die Veränderung von Einstellungen relevant sind, sondern vielmehr um die Frage, was Einstellungen sind und wie Einstellungen verändert werden (Rosenberg, 1960). In einer früheren Arbeit definiert Rosenberg Einstellungen als affektive Antwort auf ein Objekt und sieht sie in struktureller Verbindung mit Überzeugungen, mittels dieses Einstellungsobjekts bestimmte Werte zu realisieren (Rosenberg, 1956). Später erweitert er dann den Einstellungsbegriff, indem er der affektiven Komponente eine kognitive Komponente hinzufügt, wobei er die erwähnte Überzeugung über Beziehungen zwischen dem Einstellungsobjekt und anderen Objekten von subjektiver Bedeutung oder Werten als besonders wichtige Kognition nennt (Rosenberg, 1960). Wie Rosenberg (1956) ausführt und empirisch belegt, hängt die affektive Reaktion auf ein Einstellungsobjekt sowohl von der Instrumentalität dieses Objektes zur Realisierung wichtiger Werte als auch von der Wichtigkeit dieser Werte ab: Je höher die Instrumentalität und je wichtiger die betroffenen Werte, desto stärker die affektive Reaktion. Während so bereits nachgewiesen werden konnte, dass affektive und kognitive Einstellungskomponenten bei bestehenden Einstellungen zusammenhängen, erbrachte Rosenberg später den Beleg dieser Beziehung bei der Veränderung von Einstellungen (Rosenberg, 1960). Diesbezüglich stellt Rosenberg die Hypothese auf, dass "*the production of inconsistency between the affective and cognitive portions of an attitude will culminate in a general attitude reorganization (...) when (1) the inconsistency exceeds the individual's present tolerance limit and (2) the force producing it cannot be ignored or avoided*" (S. 323, Hervorhebung im Original). Hierbei kann die Inkonsistenz von der Veränderung einer kognitiven Komponente hervorgerufen werden und zu einer Veränderung der affektiven Komponente führen und umgekehrt. Für beide Richtungen führt Rosenberg empirische Befunde an. Diese wechselseitige Einflussnahme kann laut Rosenberg (1956) auch das Lernen von Einstellungen erklären: So kann einerseits die Aneignung von Überzeugungen zu einer affektiven Antwort führen, andererseits kann auch die Imititation

einer affektiven Reaktion eine kognitive Begründung einleiten. Insbesondere die Einnahme einer strukturellen Perspektive und die Berücksichtigung der affektiven Komponente von Einstellungen zeichnen diesen Ansatz im Rahmen der Konsistenzforschung aus.

Äußerungen verschiedener Autoren belegen, welchen Stellenwert Abelsons Beiträge für sie hatten (z.B. Langer, 1994; Read und Miller, 1994). Für den Bereich der Konsistenztheorien war das von ihm und Milton Rosenberg entwickelte Modell der Psycho-Logik (1958) von besonderer Relevanz, in dem sie die Struktur und die Prozesse des Denkens abbildeten. Dem Problem der mangelnden Spezifikation von Konsistenztheorien begegneten sie mit einem spezifischeren Ansatz (Abelson, 1968). So stellten Abelson und Rosenberg (1958) nicht nur ihr psychologisches Modell, sondern zugleich dessen mathematische Formalisierung vor. Im Folgenden soll nun lediglich das psychologische Modell nach Abelson und Rosenberg (1958) skizziert werden: Ausgehend von der Annahme, dass eine Einstellung aus einer affektiven und einer kognitiven Komponente besteht und aufgrund deren Interaktion als kognitive Struktur zu sehen ist, schlagen sie ein Modell zur Erfassung dieser Einstellungsstruktur vor. Zentrale Bestandteile dieses Modells sind kognitive Elemente (Akteure, Mittel, Ziele), die zwischen diesen Elementen bestehenden Beziehungen (positiv, negativ, null, ambivalent) sowie kognitive Einheiten, die durch den Aufbau einer bestimmten Beziehung zwischen zwei Elementen entstehen. Um nun die kognitive Struktur zu erfassen, werden Personen zu ihrer sogenannten *conceptual arena*, bestehend aus Wörtern oder Sätzen, die sie mit einem bestimmten Thema verbinden, befragt. Hieraus kann eine Strukturmatrix gebildet werden, die die Relationen zwischen den relevanten Elementen anzeigt. Wenn eine Person über ein Thema nachdenkt, können durch die Kombination bestehender kognitiver Einheiten, die ein gemeinsames kognitives Element aufweisen, weitere kognitive Einheiten generiert werden. Da entsprechende Schlussfolgerungen nicht zwangsläufig den Kriterien formaler Logik entsprechen, wählen Abelson und Rosenberg die Bezeichnung *Psycho-Logik*. Das Nachdenken über ein Thema, für das die Autoren Motivation voraussetzen, kann je nach kognitiver Struktur Inkonsistenz offenbaren oder nicht. Ersteres ist im Fall unbalancierter Strukturen, letzteres im Fall balancierter Strukturen gegeben, wobei die Definitionen von Balance und Inbalance denen Heiders (1946; 1958) entsprechen. Als Möglichkeiten, die Inbalance aufzuheben, nennen Abelson und Rosenberg die Veränderung der Relationen, die Änderung von Elementen sowie das Beenden des Denkprozesses, wobei die Wahl der Möglichkeit von dem Kontext und der Veränderungsresistenz der kognitiven Struktur abhängt. In einer späteren Arbeit setzt sich Abelson (1968)

mit den Stärken und Schwächen der Psycho-Logik auseinander und stellt eine theoretische Weiterentwicklung vor. So verweist er u.a. auf weitere Möglichkeiten der Reduktion von Inbalance, z.B. durch die Verstärkung eines Elementes, die Differenzierung eines Elementes in einen „guten“ und einen „schlechten“ Teil sowie die Rationalisierung. Zudem erläutert er das Prinzip des evaluativen Transfers, nach dem sich die Elemente wechselseitig in ihrer affektiven Ladung beeinflussen, was auch für die Einstellungsänderung maßgeblich ist. Schließlich fasst er basale psychologische Implikationsprinzipien zusammen, die auf zusammengehörige Elemente, die Implikationsmoleküle, wirken. Hier nimmt Abelson direkt Bezug zum gestaltpsychologischen Prinzip der Prägnanz, wenn er von einer *“tendency to complete the molecule*” (1968, S. 133, Hervorhebung im Original) spricht. Interessanterweise verbindet Abelson als Person auch in gewisser Hinsicht die kognitive Konsistenzforschung mit den PCS-Ansätzen, da Read und Miller als Vertreter des letztgenannten Ansatzes durch Abelson maßgeblich beeinflusst wurden: “Bob’s influence runs like a thread (sometimes a fairly thick one) through much of our subsequent work (…)” (Read und Miller, 1994, S. 210). Rückblickend erfährt gerade diese Konsistenztheorie bei Simon und Holyoak (2002) eine besondere Wertschätzung für ihre Beschreibung kognitiver Restrukturierungen. Die zweite Theorie, die Simon und Holyoak für diesen Aspekt zu Recht würdigen, ist diejenige McGuires, die im Folgenden vorgestellt wird.

“In my general strategy, the need for consistency was taken for granted, as a means rather than an end of the study” (McGuire, 1968, S. 140). Mit dieser Aussage charakterisiert McGuire nicht nur sein eigenes Forschungsprogramm im Rahmen der Konsistenzforschung, sondern grenzt sich auch von den anderen Konsistenztheoretikern ab. Sein Ziel war es nicht, das Streben nach Konsistenz zu erforschen, sondern vielmehr die zugrunde liegende Struktur des Denkens zu ergründen. Obgleich er in einigen Modellen wie dem der Psycho-Logik entsprechende Ansätze sieht, bemängelt er doch die mangelnde Fokussierung auf die kognitive Struktur. Folglich zielten seine Studien zum Einfluss von Überzeugungen auf die eigene Meinung auf das Testen von Hypothesen über den Aufbau und die Funktionsweise des kognitiven Systems ab (McGuire, 1960a; 1960b). Dass ein Streben nach Konsistenz vorliegt, ist für ihn nicht der Gegenstand, sondern Bestandteil seiner zentralen Annahme: Die Komponenten des kognitiven Systems sind miteinander verbunden und es besteht die Tendenz, die internale Kohärenz unter diesen Komponenten aufrechtzuerhalten (McGuire, 1968). Unter Komponenten versteht McGuire die verbalisierbare Aussage, die Proposition, über ein bestimmtes Objekt, das hierfür zuvor auf einer Urteilsdimension bewertet wurde.

Die Verbindung der Propositionen erfolgt laut McGuire nach den Gesetzen der formalen Logik sowie der Wahrscheinlichkeitstheorie. Durch den Einbezug der Wahrscheinlichkeitstheorie konnten graduierte und damit adäquatere Urteile abgegeben werden – ein Vorzug, den McGuire bei den übrigen Konsistenztheorien vermisst. Hier fragt McGuire nicht nur nach der Wahrscheinlichkeit, dass eine Proposition wahr ist, sondern auch nach deren Attraktivität. In seinen experimentellen Studien (1960a; 1960b) wandte McGuire seinen Ansatz auf den Bereich der Überzeugung an, wobei er als Untersuchungsmaterial auf Syllogismen zurückgriff. So untersuchte McGuire (1960a) etwa, ob allein das Verbalisieren von Meinungen einer Person diese (unbewusst) für mögliche Inkonsistenzen sensibilisiert und dadurch Veränderungen induziert werden. McGuire bezeichnet dies als *Sokratische Methode.* In der Tat fand McGuire Bestätigung für diese Hypothese, dass mittels der Sokratischen Methode das Bedürfnis nach Konsistenz geweckt und somit eine Woche später ein höheres Maß an Konsistenz zwischen den Meinungen zu finden war. Des Weiteren ging McGuire in dieser wie auch in einer weiteren Studie (1960b) der Frage nach, ob der Effekt einer Überzeugungsaussage auch auf logisch zugehörige, aber ungenannte Propositionen ausstrahlt. Für beide Aspekte fand McGuire empirische Stützung. Ein weiterer wichtiger Befund war der sogenannte "delayed-action effect" (McGuire, 1960a, S. 350), der besagt, dass der Einfluss von Überzeugungsaussagen auf ungenannte Propositionen erst nach einer gewissen Zeit, in diesem Fall im Laufe einer Woche, zum Tragen kommt. Die Befunde, dass der Einfluss auf ungenannte Propositionen schwächer ist als der gemäß der formalen Logik zu erwartende sowie dass dieser Einfluss graduell, das heißt über einen längeren Zeitraum, ausgeübt wird, bezeichnet McGuire (1960a) als kognitive Trägheit. In einer späteren Überblicksarbeit diskutiert McGuire (1968) die kognitive Trägheit als wichtigen Einflussfaktor für das menschliche Denken, der nicht durch die Gesetze der formalen Logik erfasst wird. Während McGuire prinzipiell einen weiteren Unterschied zu den anderen Ansätzen in seiner Annahme sieht, dass nicht Inkonsistenz minimiert, sondern Konsistenz maximiert werden soll, weist er doch auch darauf hin, dass nicht maximale Kongruenz im Sinne formaler Logik angestrebt werden muss, sondern "some kind of congruence" (S. 160), und dass auch Inkongruenz einen Reiz zu haben scheint. Insbesondere McGuires Kritik an der damaligen Konsistenzforschung sowie seine daraus abgeleitete Forderung, eher die Struktur des menschlichen Denkens zu ergründen, ist – wie im Folgenden deutlich werden wird – für die anschließende Forschung von Relevanz gewesen.

Zusammenfassend ist festzuhalten, dass sich die dargestellten Theorien zwar durchaus durch eine unterschiedliche Breite und Spezifikation sowie durch eine unterschiedliche Schwerpunktsetzung auszeichnen, dass ihr Kernelement aber dennoch identisch ist. So besteht in allen Konsistenztheorien die Annahme, dass Personen Inkonsistenz im kognitiven System zu vermeiden oder zu reduzieren suchen und stattdessen nach Konsistenz streben. Damit spiegelt sich das ursprünglich für den Bereich der Wahrnehmung formulierte gestaltpsychologische Prinzip der Tendenz zur guten Gestalt in diesen Ansätzen deutlich wider. Dass Konsistenz durch Reorganisationen und Reinterpretationen erreicht werden kann und damit auf einem dynamischen Prozess basiert, verdeutlicht ebenso den Einfluss der Gestaltpsychologie. Obgleich diese Kernaspekte die verschiedenen Konsistenztheorien verbinden, wurden doch auch die Unterschiede zwischen ihnen deutlich. Neben dem offensichtlichen Unterschied hinsichtlich der Breite und der Art des Anwendungsbereiches fällt insbesondere auf, dass in einigen Theorien (Abelson und Rosenberg, 1958; McGuire, 1968; Rosenberg, 1956) der Anspruch erhoben wird, über das reine Untersuchen des Konsistenzstrebens hinaus die Struktur und Funktionsweise des kognitiven Systems zu erfassen. Insbesondere die Ansätze von Abelson und Rosenberg sowie von McGuire werden von „modernen“ Konsistenzforschern wie Simon und Holyoak (2002) für diese Sichtweise gewürdigt. Besonders deutlich propagiert diese Perspektive McGuire (1968), der sich von den übrigen Konsistenzforschern abgrenzen möchte. Dass er hierbei das Streben nach Konsistenz als gegeben sieht, zeigt, dass er das Konsistenzbedürfnis nicht an besondere Voraussetzungen knüpft. Diese Interpretation ist insofern wichtig, als dass sie sich in späteren Konzeptionen des Konsistenzstrebens wiederfindet. Dies betrifft auch eine weitere entscheidende Annahme McGuires, nämlich dass nicht Inkonsistenz minimiert, sondern Konsistenz maximiert werden soll (vgl. auch Simon und Holyoak, 2002). Interessant erscheint zudem die in verschiedenen Arbeiten geäußerte Annahme, situative und personale Faktoren könnten einen Einfluss auf den Veränderungsprozess haben (z.B. Abelson und Rosenberg, 1958; Festinger, 1957; Rosenberg, 1960).

Wie bereits angesprochen, ging das Interesse an konsistenzbezogener Forschung nach 1960 deutlich zurück. Als Ursachen werden generelle Schwächen der Konsistenztheorien diskutiert. So erwähnen Read und Miller (1994) als zwei schwerwiegende Unzulänglichkeiten der kognitiven Konsistenztheorien zum einen, dass diese der Komplexität und der inhaltlichen Bedeutung der untersuchten Konzepte nicht gerecht wurden. Zum anderen bemängeln sie das Fehlen eines umfassenden formalen Modells, anhand dessen das Ausmaß

an Konsistenz bewertet werden könnte. Hinter dem letzten Punkt steht der auch von Simon, Snow et al. (2004) vorgebrachte kritische Hinweis, dass vielfach weder die Bedeutung der Konzepte noch die Stärke der Relationen berücksichtigt wurden. Osgood und Tannenbaum (1955) haben immerhin die Einstellung zu den Konzepten erfasst, allerdings nicht die Stärke der Verbindung zwischen diesen. Obgleich die Konsistenztheorien im Vergleich zur Gestaltpsychologie stärker präzisiert sind und ihre Annahmen einer experimentellen Testung unterzogen wurden, leiden sie doch ebenfalls unter einer begrifflichen Unschärfe und einer methodischen Begrenztheit. Sowohl die Stärken als auch die Schwächen der bisherigen Konsistenzforschung gaben nach einer Phase des Stillstands den Impuls für die moderne Konsistenzforschung in Gestalt von den sogenannten PCS-Netzwerken, die im folgenden Kapitel vorgestellt werden.

2.2.3 Parallel constraint satisfaction-Netzwerke

Die Konsistenzforschung stand vor einer Herausforderung: Auf der einen Seite hatte sich die Rezeption der gestaltpsychologischen Annahme struktureller Dynamik als fruchtbar erwiesen und interessante Konsistenzphänomene offenbart. Auf der anderen Seite konnten auch die Konsistenztheorien dem Anspruch, komplexe dynamische Prozesse zu erfassen und zu formalisieren, nicht gerecht werden. Die Antwort auf diese Herausforderung waren sogenannte *Parallel Constraint Satisfaction* (PCS)- Netzwerke (z.B. Read et al., 1997; Spellman und Holyoak, 1992). In entsprechenden Forschungsarbeiten finden sich beide bedeutenden und miteinander verknüpften Stränge der Konsistenzforschung, nämlich die Gestaltpsychologie und die kognitiven Konsistenztheorien, wieder. So ermöglichen PCS-Netzwerke nicht nur eine rechenbetonte Umsetzung der gestaltpsychologischen Prinzipien (z.B. Read et al., 1997; Spellman und Holyoak, 1992), sondern auch eine Modellierung der aus kognitiven Konsistenztheorien bekannten Phänomene (z.B. Monroe und Read, 2008; Read und Miller, 1994; Shultz und Lepper, 1996; Simon, Snow et al., 2004). Zudem wurde der ehemals vage Begriff der Konsistenz bzw. Kohärenz formal definiert und es wurde aufgezeigt, weshalb ein konnektionistischer Algorithmus am besten geeignet ist, um Kohärenz zu messen (Thagard und Verbeurgt, 1998). Da PCS-Modelle folglich einen wichtigen Fortschritt in der Konsistenzforschung markieren, sollen im Folgenden sowohl ihre Struktur und Funktionsweise erläutert sowie einige Modelle im Überblick vorgestellt werden.

PCS-Modelle gehören zur größeren Klasse der konnektionistischen Netzwerke, die insbesondere durch die Arbeiten von McClelland und Rumelhart bekannt wurden (McClelland und Rumelhart, 1981; McClelland, Rumelhart & The PDP Research Group, 1986; Rumelhart, McClelland & The PDP Research Group, 1986). Diese Autoren sprechen selbst von *Parallel Distributed Processing* (PDP)-Modellen, um die Kernaspekte dieses Ansatzes – paralle Verarbeitung, verteilte Repräsentation, generelle Prozesse – hervorzuheben (McClelland und Rumelhart, 1986). Als zentrales Charakteristikum von PDP-Modellen nennen McClelland, Rumelhart und Hinton (1986) "that information processing takes place through the interactions of a large number of simple processing elements called units, each sending excitatory and inhibitory signals to other units" (S. 10). Durch die Annahmen der Vernetztheit und Parallelität wird der Tatsache Rechnung getragen, dass häufig viele Informationen, die sich gegenseitig beeinflussen und damit Beschränkungen – sogenannte *constraints* – darstellen, gleichzeitig zu beachten sind. Was die Einheiten repräsentieren, ist eine inhaltliche Fragestellung. So können sie etwa Wörter, Ziele oder auch Entscheidungsalternativen darstellen. Gestützt wird diese Behauptung dadurch, dass McClelland und Rumelhart bereits 1986 eine Vielzahl an Beispielen für PDP-Modelle aufführen konnten (McClelland; Rumelhart & The PDP Research Group, 1986; Rumelhart, McClelland et al., 1986). Diese verschiedenen Modelle werden als Spezialfälle eines allgemeinen Modells, das Rumelhart, Hinton und McClelland (1986) vorstellen, betrachtet. Wichtige Aspekte eines PDP-Modells sind demnach (1) ein Set verarbeitender Einheiten, (2) ein Aktivationsstatus des Systems bzw. der einzelnen Einheiten, (3) eine Output-Funktion für jede Einheit, (4) Konnektivität der einzelnen Einheiten, (5) eine Regel zur Berechnung des Inputs, (6) eine Aktivationsregel zur Berechnung des Aktivationslevels, (7) eine Lernregel und (8) eine Umgebung, in der das System arbeitet. Die Autoren führen jeden der Aspekte näher aus – an dieser Stelle sei lediglich auf drei bedeutende Details hingewiesen: Erstens können sich die Einheiten danach unterscheiden, ob sie selbst ganze Konzepte repräsentieren oder ob das Aktivationsmuster mehrerer Einheiten erst dieses Konzept darstellt. Im ersten Fall spricht man von lokaler, im zweiten Fall von verteilter Repräsentation. Hinton, McClelland und Rumelhart (1986) favorisieren zwar Modelle mit verteilter Repräsentation, räumen aber ein, dass auch lokale Modelle wie das *Interactive Activation Model* (McClelland und Rumelhart, 1981) ähnliche Vorzüge haben können. Zudem können die Modelle danach unterschieden werden, ob sie neben Input-Einheiten, die externalen Input erhalten, und Output-Einheiten, die Signale entsenden, weitere Einheiten, sogenannte *hidden units*, besitzen. Diese Einheiten

werden zwischen den Input- und Output-Einheiten angeordnet und sind nicht sichtbar. Eine zweite wichtige Unterscheidung, auf die Rumelhart, Hinton et al. (1986) hinweisen, bezieht sich auf die Konnektivität zwischen Einheiten. Grundsätzlich werden exzitatorische Verbindungen, in denen zwei Einheiten in positiver Relation (sich unterstützen, zusammengehören etc.) zueinander stehen, von inhibitorischen Verbindungen, in denen zwei Einheiten in negativer Relation zueinander stehen (sich widersprechen, nicht zusammengehören etc.) differenziert. Die Summe der Inputs, die eine Einheit von anderen, mit ihr verbundenen Einheiten erhält, kann nun entweder über alle Einheiten hinweg oder separat für Einheiten mit exzitatorischer und inhibitorischer Verbindung berechnet werden. Drittens ist der Hinweis der Autoren hervorzuheben, dass Lernen erst durch die Modifikation der Verbindungen erzielt wird. Dies ist nicht zu verwechseln mit der Veränderung der Aktivation der Knoten. Diese Aspekte der PDP-Modelle geben ihnen bedeutende Eigenschaften wie die Fähigkeiten, trotz Fehler in dem Informationsmuster die richtige Lösung zu generieren, unvollständige Muster zu ergänzen und Muster wiederherzustellen (McClelland, Rumelhart & Hinton, 1986). Ein weitere wichtige Eigenschaft von PDP-Modellen ist, dass sie in ihrem Aufbau und ihrer Funktionsweise Netzwerken aus Neuronen ähneln, wobei explizit darauf hingewiesen wird, dass keine Modellierung von Neuronen angestrebt wird, sondern eine durch Neuronen inspirierte Modellierung kognitiver Prozesse (McClelland, Rumelhart & Hinton, 1986). Die von Rumelhart, Hinton et al. (1986) vorgenommene Einteilung in bottom-up-, top-down- und interaktive Systeme kann dahingehend vergröbert werden, dass zwischen Modellen ohne Feedback und mit Feedback differenziert wird (vgl. Read et al., 1997). Feedback kann dadurch realisiert werden, dass sich Einheiten wechselseitig beeinflussen, die Verbindung zwischen ihnen also bidirektional ist. Nach Read et al. (1997) kann hinzugefügt werden, dass zudem die Aktivationen der Knoten über mehrere Durchgänge hinweg aktualisiert werden, während in *feed-forward*-Netzwerken nur ein Durchgang erfolgt. *Feedback*-Netzwerke werden auch als *Parallel Constraint Satisfaction* (PCS)-Modelle bezeichnet, da sie eine parallele Befriedigung multipler *constraints* ermöglichen.[3] Das Finden dieser Lösung erstreckt sich über mehrere *Updating*-Durchgänge mit entsprechenden Feedbacks, so dass diese Modelle dynamisch sind. Smith (1996) ergänzt bezüglich der Klassifikation der Netzwerke, dass sowohl *feed-forward*- als auch *feedback*-Netzwerke selbstverständlich ein Teil bzw. ein Modul eines

[3] Zu unterscheiden ist zwischen *soft constraints*, deren Erfüllung wünschenswert ist, und *hard constraints*, deren Erfüllung unverzichtbar ist (Shultz und Lepper, 1996). In den folgenden Modellen liegen *soft constraints* vor.

größeren Systems sein können. Da PCS-Modelle zu den PDP-Modellen gezählt werden können, könnte es aufgrund der Begrifflichkeiten zu einer Fehlinterpretation kommen. So bedeutet die Zugehörigkeit zur Klasse der PDP-Modelle nicht zwangsläufig, dass die Repräsentationen verteilt (*distributed*) sind (für eine abweichende Sichtweise vgl. Smith, 1996). Vielmehr weisen die im Folgenden erläuterten Modelle lokale Repräsentationen auf. Read et al. (1997) begründen den damit verbundenen Verlust an Vorteilen vor allem mit der besseren Beantwortbarkeit der jeweils zugrunde liegenden Fragestellungen.

Die Relevanz und der Nutzen von PCS-Modellen für die Beschreibung und Erklärung kognitiver Prozesse werden an der Vielzahl an Anwendungsbeispielen deutlich. Das Spektrum reicht von der Wortwahrnehmung (McClelland und Rumelhart, 1981) über die Bewertung von Erklärungen (Thagard, 1989) und analogisches Schließen (Holyoak und Thagard, 1989) hin zur sozialen Wahrnehmung (Read und Miller, 1998a). Mittels auf Basis des PCS-Ansatzes entwickelter Programme wurden hierbei sowohl bekannte Forschungsergebnisse (z.B. Kunda und Thagard, 1996; Monroe und Read, 2008; Read und Miller, 1994) als auch lebensnahe Ereignisse (z.B. Spellman, Ullman & Holyoak, 1993; Thagard, 1989; Thagard, 2003) erfolgreich simuliert. Von besonderem Einfluss für die PCS-Forschung im Allgemeinen und die im Folgenden exemplarisch dargestellten Modelle im Speziellen war das bereits erwähnte *Interactive Activation Model* der Wortwahrnehmung (IAM) von McClelland und Rumelhart (1981). Aus diesem Grund soll anhand dieses Modells exemplarisch die Struktur und Funktionsweise eines PCS-Modells erläutert werden. Das IAM besteht aus Knoten, die in drei verschiedenen Ebenen angeordnet und über Verbindungen miteinander verknüpft sind. Auf der ersten Ebene repräsentieren die Knoten visuelle Merkmale von Buchstaben, auf der zweiten Ebene Buchstaben selbst und auf der dritten Ebene Wörter. Verbindungen können prinzipiell exzitatorisch oder inhibitorisch sein und bestehen zum einen zwischen angrenzenden Ebenen und zum anderen innerhalb einer Ebene. Eine Verbindung zwischen zwei Knoten aus angrenzenden Ebenen ist exzitatorisch, wenn diese beiden Knoten miteinander konsistent sind. Ein Beispiel hierfür ist, dass der Buchstabe *A* an der ersten Position des Wortes *Apfel* steht. Dementsprechend ist eine Verbindung zwischen zwei Knoten aus angrenzenden Ebenen inhibitorisch, wenn die Knoten inkonsistent ist. So sind etwa der Knoten für den Buchstaben *A* an der ersten Stelle eines Wortes und das Wort *Buch* inkonsistent. Die Verbindungen innerhalb der Buchstaben- und Wortebene sind stets inhibitorisch, da nur ein Buchstabe bzw. ein Wort richtig sein kann. Alle Verbindungen sind zudem gewichtet. Von besonderer Bedeutung ist, dass der Input nicht nur

von der niedrigeren zur nächsthöheren Ebene verläuft, sondern dass die Verbindungsrichtung auch von Knoten einer höheren Ebene zu Knoten der nächstniedrigeren Ebene verläuft. Hierdurch wird Feedback implementiert. Bezüglich der Funktionsweise ist zunächst wichtig, dass die Verarbeitung parallel erfolgt und dass die Kommunikation innerhalb des Netzwerks durch Aktivationsausbreitung erfolgt. Wird etwa ein visueller Input gegeben, werden die Merkmale einzelner Buchstaben je nach visueller Qualität entdeckt und aktivieren korrespondierende Buchstaben auf der nächsthöheren Ebene. Inkonsistente Buchstaben werden inhibiert. Der Input (*Net Input*), den ein bestimmter Knoten *j* erhält, kann nach folgender Formel berechnet werden (Rumelhart, Hinton et al., 1986, S. 71):

$$net_j = \sum w_{ij} a_i$$

Hierbei repräsentieren *i* die mit *j* verbundenen Knoten, a_i stellt die Aktivation eines Knotens *i* und w_{ij} das Gewicht zwischen den Knoten *i* und *j* dar. Folglich wird bei der Berechnung des Net Inputs nicht zwischen exzitatorischen und inhibitorischen Verbindungen getrennt.

Je nach dem, ob der Net Input positiv oder negativ ist, wird die Aktivation des Knotens erhöht oder erniedrigt. Zu beachten ist, dass aufgrund der Implementierung von Feedback sowohl Knoten der nächsthöheren als auch Knoten der nächstniedrigeren Ebene Input geben und somit interagieren (McClelland und Rumelhart, 1981). Das Aktivationsniveau eines Knotens *j* lässt sich folgendermaßen berechnen (Rumelhart, Hinton et al., 1986, S. 72):

$$a_j(t+1) = a_j(t)(1-\Theta) + \begin{cases} net_j(\max - a_j(t)) & net_j > 0 \\ net_j(a_j(t) - \min)) & otherwise \end{cases}$$

Mit *max* und *min* werden die maximal und minimal mögliche Aktivation eines Knotens bezeichnet. Bei Knoten, die keinen Input erhalten, fällt das Aktivationsniveau mit einer bestimmten *decay*-Rate Θ auf ein Ruheniveau ab.

Die Aktivation breitet sich dann durch das Netzwerk aus, indem Buchstaben konsistente Wörter aktivieren und inkonsistente Wörter inhibieren. Wie bereits angesprochen, wirkt deren Aktivation auf die nächstniedrigere Ebene zurück, so dass der Prozess dynamisch ist. Konvergieren die Aktivationen zu einem stabilen Muster, so bedeutet dies, dass ein Wort wahrgenommen wird. Ein Wort und dessen Buchstaben mit seinen visuellen Merkmalen weisen dann im Vergleich zu den übrigen Knoten schließlich die höchste

Aktivation auf (McClelland und Rumelhart, 1981). Ein an die Originalarbeit angelehntes vereinfachtes Beispiel mag die Funktionsweise des Netzwerks illustrieren: Nehmen wir an, einer Person wird das Wort TIER gezeigt. Die Qualität des visuellen Inputs ist aber beschränkt – während die Person die Buchstaben TIE einwandfrei identifizieren kann, ist sie sich bezüglich des letzten Buchstabens nicht sicher, ob der Buchstabe ein R oder ein K darstellt. Es werden aufgrund der Buchstaben verschiedene Wörter aktiviert, der Einfachheit halber sei die Fülle möglicher Wörter auf die vier Wörter TIER, TIEF, TANK und TEER beschränkt. Da die Buchstaben TIE identifiziert wurden, werden die Wörter TIER und TIEF eine höhere Aktivation erhalten als die anderen beiden Wörter. Als vierter Buchstabe kommen lediglich ein R oder ein K in Frage, weshalb das Wort TIER stärker aktiviert wird. Dadurch wiederum erhält der Buchstabe R Feedback. Auch wenn dies zur Folge hat, dass das Wort TEER ebenfalls stärker aktiviert wird, behält das Wort TIER die Dominanz und wird als Lösung ausgegeben. Die Funktionsweise des Modells verdeutlicht das Prinzip der *parallel constraint satisfaction*, das diesen Modellen den Namen gab: "In a feedback or parallel constraint satisfaction network, activation passes around symmetrically connected nodes until the activation of all the nodes asymptotes or 'relaxes' into a state that satisfies the constraints among the nodes" (Read et al., 1997, S. 26). Dieses Verhalten eines Systems lässt sich gemäß Hopfield (1982) als Minimierung von Energie verstehen. Hopfield machte zudem den Vorschlag, die Energie des Systems als Landschaft zu visualisieren. Eine anschauliche Darstellung geben Read et al. (1997): In der Energielandschaft repräsentieren die einzelnen Dimensionen die Aktivationen der Knoten und eine weitere Dimension gibt die Energie des Systems an. Folglich bestimmen die sich aus den verschiedenen Aktivationen der Knoten ergebenden Energie-Zustände des Systems die Oberfläche der Landschaft. Da die Aktivation in den Tälern am niedrigsten ist, besteht das Ziel von PCS-Prozessen darin, ein solches zu erreichen, um die Energie zu minimieren und so Konsistenz zu maximieren. Solche Minima werden auch als Attraktoren bezeichnet, da sie quasi „anziehend" wirken. Für die Berechnung der Höhe der Energie schlägt Hopfield (1982) eine Formel vor, deren Inverse ein Maß für die Harmonie oder *Goodness-of-fit* des Systems ist (Rumelhart, Smolensky et al., 1986). Vor dem Hintergrund der bisherigen Ausführungen und der vorgestellten Arbeitsdefinition wird in der vorliegenden Arbeit die Bezeichnung Konsistenz gewählt, die sich vereinfacht[4] folgendermaßen berechnen lässt:

[4] Die Vereinfachung besteht darin, dass im Gegensatz zu Rumelhart, Smolensky et al. (1986) der neue Input nicht berücksichtigt wird.

$$K(t) = \sum_i \sum_j w_{ij} a_i(t) a_j(t)$$

Die Konsistenz *K* eines Systems wird folglich berechnet, indem das Produkt der Aktivationen eines jeden Knotenpaares gebildet und mit dem Gewicht zwischen den beiden Knoten multipliziert wird. Die resultierenden Produkte werden für alle Knotenpaare aufsummiert. Jede lokale Änderung der Aktivation eines Knotens dient damit der Veränderung des globalen Systemzustandes in Richtung maximaler Befriedigung der *constraints* (Rumelhart, Smolensky et al., 1986). Hinsichtlich der Maximierung von Konsistenz bzw. der Minimierung von Energie gilt zu beachten, dass "such minima are not guaranteed to be global minima of the entire system, but may instead be local" (Read et al., 1997, S. 30).

An der Popularisierung von PCS-Modellen für verschiedene Bereiche der Psychologie hat Paul Thagard maßgeblichen Anteil gehabt, der seinerseits die breite Anwendbarkeit von PCS-Modellen betont (Thagard, 2000). Im Jahr 1989 veröffentlichte er die *Theory of Explanatory Coherence*, die sich auf die Akzeptanz bzw. Ablehnung sowohl wissenschaftlicher Hypothesen als auch alltäglicher Schlussfolgerungen bezieht (Thagard, 1989). Als zentral für diese Entscheidung sieht er das Ausmaß an *explanatory coherence* an. Diese liegt laut Thagard vor, wenn bestimmte Propositionen aufgrund erklärender Beziehungen verbunden sind: "Propositions P and Q cohere if there is some explanatory relation between them" (S. 436). Dementsprechend ist laut Thagard Inkohärenz dadurch gekennzeichnet, dass sich zwei Propositionen widersprechen oder ihre Erklärungen nicht mit dem Hintergrundwissen vereinbar sind. Er postuliert sieben Prinzipien, anhand derer die Kohärenz einer Erklärung beurteilt werden kann: (1) Symmetrie, (2) Erklärung, (3) Analogie, (4) Priorität der Daten, (5) Widerspruch, (6) Akzeptabilität und (7) Systemkohärenz. In späteren Arbeiten (z.B. Thagard, 2000) bezeichnet das sechste Prinzip Wettbewerb und das siebte Prinzip die Akzeptabilität. Diese allgemeinen Prinzipien kommen in ähnlicher Form auch bei der Beurteilung weiterer Kohärenzarten, etwa der *analogical*, der *deductive*, der *perceptual*, der *conceptual* sowie der *deliberative coherence* zum Tragen (Thagard, 2000). Diese Prinzipien implementiert Thagard (1989) in einem konnektionistischen Netzwerk, genauer gesagt einem PCS-Netzwerk, mit dessen Hilfe eine Evaluation der Hypothesen und eine Bestimmung der Gesamtkohärenz des Systems ermöglicht werden. Hierbei stellen die Hypothesen und die zu erklärende Evidenz die Knoten dar. Durch einen zusätzlichen Knoten, die *special evidence unit*, wird die Priorität der Daten implementiert. Das zugrunde

liegende Simulationsprogramm trägt den Namen ECHO: *Explanatory Coherence by Harmany Optimization*, wobei mit *Harmany Optimization* die Maximierung der Systemkohärenz ausgedrückt wird. Zur Berechnung der Gesamtkohärenz verwendet Thagard obige Harmonie-Formel. Thagard unterstreicht die Bedeutung eines besonderen Kennzeichens der PCS-Netzwerke, nämlich dass die Aktivation eines Knotens jeden anderen Knoten im Netzwerk beeinflussen kann. Dadurch wird implementiert, dass eine Struktur holistisch wahrgenommen wird. Basierend auf diesem Netzwerk können für eine Evaluation wichtige Aspekte wie z.B. die Breite einer Erklärung, dass die Erklärung ihrerseits erklärt wird, die Präferenz für einfache Erklärungen und das Bestehen von Wettbewerb berücksichtigt werden. Dass jeder dieser Aspekte eine eigene Bedeutung für die Gesamtkohärenz hat und dass sie auch in alltäglichen Entscheidungen relevant sind, haben Read und Marcus-Newhall (1993) gezeigt. In einer Weiterentwicklung von ECHO wurde das Modell um einen Lernmechanismus und die Fähigkeit, sich dynamisch an die Umwelt anzupassen, erweitert (Wang, Johnson & Zhang, 1998). Holyoak hat in Zusammenarbeit mit Thagard eine Theorie des analogischen Mappings aufgestellt (Holyoak und Thagard, 1989). Unter Berücksichtigung struktureller, semantischer und pragmatischer Beschränkungen soll ein angemessenes Mapping erreicht werden. Sie bilden das Analogie-Problem in einem PCS-Netzwerk ab, wobei die Knoten zum einen Paarungen der Analogie-Quelle und des Analogie-Ziels und zum anderen sogenannte *special units*, die die semantische Ähnlichkeit von Quelle und Ziel bzw. die Bedeutung einzelner Konzepte abbilden, repräsentieren. Je nach dem, ob die Paarungen gemeinsam auftreten können oder nicht, bestehen exzitatorische oder inhibitorische Verbindungen zwischen den Knoten. Simuliert wird das analogische Mapping mit dem Programm ACME (*Analogical Constraint Mapping Engine*). Im Gegensatz zu ECHO basiert der Updating-Algorithmus hier nicht auf der Formel von McClelland und Rumelhart (1981), sondern auf derjenigen von Grossberg (1978). Die Autoren begründen dies mit einer höheren Effektivität der Formel von Grossberg auch für komplexere Probleme. Bei dieser Formel werden der exzitatorische und der inhibitorische Input separat aufsummiert (Grossberg, 1978).[5,6] Spellman und Holyoak (1992) gelang der wichtige Nachweis, dass die Annahmen

[5] Rumelhart, Hinton et al. (1986, S. 71) stellen die typische Form dieser Aktivationsregel anschaulich dar:
$a_j(t+1) = a_j(t)(1-A) + (B - a_j(t))net_{ej}(t) - (a_j(t) + C)net_{ij}(t)$
Hierbei repräsentieren net_{ej} (t) und net_{ij} (t)den exzitatorischen bzw. inhibitorischen Input zum Zeitpunkt t, A entspricht dem Zerfallsparameter und B und C repräsentieren die maximal mögliche Exzitation bzw. Inhibition.

[6] Weitere Arbeiten, die sich auf die Formel von Grossberg (1978) stützen, stammen insbesondere von Read und Kollegen (Read und Marcus-Newhall, 1993; Read und Miller, 1993, 1994; Read et al., 1997). In einer späteren Arbeit aus dem Jahr 1998 greifen Read und Miller hingegen auf den oben genannten Algorithmus von McClelland und Rumelhart (McClelland und Rumelhart, 1981; Rumelhart, Hinton et al., 1986) zurück.

dieser Theorie auch unter Verwendung eines Fallbeispiels, nämlich des Zweiten Golfkriegs, ihre Gültigkeit haben. Des Weiteren führen sie die von Holyoak und Thagard formulierte Ähnlichkeit von PCS-Prozessen und gestaltpsychologischen Prinzipien näher aus, insbesondere das Wirken dynamischer Kräfte mit dem Ziel, Konsistenz zu maximieren und so eine Lösung eines bestimmten Problems (z.B. visuelle Wahrnehmung, Entscheidung) zu erzielen. Die Autoren betrachten PCS-Modelle als deutlichen Fortschritt, wenn sie schließen: "Models of parallel constraint satisfaction (…) provide computational realizations of core Gestaltist ideas concerning the emergence of coherence" (S. 926). Eine PCS-Theorie zur Eindrucksbildung schlagen Kunda und Thagard (1996) vor und simulieren bekannte diesbezügliche Phänomene mittels des Programms IMP (*IMPression formation*). Die Knoten repräsentieren Konzepte wie ein Verhalten (hat jemanden gestoßen), mögliche Interpretationen (gewaltsamer Schubs vs. wohlwollender Schubs) und Kategorien von Stereotypen (weiße Hautfarbe vs. dunkle Hautfarbe). Die Verbindungen repräsentieren die Überzeugungen des Beobachters bezüglich der Zusammenhänge der Konzepte. Ein weiterer Knoten verleiht beobachtetem Verhalten einen höheren Status und grenzt es so von Inferenzen ab. Evaluiert wird bei diesem Modell die *conceptual coherence* (Thagard und Kunda, 1998). Thagard und Kunda verstehen die *explanatory coherence*, die *analogical coherence* und die *conceptual coherence* als Beispiele für ein übergreifendes Kohärenzproblem. Die Mechanismen, um die jeweilige Kohärenz zu erreichen, nämlich konzeptuelle Integration, Erklärung und Analogie, erachten sie als Wege der sozialen Wahrnehmung, wobei sich diese Wege nicht ausschließen. Diesen Versuch, verschiedene Kohärenztheorien in einen einheitlichen Kohärenzansatz zu integrieren, intensiviert Thagard (2000) explizit und verdeutlicht damit, für wie allgemeingültig er das Prinzip der Kohärenz bzw. Konsistenz hält. Thagards in diesem Zusammenhang geäußerter Wunsch, die Kognition und die Emotion in einer Theorie zu verbinden, kommt auch in seiner Arbeit aus dem Jahr 2003 zum Ausdruck, in der er die *emotional coherence* einführt. In dieser Arbeit legt Thagard dar, dass bei der Urteilsbildung nicht nur rein kognitive Inferenzen gezogen werden, sondern auch emotionale Prozesse eine entscheidende Rolle spielen können. Dies veranschaulicht er am Freispruch von O. J. Simpson, der wegen Mordes an seiner Ex-Frau und ihrem Freund angeklagt war. Entsprechende Simulationen führt er mit dem Programm HOTCO2 durch, das er als Weiterentwicklung von ECHO begreift. In dem zugrunde liegenden Netzwerk können die in Form von Knoten repräsentierten Fakten und Hypothesen nicht nur hinsichtlich ihrer Erklärungskraft variieren, wie das bei ECHO der Fall war, sondern zudem

hinsichtlich ihrer Valenz. Dies wurde durch einen zusätzlichen speziellen Valenz-Knoten erreicht. Eine gezielt für den Entscheidungsbereich konzipierte Theorie ist die *Coherence Theory of Decision*, die Thagard und Millgram (1995) gemeinsam mit dem zugehörigen Programm DECO (*DEliberative COherence*) vorstellen. Aufgrund ihres direkten Bezugs zur Entscheidungspsychologie wird sie in dem entsprechenden späteren Kapitel dargestellt. Neben Thagard haben auch Read und Miller (Read und Miller, 1993; Read und Miller, 1994; Read und Miller, 1998a; Read et al., 1997) wichtige Beiträge zur Erforschung und Verbreitung von PCS-Modellen geleistet. Nicht zuletzt deutlich wird dies an dem von ihnen herausgegebenen Buch *Connectionist Modeling of Social Reasoning and Social Behavior* aus dem Jahr 1998. Sie wandten PCS-Prozesse auf den Bereich der sozialen Interaktion an. Ausgehend von der Annahme, dass PCS-Prozesse zentral für die Generierung kohärenter Repräsentationen sozialer Interaktionen sind, erläutern Read und Miller (1993) ihr zweistufiges Modell für die Erklärung sozialer Interaktionen. In einem ersten Schritt werden zunächst relevante Konzepte aktiviert und in einem Netzwerk verbunden. Im zweiten Schritt kommen dann PCS-Prozesse zum Tragen, wobei sich die Autoren auf Thagards (1989) *Theory of Explanatory Coherence* beziehen. Unter Rückgriff auf ein Fallbeispiel zeigen Read und Miller auf, dass das Bilden eines kohärenten Eindrucks und das Finden einer kohärenten Erklärung für ein Verhalten mit PCS-Prozessen zu erklären sind. Ein elaborierteres Modell der sozialen Wahrnehmung stellen Read und Miller im Jahr 1998 vor, nämlich ein *Interactive Activation and Competition* (IAC)-Modell: *Social Dynamics*. Sie gründen ihre Arbeit explizit auf das IAM von McClelland und Rumelhart (1981) und heben die im Vergleich zu ihren vorherigen Arbeiten (Read und Miller, 1993) und derjenigen von Kunda und Thagard (1996) detailliertere Charakterisierung der Struktur und der Repräsentationen hervor. Darüber hinaus erörtern sie wichtige Mechanismen des Lernens und Eigenschaften von PDP-Modellen wie die Ergänzung fehlender Informationen und die Konstruktion neuer Kombinationen, die sie dann auf den Bereich der sozialen Wahrnehmung beziehen. Unter Berücksichtigung der deutlichen Bezüge zur Gestaltpsychologie und zu den kognitiven Konsistenztheorien betonen die Autoren, dass PCS-Modelle in der Lage sind, Charakteristika der sozialen Wahrnehmung wie Ambiguität und das Vorliegen vieler unterschiedlicher Informationen adäquat abzubilden. So kommen sie zu dem Schluss: “Although connectionist models provide strong conceptual ties to our past, they also hold great promise for our future” (S. 65). Ein auf ECHO basierendes PCS-Modell der Einstellungsänderung, das *Coherence Model of Cognitive Consistency* (Co3), stellen Spellman et al. (1993) vor. Sie

zeigen anhand eines tatsächlichen Ereignisses, nämlich anhand des Zweiten Golfkriegs, dass die bereits in früheren Konsistenztheorien formulierten Einstellungsänderungen mit dem Ziel kognitiver Konsistenz durch PCS-Prozesse erklärt werden können. In einer aktuellen Arbeit unternehmen Monroe und Read (2008) den Versuch, ein einheitliches Modell für die gesamte Einstellungspsychologie zu entwerfen. Sie möchten folglich nicht ein bestimmtes Phänomen, z.B. Einstellungsänderung, sondern eine Vielzahl an Phänomenen erklären. Der Ausgangspunkt ihrer Arbeit ist die Fokussierung auf die Struktur einer Einstellung, zu deren wichtigen Aspekten sie die Vernetzung verschiedener Konzepte und das Streben nach Konsistenz sehen. "A more sophisticated view of consistency" (S. 735) sehen sie in der PCS-Perspektive, die mit entsprechenden Modellen einhergeht. Da sie konnektionistische Netzwerke für besonders geeignet halten, um Einstellungsphänomene zu modellieren, liegen auch ihrem Modell, dem *Attitudes as Constraint Satisfaction* (ACS)-Modell, PCS-Prozesse zugrunde. Ohne auf die detaillierte Architektur des Modells einzugehen, sei lediglich erwähnt, dass die Autoren erfolgreich verschiedene basale Eigenschaften, z.B. die Polarisierung von Einstellungen, *motivated reasoning* und Einstellungsänderung, simulieren konnten. Des Weiteren sei auf eine im Vergleich zu den vorherigen Modellen interessante Weiterentwicklung hingewiesen: Im ACS-Modell sind die Gewichte der Verbindungen nicht fixiert, sondern können modifiziert werden, sodass Lernen möglich ist. Monroe und Read nennen weitere Punkte, die in künftigen Modellen berücksichtigt werden könnten, etwa die Bedeutung automatischer und kontrollierter Prozesse, und verdeutlichen so erneut ihren Anspruch, ein umfassendes und übergreifendes Modell zu entwickeln. Einen derartigen Anspruch haben sich auch Read et al. (2010) bei der Konzeption ihres neuronalen Netzwerkmodells der Persönlichkeit gesetzt. Sie verfolgen hierbei das Ziel, strukturelle und dynamische Ansätze der Persönlichkeitsforschung in einem Modell zu vereinen. Die Autoren betonen, dass dieses Modell sehr konkret und realistisch ist, was die Simulation einer Vielzahl an spezifischen Phänomenen der Persönlichkeitsforschung erlaubt. Neben der Formulierung von PCS-Netzwerken für bestimmte Anwendungsbereiche bzw. Fallbeispiele werden sie auch zur gezielten Simulation konsistenztheoretischer Befunde eingesetzt: Read und Miller (1994), deren Beeinflussung durch Bob Abelson bereits angesprochen wurde (siehe Kapitel 2.2.2), simulieren mittels ECHO erfolgreich Befunde der Balancetheorie und der Dissonanztheorie. Darüber hinaus zeigen sie auf, wie weitere, von Abelson vorgeschlagene Wege, um Konsistenz zu erreichen, in einem konnektionistischen Netzwerk implementiert werden könnten. Die Autoren machen zudem auf einen wichtigen Aspekt

aufmerksam, nämlich dass die Evaluation der Kohärenz eines sehr großen Netzwerks aufgrund kognitiver Begrenzungen lokal und nicht global ist. Das heißt, dass Subsets evaluiert werden, wenn die Komplexität eines Netzwerks zu groß wird. Shultz und Lepper (1996) erläutern ihr *Consonance Model*, das eine Formalisierung der Dissonanztheorie als PCS-Netzwerk darstellt. Mittels dieses Modells konnten die Autoren nicht nur bestehende dissonanztheoretische Befunde simulieren und erklären, sondern auch neue Phänomene erfolgreich vorhersagen. Die Tatsache, dass mittels ihres Modells die empirischen Daten sogar besser abgebildet werden konnten als mittels der Dissonanztheorie, spricht für Shultz und Lepper für den Mehrwert von PCS-Modellen. Folglich halten sie die Ausweitung ihres Modells auf weitere Konsistenztheorien für möglich.

Zusammenfassend ist festzuhalten, dass die Entwicklung und Popularisierung von PCS-Modellen der Konsistenzforschung neuen Auftrieb gab. Die bereits diskutierten Unzulänglichkeiten der Gestaltpsychologie und der kognitiven Konsistenztheorien wurden beseitigt: Das zuvor eher abstrakte gestaltpsychologische Prinzip der Tendenz zur guten Gestalt konnte formal dargestellt werden. Hierbei wurde zugleich ein wichtiger Forschritt hinsichtlich des Umgangs mit Komplexität erzielt, indem erstens dynamische Prozesse abgebildet werden konnten, zweitens die Verarbeitung widersprüchlicher Informationen erklärt werden konnte und drittens größere Strukturen betrachtet werden konnten. Holyoak und Thagard (1989) berichten etwa von einem Netzwerk mit 214 Knoten und 1783 Verbindungen. Zudem wurde nun Konsistenz messbar. Die Vielzahl an PCS-Modellen, die in den vergangenen 25 Jahren für verschiedenste Fragen der Psychologie entwickelt wurde, verdeutlicht die breite Anwendbarkeit von PCS-Prozessen. Der Trend, mehr oder weniger breite und übergreifende Modelle zu entwickeln (z.B. Monroe und Read, 2008; Thagard, 2000) scheint angesichts der Befunde nur konsequent (vgl. auch Simon, Snow et al., 2004). Hierbei stellen die von verschiedenen Autoren (z.B. Kunda und Thagard, 1996; Monroe und Read, 2008) angeregten Weiterentwicklungen, etwa die Berücksichtigung von automatischen und kontrollierten Prozessen, von Motivation und Affekt, interessante Herausforderungen dar.

Nachdem die allgemeine Entwicklung der Konsistenzforschung von den Anfängen in der Gestaltpsychologie bis zu den neueren PCS-Ansätzen nachgezeichnet worden ist, wird das Thema der Konsistenz in den folgenden Ausführungen speziell für einen Bereich betrachtet: die Entscheidungsforschung.

2.3 DAS STREBEN NACH KONSISTENZ: ENTSCHEIDUNGSTHEORIEN UND -MODELLE

Dass das Bedürfnis nach Konsistenz eine zentrale Rolle in der Entscheidungsfindung einnimmt, wurde bereits in frühen Konsistenztheorien (Festinger, 1957; Festinger, 1964) formuliert. Seit Mitte der 80er Jahre findet dieser Aspekt explizite Berücksichtigung in verschiedenen Entscheidungstheorien und -modellen (z.B. Glöckner und Betsch, 2008b; Holyoak und Simon, 1999; Montgomery, 1989; Montgomery und Willén, 1999; Pennington und Hastie, 1986; Pennington und Hastie, 1988; Thagard und Millgram, 1995). Entscheidende Fortschritte in dieser Entwicklung wurden durch die Erforschung und Popularisierung von PCS-Modellen erzielt. Im folgenden Kapitel werden zunächst wichtige Ansätze der Entscheidungsforschung, in denen das Thema Konsistenz besondere Bedeutung erfährt, erläutert (Kapitel 2.3.1). In Kapitel 2.3.2 wird ausführlich auf das PCS-Modell für probabilistische Inferenzentscheidungen von Glöckner und Betsch (2008b) eingegangen, da die vorzustellenden empirischen Forschungsarbeiten auf diesem Modell gründen.

2.3.1 Bisherige Ansätze

DIE ENTSCHEIDUNGSFINDUNG ALS KONSTRUKTIONSPROZESS

Bereits im Jahr 1992 hielten Payne, Bettman und Johnson in ihrem Review der *Behavioral Decision Research* fest, dass ein neuer Trend in der Beschäftigung mit der *Konstruktion* von Urteilen und Entscheidungen besteht. Hinter diesem Trend steht die Annahme, dass eine Entscheidung nicht einfach offen gelegt wird, sondern in Abhängigkeit von Aufgaben-, Kontext- und Personenfaktoren konstruiert wird. Dies beinhaltet laut Payne et al. auch, dass Entscheidungsprobleme re-konstruiert werden, und zwar mit dem Ziel "to reduce either the amount of perceived conflict in the decision or the degree of complexity" (S. 101). Für die vorliegende Fragestellung ist insbesondere das erste Ziel entscheidend, da es das Streben nach Konsistenz impliziert. Auch Svenson (1996) ist der Ansicht, dass die Entscheidungsforschung ein Stadium erreicht hat, in der Änderungen der Problemrepräsentation als Möglichkeit zur Lösung von Entscheidungsproblemen bzw. -konflikten gesehen werden. Prinzipiell können solche Rekonstruktionsprozesse sowohl postdezisional als auch prädezisional stattfinden und sowohl Umstrukturierungen als auch Umwertungen beinhalten. Im Folgenden soll allerdings lediglich die Entscheidungsfindung selbst als Konstruktions-

prozess beleuchtet werden, nicht die nachträgliche private oder öffentliche Rechtfertigung einer Entscheidung. Folglich werden etwa Prozesse der postdezisionalen Dissonanzreduktion (Festinger, 1957; Festinger, 1964) ausgeklammert. Eine Vielzahl an Entscheidungstheorien und -modellen berücksichtigt den konstruktiven Charakter der Entscheidungsfindung in verschiedener Hinsicht. Brownstein (2003) gibt einen umfassenden Überblick über die verschiedenen Ansätze. Diese können sowohl nach der Art der Rekonstruktion als auch nach dem Stellenwert im Entscheidungsprozess unterschieden werden. So kann der Rekonstruktionsprozess lediglich als Umstruktierung des Entscheidungsproblems mit dem Ziel, eine für ein Entscheidungsproblem passende Entscheidungsregel zu finden, aufgefasst werden (Payne, Bettman, Coupey & Johnson, 1992; Payne et al., 1988). Er kann aber auch inhaltliche Umwertungen umfassen und hierbei vorrangig einem bestimmten Ziel dienen, z.B. der Vermeidung von Konflikt (Janis und Mann, 1977), dem Gewinnen von Sicherheit (Mills, 1968; Mills und O'Neal, 1971) oder dem Ermöglichen von Handeln (Beckmann und Kuhl, 1984). Im Unterschied zu diesen Theorien wird in einigen anderen Theorien dem Rekonstruktionsprozess eine zentrale Rolle beigemessen, indem er explizit als wesentlicher Bestandteil des Entscheidungsprozesses aufgefasst wird (Montgomery, 1983; Montgomery, 1989; Montgomery und Willén, 1999; Pennington und Hastie, 1986; Pennington und Hastie, 1988; Pennington und Hastie, 1992; Svenson, 1992; Svenson, 1996). Selbstverständlich können durch die Umwertungen auch in diesen Ansätzen weiterführende Ziele erfüllt werden, in Abgrenzung zur vorherigen Gruppe liegt aber das Hauptaugenmerk auf dem Rekonstruktionsprozess und dem damit verbundenen Streben nach Konsistenz. Im Folgenden wird nur diese letzte Gruppe an Theorien, in der der Aspekt der inhaltlichen Rekonstruktion zentral ist, betrachtet.

Ein erstes Beispiel für eine solche Theorie stellt das *Story Model* von Pennington und Hastie (1986; 1988; 1992) dar. Ausgehend von der Feststellung, dass verschiedene Aspekte juristischer Entscheidungen – wie die Verfügbarkeit einer Vielzahl an Informationen, deren ungeordnete und unvollständige Präsentation sowie das Bestehen von Interdependenz zwischen einzelnen Informationen – nicht adäquat durch bestehende Modelle abgebildet werden konnten, entwickelten sie das Story Model. Dessen zentrale Annahme besteht darin, dass Geschworene die verfügbare Evidenz zu einer Geschichte konstruieren, in der die einzelnen Informationen kausal verknüpft werden. Sie unterscheiden in ihrem Modell drei Komponenten: (1) die Konstruktion einer Geschichte, (2) die Repräsentation der Urteilsalternativen und (3) die Klassifikation der Geschichte. Im Prozess der Konstruktion wird auf

Basis der dargebotenen Informationen und auf Basis des Wissens über ähnliche Sachverhalte sowie über Charakteristika einer vollständigen Geschichte eine solche gebildet. Diese ist durch einen hierarchischen Aufbau gekennzeichnet, wobei das zentrale Ereignis auf der höchsten Ebene anzufinden ist. Da mehr als eine Geschichte konstruiert werden kann, bedarf es einer Einschätzung der Akzeptanz der Geschichte und der Konfidenz in diese. Hierzu werden die vier Prinzipien der Abdeckung (*coverage*), der Kohärenz (*coherence*), der Einzigartigkeit (*uniqueness*) und der Güte der Passung zwischen der Geschichte und der Urteilskategorie (*goodness of fit*) zugrunde gelegt. Es gilt zu beachten, dass eine stärkere Erfüllung dieser Prinzipien nicht per se zu höherer Akzeptanz und Konfidenz führt. Vielmehr konnten Pennington und Hastie in einem Experiment aus dem Jahr 1988 zeigen, dass die Konfidenz und die subjektive Wichtigkeit einzelner Informationen durch die Kohärenz der nicht gewählten Geschichte beeinflusst werden: Je höher diese ist, desto geringer sind Konfidenz und subjektive Wichtigkeit der Informationen, die Bestandteil der finalen Geschichte sind. Bei der Repräsentation der Urteilsalternativen geht es um das Lernen der einzelnen Urteilskategorien und ihrer Merkmale. Im Klassifikationsprozess wird schließlich bestimmt, welche Urteilskategorie der Geschichte bzw. dem kausalen Modell entspricht. Es wird also auf Basis der Geschichte eine Entscheidung getroffen. In verschiedenen Experimenten konnten Pennington und Hastie (1986; 1988; 1992) Unterstützung für ihre Annahmen finden. So konnten sie z.B. zeigen, dass die spontane Konstruktion einer Geschichte tatsächlich die Entscheidung determiniert sowie zu einer höheren Einschätzung der Wichtigkeit der für die jeweilige Geschichte relevanten Informationen führt. Dass sich eine Geschichte nicht nur aus der kausalen Verknüpfung von präsentierten Informationen ergibt, sondern dass diese tatsächlich *konstruiert* wird, wird aus der folgenden Aussage der Autoren deutlich: “Jurors typically inferred missing components for these structures when they were not contained in direct testimony at trial. Evidence not related to a story of what happened was systematically deleted from discussion” (Pennington und Hastie, 1986, S. 252). Obgleich das Story Model für juristische Entscheidungen konzipiert wurde, wird die Anwendung der zugrunde liegenden erklärungsbasierten Strategie auf weitere komplexe Entscheidungsaufgaben für möglich gehalten, wobei jeweils bereichsspezifische kausale Modelle konzipiert würden (Pennington und Hastie, 1986; Pennington und Hastie, 1988; Pennington und Hastie, 1992). Während in diesem Modell die Restrukturierung von Informationen im Vordergrund steht, werden im folgenden Ansatz sowohl Restrukturierungen als auch Reinterpretationen thematisiert.

Eine weitere Theorie, die den Aspekt der Konstruktion im Entscheidungsprozess betont, ist die *Search for a Dominance Structure* (SDS)-Theorie (Montgomery, 1983; Montgomery, 1989; Montgomery und Willén, 1999), deren Grundzüge sich folgendermaßen zusammenfassen lassen: Montgomery sieht den Entscheidungsprozess als die Suche nach einer sogenannten Dominanzstruktur an, das heißt einer Struktur, in der eine Alternative (eine) andere Alternative(n) dominiert. Genauer gesagt bedeutet dies, dass diese Alternative auf keinem Attribut den anderen Alternativen unterlegen ist und diesen auf mindestens einem Attribut überlegen ist. Die Suche nach Dominanz konzipiert Montgomery in vier Phasen: In der ersten Phase, dem *pre-editing*, finden sowohl eine Bewertung und Selektion von Attributen hinsichtlich ihrer Wichtigkeit als auch ein Screening der Alternativen statt. Nur die relevanten Attribute und Alternativen werden in der Repräsentation des Entscheidungsproblems berücksichtigt. Auf Basis von Attraktivitäts-Einschätzungen wird eine der Alternativen in der zweiten Phase, dem *finding a promising alternative*, als Kandidat für eine dominante Alternative gewählt. Diese Alternative wird dann in der dritten Phase, dem *dominance-testing*, dahingehend geprüft, ob sie tatsächlich die andere(n) Alternative(n) dominiert. Ist eine Alternative dominant, wird diese gewählt und der Entscheidungsprozess beendet. Ist dies hingegen nicht der Fall, setzt die vierte Phase, das *dominance-structuring*, ein. Im Hinblick auf die Rekonstruktion von Entscheidungsproblemen und das damit verbundene Streben nach Konsistenz ist diese Phase von besonderer Bedeutung. So werden die Informationen rekonstruiert, um eine dominierende Alternative zu ermitteln. Montgomery unterscheidet hierbei vier Rekonstruktions-Methoden, die einer Alternative zu Dominanz verhelfen können: *De-emphasizing* bedeutet, dass ein Nachteil der Alternative in seiner Relevanz abgemildert wird. Durch *bolstering* kann hingegen die Bedeutung eines erwünschten Attributs erhöht werden. Wird ein Nachteil dadurch kompensiert, dass er als logische Konsequenz eines Vorteils betrachtet wird, spricht Montgomery von *cancellation*. *Collapsing* bezeichnet den Zusammenschluss mehrerer Attribute, wodurch die Ausprägung eines Attributs kompensiert werden kann. Eine Differenzierung der einzelnen Methoden nimmt Montgomery in einer Arbeit aus dem Jahr 1989 vor: Während er *cancellation* und *collapsing* für rationale Methoden hält, sieht er *de-emphasizing* und *bolstering* als irrationale Methoden an. Neben diesen kognitiven Prozessen können auch Handlungen, mit denen in die externale Umwelt eingegriffen wird, für die Konstruktion einer Dominanzstruktur herangezogen werden (Montgomery und Willén, 1999). Einen Überblick über Forschungsergebnisse, die die Annahme der beschriebenen Prozesse und Rekonstruktions-Methoden

stützen, geben Montgomery und Willén (1999). In seiner ersten Arbeit zu dieser Theorie legt Montgomery (1983) vor allem dar, welche Stellung non-kompensatorische und kompensatorische Entscheidungsregeln im Rahmen des SDS-Modells einnehmen. Da das Erreichen von Dominanz für ihn den Entscheidungsprozess charakterisiert, ist die Dominanzregel die zentrale Regel, während die anderen Entscheidungsregeln dazu dienen, Dominanz zu erreichen. Die Nachteile non-kompensatorischer Regeln, z.B. ihre begrenzte Anwendbarkeit oder die mit ihrem Einsatz verbundene Vernachlässigung von wichtigen Informationen, und kompensatorischer Regeln, z.B. die Schwierigkeit, eine Vielzahl an Informationen zu überblicken, oder das sich aus den trade-offs ergebende Bewusstsein, eine andere gute Alternative aufgegeben zu haben, können laut Montgomery gelindert werden, indem die Regeln als Mittel zum Zweck gesehen werden. In einem solchen Ansatz sieht er etwa die Möglichkeit, die Probleme non-kompensatorischer Probleme durch den kombinierten Einsatz non-kompensatorischer und kompensatorischer Regeln und die angesprochenen Schwächen kompensatorischer Regeln durch die Vorstrukturierung und die Taktiken des *dominance structuring* zu beheben. In späteren Arbeiten hebt Montgomery (1989; Montgomery und Willén, 1999) insbesondere den Wert einer Dominanzstruktur für das spätere Handeln hervor. Er begreift den Entscheidungsprozess als eine Vorbereitung für späteres Handeln, indem eine Handlungsintention so stabil ist, dass sie gegen konkurrierende Tendenzen abgeschirmt werden kann (Montgomery, 1989). Letztlich wird dadurch eine Konsistenz des Verhaltens gewährleistet (Montgomery und Willén, 1999). Einen wichtigen Aspekt der SDS-Theorie stellt die Interpretation von Dominanz dar: Schon 1983 weist Montgomery darauf hin, dass die Konstruktion einer Dominanzstruktur nicht zwingend bedeutet, dass eine Alternative im strengen Sinn, das heißt gemäß der obigen Definition, dominant ist. Das Erreichen von Dominanz stellt ein Ziel dar, das zwar angestrebt wird, aber mehr oder weniger erfolgreich erreicht werden kann (Montgomery, 1989). Montgomery und Willén (1999) schlagen daher vor, in Bezug auf die Dominanz in der SDS-Theorie von subjektiver Dominanz zu sprechen. Dadurch wird zum einen berücksichtigt, dass eine Alternative, die sich nur geringfügig von anderen Alternativen abhebt, zwar streng genommen dominant ist, subjektiv aber nicht so wahrgenommen wird. Zum anderen erlaubt diese Konzeption, dass eine Alternative, die einen entscheidenden Vorteil hat, subjektiv als dominant wahrgenommen wird, auch wenn sie die Kriterien einer strengen Dominanz nicht erfüllt. Wie der Entscheider merkt, dass eine Dominanzstruktur erzielt ist, beschreibt Montgomery (1983) folgendermaßen: “I believe that a decision is often preceded by a ‘click’

experience, a feeling of confidence – now I know what to do" (S. 343). Laut Montgomery wird dieses Gefühl dann wahrgenommen, wenn die Gründe für die entsprechende Entscheidung überzeugend sind. Das Finden einer Dominanzstruktur erachtet Montgomery (1989) als sehr schnellen und (mehr oder weniger) automatischen Prozess.

Während laut SDS-Theorie das Ziel des Entscheidungsprozesses darin besteht, eine dominante Alternative zu finden, geht es bei der *Differentiation und Consilidation* (Diff Con)-Theorie von Svenson (1992; 1996) darum, dass die gewählte Alternative hinreichend überlegen sein sollte. Damit dieses Ziel erreicht werden kann, postuliert Svenson prädezisionale Differenzierungsprozesse (*differentiation*) und postdezisionale Konsolidierungsprozesse (*consolidation*). Die Art der Prozesse ist in beiden Fällen gleich: Erstens kann auf Basis einer schnellen holistischen Bewertung differenziert werden. Zweitens kann Differenzierung auf der Prozessebene stattfinden, und zwar durch die Wahl verschiedener Entscheidungsregeln, die alle für eine Alternative sprechen und diese damit stützen, und durch die Modifizierung des Entscheidungskriteriums. Drittens ist eine Differenzierung auf der Strukturebene möglich, wobei die Attribute hinsichtlich ihrer Attraktivität und ihrer Wichtigkeit umgewertet, Fakten restrukturiert – z.B. durch verzerrte Erinnerung – und Problemrepräsentationen etwa durch das Hinzufügen neuer Alternativen modifiziert werden können. Durch diese Differenzierung wird die Anwendung bestimmter Entscheidungsregeln möglich gemacht. Ziel aller drei genannten Prozesse ist es, die gewählte Alternative gegen mögliche zukünftige Bedrohungen, etwa Bedauern (*regret*) aufgrund des Verlusts positiver Aspekte der nicht gewählten Alternative, abzuschirmen. Der Entscheidungsprozess lässt sich nach Svenson folgendermaßen umreißen: Nach einer eventuell notwendigen Eingrenzung des Alternativen-Pools durch die Eliminierung von Alternativen wird eine vorläufige Entscheidung getroffen. Diese vielversprechende Alternative (*promising alternative*) wird noch vor der endgültigen Entscheidung einem Differenzierungsprozess unterzogen. Gelingt die Differenzierung, wird diese Alternative gewählt und es können postdezisionale Konsolidierungsprozesse einsetzen. Konnte keine ausreichende Differenzierung erreicht werden, wird ein neuer Kandidat gesucht und versucht zu differenzieren. Der erforderliche Grad an Differenzierung ist laut Svenson (1992) abhängig vom Entscheidungsproblem und personalen Präferenzen. Er nimmt an, dass unter anderem die Wichtigkeit einer Entscheidung, ihre Irreversibilität und die Verantwortlichkeit für diese Entscheidung zu einer höheren Differenzierung führen. Dass die Differenzierung und Konsolidierung einer Alternative dem übergeordneten Ziel der Konsistenz dienen, wird in Svensons (1992)

Charakterisierung des Entscheidungsprozesses deutlich: “In this process differentiation processes of different kinds operate in order to harmonize the decision maker with the decision problem, his or her own decision and its consequences” (S. 164).

Zusammenfassend betrachtet eint diese Theorien die Annahme, dass im Entscheidungsprozess die Rekonstruktion von Informationen zentral ist. Dieser Rekonstruktionsprozess kann in verschiedener Weise ausgestaltet sein: Während im Story Model das Weglassen und Hinzufügen von Informationen nachgewiesen werden konnten, sind inhaltliche Umwertungen ein Bestandteil von SDS und Diff Con. Die Konstruktion einer kohärenten Geschichte, einer dominanten Alternative oder einer überlegenen Alternative werden aber stets durch die Erhöhung der Konsistenz erzielt. Diese Ansätze stellen aufgrund ihrer konstruktiven Sichtweise einen Mehrwert für die Entscheidungsforschung dar und liefern zudem wichtige Beiträge für die Konsistenzforschung. Allerdings mangelt es den Ansätzen an einer Konkretisierung bzw. Formalisierung des Entscheidungsprozesses, die auch die Dynamik desselben abbildet. Eine Lösung für dieses Problem stellen PCS-Modelle dar. Interessanterweise deutet sich diese Entwicklung bereits bei Montgomery (1989) an. So diskutiert Montgomery die mentale Bereitschaft für die Konstruktion einer Dominanzstruktur und führt als entsprechende Vorgänge u.a. die Repräsentation des Wissens in semantischen Netzwerken an. Er nimmt an, dass sich die Attribute der Alternativen als semantisches Netzwerk darstellen lassen und dass sich die Aktivation in diesem Netzwerk ausbreitet. Der Aufbau einer Dominanzstruktur erfolgt laut Montgomery dadurch, dass breitere Kategorien (z.B. übergeordnete Werte anstelle einzelner Attribute) betrachtet werden, da auf diesem Wege quantitative Unterschiede auf der Attribut-Ebene unbeachtet bleiben und so Verletzungen der Dominanz nicht relevant sind. Eine Alternative kann auch dadurch dominant werden, dass die Knoten, die die Dominanz verletzen, deaktiviert werden, indem mit ihnen verbundene Knoten, die zudem einen positiven Aspekt der Alternative darstellen, aktiviert werden. Montgomery sieht hier die Methode der *cancellation* umgesetzt. Die prinzipielle Annahme, dass ein Entscheidungsproblem als Netzwerk dargestellt und eine dominante Struktur über die Aktivationsausbreitung erreicht werden kann, deckt sich mit den PCS-Modellen. Als weiteren Beleg für die mentale Bereitschaft zur Konstruktion von Dominanz führt Montgomery (1989) den Einfluss von kausalen Schemata an, die wiederum in einer „schwarz-weiß“-Art organisiert sind. Hier findet sich der Attraktor-Gedanke aus PCS-Netzwerken wieder, in denen zwei Attraktoren ein „entweder-oder“ symbolisieren. Folglich überrascht es nicht, dass Montgomery und Willén (1999) die Dominanzstruktur als

gute Gestalt interpretieren. Allerdings mündet Montgomerys (1989) Darstellung nicht in einem spezifischen Modell, zumal die Konzeption als Netzwerk nicht zentral ist, sondern nur einen Teil des Prozesses zu erklären hilft. Eine präzisere und formale Beschreibung des Entscheidungsprozesses erlauben PCS-Netzwerke, deren Anwendung auf den Entscheidungsbereich im Folgenden vorgestellt wird.

DER PARALLEL CONSTRAINT SATISFACTION-ANSATZ IN DER ENTSCHEIDUNGSFORSCHUNG

"Everyday social reasoning and decision making often rely on multitudes of judgments and inferences. (…) Despite the ubiquity and importance of these kinds of tasks, the reasoning process underlying them are not well understood" (Simon, Snow et al., 2004, S. 814).

Die Unzufriedenheit mit bisherigen Entscheidungstheorien veranlasste Forschergruppen wie diejenigen um Thagard und um Simon dazu, alternative Ansätze aufzustellen. Ein wesentlicher Kritikpunkt betraf den unzureichenden Umgang mit Komplexität und Ambiguität, der eine Folge der fehlenden Berücksichtigung von Prozessen dynamischen Zusammenwirkens war (z.B. Holyoak und Simon, 1999; Simon, Snow et al., 2004; Thagard und Millgram, 1995). Diesem Problem begegneten sie mit der Anwendung von PCS-Ansätzen auf die Entscheidungsforschung.[7]

In ihrer *Theory of Deliberative Coherence* (TDC) begreifen Thagard und Millgram (1995) die Entscheidungsfindung als Inferenz des besten Plans, die aus einer holistischen Einschätzung verschiedener Handlungen und Ziele, die sich stützen oder auch widersprechen können, resultiert. Die Kohärenz dieses Sets an Handlungen und Zielen, das gemeinsam den Plan ergibt, wird – ähnlich wie in den weiteren Theorien Thagards (siehe Kapitel 2.2.1) – nach bestimmten Prinzipien beurteilt. So basiert die Einschätzung dieser sogenannten *deliberative coherence* auf insgesamt sechs Prinzipien: (1) *Symmetry* besagt, dass (In)Kohärenz von Handlung A mit B auch (In)Kohärenz von B mit A impliziert. (2) Gemäß der *facilitation* sind Handlungen, die gemeinsam zur Erreichung eines Ziels beitragen, untereinander und mit dem Ziel kohärent. (3) Das Prinzip *incompatibility* drückt im Wesentlichen aus, dass wenn zwei Handlungen nicht gemeinsam ausgeführt und zwei Ziele nicht gemeinsam erreicht werden können, diese inkohärent sind. (4) Mit der *goal priority* wird berücksichtigt, dass etwa intrinsischen Zielen eine besondere Bedeutung zugesprochen

[7] Die Anwendung von PCS-Prozessen zur Erklärung der Dissonanzreduktion haben Shultz und Lepper (1996) vorgenommen. Da es sich bei der Dissonanzreduktion aber um ein postdezisionales Phänomen handelt, wurde dieses Modell bereits in Kapitel 2.2.3 vorgestellt.

wird. (5) Nach dem Prinzip des *judgment* ist auch die Akzeptabilität von Fakten zu berücksichtigen. (6) Schließlich besagt *decision*, dass die Entscheidung auf Basis der Gesamtkohärenz getroffen wird. Das Prinzip der Symmetrie gewährleistet etwa, dass nicht nur beispielsweise aktivierte Ziele kohärente Handlungen aktivieren, sondern dass auch Handlungen ihrerseits zur Aktivation von Zielen beitragen können. Mithilfe des Programms DECO (*DEliberative COherence*), das wie ECHO PCS-Prozesse nutzt, wird die Kohärenz des Zusammenwirkens von Handlungen und Zielen evaluiert. So werden Handlungen und Ziele als Knoten in einem konnektionistischen Netzwerk lokal repräsentiert. Zwischen den einzelnen Knoten bestehen Verbindungen, die im Falle einer Kohärenzbeziehung exzitatorisch und im Falle einer Inkohärenzverbindung inhibitorisch sind. Entsprechend der Stärke der Relation sind sie zudem gewichtet. Die Priorität gewisser Ziele wird durch eine direkte Verbindung zu einem *special node* implementiert. Das Gewicht dieser Verbindung hängt von dem Grad ihrer Wichtigkeit ab. Basierend auf den finalen Aktivationen, die die Attraktivität der Handlungen und Ziele repräsentieren, kann dann der beste Plan inferiert werden. Eine wichtige Eigenschaft dieses Modells ist, dass die Evaluation der Handlungen und Ziele holistisch und damit in wechselseitiger Abhängigkeit erfolgt. Selbst die Attraktivitätseinschätzungen priorisierter Ziele können aufgrund dieses dynamischen Zusammenwirkens revidiert werden.

Dan Simon hat sich in Zusammenarbeit mit Kollegen, insbesondere mit Keith Holyoak und Stephen Read, intensiv der Frage gewidmet, wie PCS-Prozesse bei der Entscheidungsfindung wirken (Holyoak und Simon, 1999; Simon, 2004; Simon und Holyoak, 2002; Simon, Krawczyk, Bleicher & Holyoak, 2008; Simon, Krawczyk & Holyoak, 2004; Simon et al., 2001; Simon, Snow et al., 2004). Sie argumentieren, dass tägliche Entscheidungsprobleme häufig durch komplexe, sich widersprechende und ambigue Informationen gekennzeichnet sind, die durch unidirektionales Schlussfolgern nicht gelöst werden können (Holyoak und Simon, 1999; Simon, Snow et al., 2004). Stattdessen halten sie die Annahme bidirektionaler Prozesse für unerlässlich, die in PCS-Netzwerken umgesetzt sind. Simon und Kollegen stellen keine spezifische Theorie vor, sondern betrachten den PCS-Ansatz als allgemeines Rahmenmodell, das unter anderem zur Erklärung von Entscheidungsprozessen herangezogen werden kann (z.B. Holyoak und Simon, 1999; Simon und Holyoak, 2002; Simon, Krawczyk et al., 2004). Insbesondere für rechtliche Entscheidungen (Holyoak und Simon, 1999; Simon und Holyoak, 2002; Simon et al., 2001; Simon, Snow et al., 2004), aber auch für Präferenzentscheidungen (Simon et al., 2008; Simon, Krawczyk et al., 2004)

konnten sie zeigen, dass ein Entscheidungsprozess als PCS-Prozess charakterisiert werden kann: Ein Entscheidungsproblem wird in einem konnektionistischen Netzwerk dargestellt, in dem wechselseitige Einflüsse zwischen den Informationen und den Entscheidungsalternativen bestehen, so dass sie als *constraints* wirken. Die mentale Repräsentation des Entscheidungsproblems wird dann solange rekonstruiert, bis eine kohärente bzw. konsistente Lösung erzielt ist. Folglich bezeichnen Simon und Kollegen diesen Prozess als *coherence-based reasoning*. Zu ergänzen ist der Hinweis von Simon und Holyoak (2002), dass in sehr komplexen Entscheidungen der kohärenzgeleitete Prozess zu einer anderen Struktur mit Clustern von lokal kohärenten Elementen führen kann. Dan Simon (2004) nennt acht Hauptmerkmale von kohärenzbasiertem Schlussfolgern, die durch empirische Befunde (siehe Kapitel 2.4) gestützt werden: (1) Grundlegend ist das Auftreten von sogenannten *coherence shifts*: Mehrdeutige und unvereinbare Informationen werden systematisch umgewertet, so dass eine kohärente Repräsentation des Entscheidungsproblems konstruiert wird. Hierbei werden die Informationen, die für die spätere Entscheidung sprechen, aufgewertet, während widersprechende Informationen abgewertet werden. Je stärker diese Umwertungen sind, desto höher ist die Konfidenz in die Entscheidung. (2) Personen sind sich des Auftretens dieser *coherence shifts* nicht bewusst. Sie werden als natürliche Folge der automatischen Informationsverarbeitung betrachtet. (3) *Coherence shifts* sind funktional, da sie das Treffen einer konfidenten Entscheidung ermöglichen. Sie sind keine Konsequenz postdezisionaler Dissonanzreduktion. (4) Die Umwertungen zeichnen sich sowohl durch Tiefe als auch durch Breite aus. So werden – wenn auch in geringerem Maße – grundlegende Überzeugungen beeinflusst und es sind Transfereffekte auf ein weiteres Entscheidungsproblem möglich. (5) Es bestehen indirekte Einflüsse: Die Manipulation einer Variablen beeinflusst nicht nur das Urteil, sondern auch die Einschätzung der übrigen Variablen. Dieser Effekt tritt auch auf, wenn eine bereits bestehende kohärente Lösung durch die Einführung einer neuen Information revidiert wird. (6) Der kohärenzgeleitete Entscheidungsprozess wird auch durch Motivation und Einstellungen beeinflusst. (7) Die kohärente mentale Repräsentation wird für das aktuelle Entscheidungsproblem und damit temporär konstruiert, so dass auch *coherence shifts* prinzipiell vergänglich sind. Spätere Entscheidungen werden folglich nicht beeinflusst. (8) *Coherence shifts* können reduziert werden, indem Personen instruiert werden, auch die gegenteilige Entscheidungsalternative zu berücksichtigen. Der Mehrwert, den Simon und Kollegen in der Erklärung von Entscheidungsprozessen mittels PCS sehen, wird insbesondere in drei Punkten deutlich: Erstens kann die Kenntnis

des Auftretens von *coherence shifts* in praktischen Settings hilfreich sein (Holyoak und Simon, 1999; Simon, 2004). Zweitens schätzen sie die Möglichkeit, mittels PCS-Modellen gestaltpsychologische Prinzipien zu implementieren und Konsistenztheorien, insbesondere strukturbezogene wie diejenigen von Abelson und Rosenberg (1958) sowie von McGuire (1960a; McGuire, 1960b; McGuire, 1968), wiederaufleben zu lassen (Simon und Holyoak, 2002; Simon, Snow et al., 2004). Drittens sind sie der Überzeugung, dass viele höhere kognitive Prozesse als PCS-Prozesse zu verstehen sind, so dass etwa Holyoak und Simon (1999) folgende Charakterisierung vornehmen: "(…) human understanding appears (…) like the solution to a complex puzzle in which the individual pieces must be reorganized and transformed to form a coherent whole" (S. 24). Dieses Zitat verdeutlicht exemplarisch auch, welch wichtige Rolle Simon und Kollegen den *coherence* shifts und damit auch dem Streben nach Konsistenz in der Verarbeitung von Informationen beimessen.

Die vorgestellten Arbeiten haben erfolgreich dazu beigetragen, das Wirken von PCS-Prozessen bei der Entscheidungsfindung aufzuzeigen. Glöckner und Betsch (2008b) haben diese Entwicklung mit der Vorstellung ihres konnektionistischen Rahmenmodells für Entscheidungen weiter vorangetrieben. Da dieses Modell die Grundlage für die vorzustellenden empirischen Forschungsarbeiten darstellt, soll es im folgenden Kapitel näher vorgestellt werden.

2.3.2 Das parallel constraint satisfaction-Modell von Glöckner und Betsch (2008)

Glöckner und Betsch (2008b) erachten den Prozess der automatischen Konsistenzmaximierung als "the core information integration process in decision making" (S. 217), weshalb er in ihrem integrativen Modell der automatischen und deliberaten Entscheidungsfindung eine zentrale Stellung einnimmt. Bereits in früheren Arbeiten haben Betsch und Glöckner Entscheidungsprobleme als PCS-Netzwerke dargestellt. In der *Preference Theory* von Betsch (2005a) hängt die Entscheidung selbst von affektiven Reaktionen ab, wobei die Veränderung der Aktivation der Knoten durch automatische PCS-Prozesse entscheidend ist. Die Bedeutung von Affekt betrachten Betsch, Haberstroh, Molter und Glöckner (2004) als deutliche Abgrenzung zum Ansatz von Thagard und Millgram (1995). Betsch (2005a) betont die Bedeutung automatischer Prozesse, erläutert aber auch, wann deliberate Prozesse einsetzen. Glöckner (2006) hat eine intuitive Entscheidungsregel, die *Consistency Maximizing Strategy*, für probabilistische Entscheidungsregeln formuliert und empirisch

erfolgreich getestet. Gemeinsam führen Glöckner und Betsch (2008b) nun ein integratives Rahmenmodell der Entscheidungsfindung aus. Die Autoren nehmen eine grundlegende Unterscheidung zwischen Strategien für die Informationssuche und Strategien für die Entscheidung vor. Während die erste Kategorie verschiedene Strategien umfassen kann, postulieren sie, dass es lediglich eine einzige Entscheidungsregel gibt, die PCS-Regel. Eine konstruktive Sichtweise des Entscheidungsprozesses setzt demnach nicht die Existenz eines Repertoires an Entscheidungsstrategien voraus (vgl. Payne, Bettman, Coupey et al., 1992), sondern lässt sich auch direkt als Bestandteil einer universalen Entscheidungsregel implementieren. Damit ist der Rekonstruktionsprozess in diesem Modell zentral und von hoher Bedeutung. Demnach wird ein Entscheidungsproblem in eine mentale Repräsentation übersetzt, die sich als ein konnektionistisches Netzwerk, das primäre Netzwerk, darstellen lässt. Ein derartiges Entscheidungsproblem kann sowohl eine Präferenz- als auch eine probabilistische Inferenzentscheidung darstellen. Für letzteren Entscheidungstyp stellen sie ein spezifiziertes Netzwerk vor, das in den folgenden Ausführungen im Vordergrund steht. Eine probabilistische Inferenzentscheidung lässt sich nach Bröder (2000) folgendermaßen beschreiben: Es ist eine bewusste Entscheidung zwischen zwei oder mehr *Optionen* zu treffen. Diese Entscheidungsaufgabe hat eine richtige Lösung, die jedoch nicht bekannt ist. Daher wird auf bestimmte Merkmale (*Cues*) zurückgegriffen, die mit einer bestimmten Wahrscheinlichkeit (*Validität*) die richtige Option vorhersagen. Auf diesem Wege kann die richtige Lösung inferiert werden. Die Optionen und die Cues stellen die Knoten in dem von Glöckner und Betsch (2008b) konzipierten Netzwerk dar (siehe Abbildung 1).

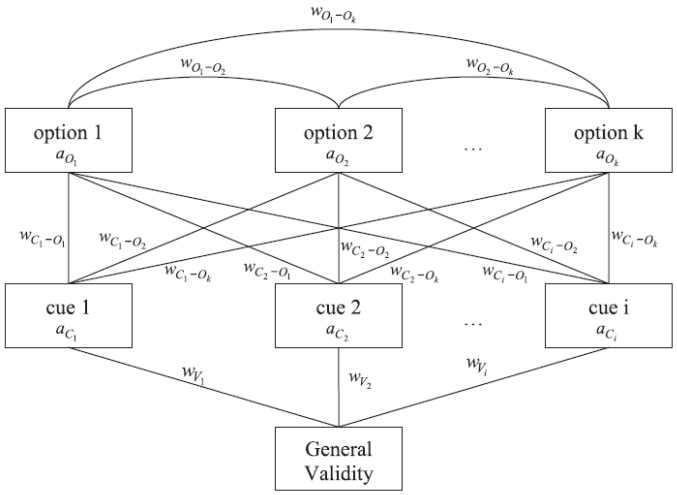


Abbildung 1: Das PCS-Modell für probabilistische Inferenzentscheidungen (aus Glöckner und Betsch, 2008b, S. 218). Erläuterung im Text.

Folglich liegt dem Modell eine lokale bzw. symbolische Präsentation (Read et al., 1997) zugrunde. Wie in den im vorherigen Kapitel beschriebenen PCS-Modellen handelt es sich bei den *constraints* um *soft constraints*. Die Optionen und Knoten sind durch exzitatorische – wenn ein Cue eine positive Aussage über eine Option macht – und inhibitorische – wenn ein Cue eine negative Aussage über eine Option macht – Verbindungen miteinander verknüpft. Das Gewicht der Verbindungen hängt von der Höhe der Vorhersagekraft der Cues ab. Zwischen den Optionen besteht zudem eine starke inhibitorische Verbindung, da nur eine Option gewählt werden kann. All diese Verbindungen sind bidirektional. Ein allgemeiner Validitätsknoten aktiviert das Netzwerk, indem er mit den Cues exzitatorisch verbunden ist. Das Gewicht dieser Verbindungen repräsentiert die initiale, das heißt die bei der Konfrontation mit dem Entscheidungsproblem vorliegende Validität der Cues. Da diese Verbindungen unidirektional sind, wird der allgemeine Validitätsknoten nicht aktualisiert. Im Gegensatz zu den anderen Knoten, die eine initiale Aktivation von 0 haben, besitzt dieser Knoten eine konstante Aktivation von 1.

Die Funktionsweise des Netzwerks gleicht derjenigen der bereits beschriebenen PCS-Netzwerke: Bei Aktivierung des Netzwerks wird Aktivation im Netzwerk ausgebreitet und die Aktivation jedes Knoten wird iterativ verändert, bis eine stabile und damit konsistente Lösung erreicht ist. Für die Berechnung der Aktivation eines Knotens greifen Glöckner und Betsch (2008b, S. 218) auf den häufig verwendeten Algorithmus von McClelland und Rumelhart (McClelland und Rumelhart, 1981; Rumelhart, Hinton et al., 1986) zurück:

$$a_i(t+1) = a_i(t) * (1 - decay) + \begin{cases} if \quad input_i(t) < 0 & input_i(t)(a_i(t) - floor) \\ if \quad input_i(t) \geq 0 & input_i(t)(ceiling - a_i(t)) \end{cases}$$

Demnach wird die Aktivation eines Knotens *i* zum Zeitpunkt *t+1* berechnet, indem die aktuelle (zum Zeitpunkt *t* vorliegende) Aktivation des Knotens *i* mit einem *decay*-Faktor multipliziert wird und dieses Produkt je nach dem, ob der Input für den Knoten positiv oder negativ ist, durch einen Faktor vergrößert oder verkleinert wird. Es wird verkleinert, wenn der Input kleiner als 0 ist, und vergrößert, wenn der Input größer als 0 ist. Dadurch nähert sich die Aktivation eines Knotens asymptotisch entweder der minimalen Aktivation (*floor*) oder der maximalen Aktivation (*ceiling*) an. Der Input selbst wird wie in anderen PCS-Netzwerken (z.B. Rumelhart, Hinton et al., 1986; Thagard, 1989) wie folgt berechnet (Glöckner und Betsch, 2008b, S. 218):

$$input_i(t) = \sum_{j=1 \rightarrow n} w_{ij} a_j(t)$$

Die Aktivation jedes Knoten *j*, mit dem der Knoten *i* verbunden ist, wird mit dem Gewicht der Verbindung zwischen den beiden Knoten, w_{ij}, multipliziert. Diese Produkte werden für alle Knoten aufsummiert.

Um eine stabile Lösung zu erreichen, werden die Aktivationen der Knoten fortlaufend modifiziert. Aufgrund der bidirektionalen Verbindungen kann ein initialer Vorsprung und damit eine vergleichsweise höhere Aktivation einer Option auf die Cues rückwirken, indem die Aktivationen der Knoten, die diese Option stützen, erhöht werden, während die Aktivationen der widersprechenden Knoten erniedrigt werden. Wenn eine stabile Lösung vorliegt, endet das *Updating* der Aktivationen. Die Option, die dominiert, wird gewählt. Eine Veränderung der Verbindungsgewichte findet nicht statt.

Die Parameter, die das Netzwerk näher spezifizieren und in Simulationsstudien eingesetzt wurden, werden ebenfalls berichtet: Den Zerfallsparameter legen sie zumeist auf .05 fest. Das Gewicht zwischen Knoten, das heißt zwischen Optionen, $w_{O1\text{-}O2}$, bzw. zwischen Cues und Optionen, $w_{C\text{-}O}$, liegt zwischen 1.0 (exzitatorische Verbindung) und -1.0 (inhibitorische Verbindung). Die minimal und maximal mögliche Aktivation der Knoten beschränken sie auf -1 bzw. +1.

Glöckner und Betsch nehmen an, dass dieser Prozess der Konsistenzmaximierung automatisch verläuft und der Kern jeder Entscheidung ist. Erst wenn der automatische Prozess zu keiner ausreichend konsistenten Lösung führt, das heißt wenn eine bestimmte Konsistenzschwelle θ nicht überschritten wird, vermuten sie den Aufbau eines sekundären Netzwerks, in dem sogenannte *deliberate constructions* zur Verfügung stehen. Hierunter verstehen die Autoren Strategien der Informationssuche, -produktion und -veränderung. In diesem Netzwerk kann zwar tatsächlich zwischen verschiedenen Strategien entschieden werden, die Entscheidung selbst basiert aber auch hier auf PCS-Prozessen. Zu beachten ist, dass diese *deliberate constructions* lediglich der Unterstützung des primären Netzwerks dienen, das heißt, dass zwar das Netzwerk deliberat modifiziert wird, die letztliche Entscheidung aber im primären Netzwerk automatisch getroffen wird. Glöckner und Betsch geben des Weiteren zu bedenken, dass bei den meisten Entscheidungen deliberate Prozesse gar nicht notwendig sind, was die Bedeutung und auch die Möglichkeiten automatischer Prozesse unterstreicht. Bezüglich der Konsistenzschwelleθ weisen d ie Autoren darauf hin, dass diese je nach Person, Aufgabe und Kontext variieren kann. Die Struktur und die Funktionsweise des Netzwerks verdeutlichen erstens, dass jede Entscheidung auf PCS-Prozessen basiert, und zweitens, dass mehrere Informationen automatisch und kompensato-

risch verarbeitet werden können. Eine Einordnung von PCS-Modellen in den größeren Forschungskontext multiattributiver Entscheidungen nehmen Newell und Bröder (2008) vor. Sie unterscheiden hierbei vier kognitive Modelle, und zwar regelbasierte Modelle, Evidenzakkumulationsmodelle, intuitive Netzwerkmodelle und exemplarbasierte Modelle. Die Autoren stellen die drei letztgenannten Ansätze jeweils dem *adaptive toolbox*-Ansatz (Gigerenzer et al., 1999) gegenüber. Eine für die vorliegende Fragestellung wichtige Abgrenzung besteht zwischen dem *adaptive toolbox*-Ansatz und PCS-Modellen. Auch Glöckner und Betsch (2008b) weisen auf die deutlichen Unterschiede zwischen diesem Ansatz und ihrem Modell hin. Nach dem *adaptive toolbox*-Ansatz von Gigerenzer et al. (1999) verfügen Menschen über eine begrenzte kognitive Informationsverarbeitungskapazität. Durch die Anwendung simpler non-kompensatorischer Entscheidungsregeln (Heuristiken) seien sie aber in der Lage, situationsspezifisch und damit adaptiv Regeln anzuwenden. Diese bestehen aus Prinzipien für die Informationssuche, für die Beendigung der Informationssuche und für die Entscheidung. Diese Regeln werden als *fast* und *frugal* bezeichnet, da sie einfach anzuwenden sind und lediglich einen Teil an Informationen berücksichtigen. Ein prominentes Beispiel ist die *Take The Best* (TTB)-Heuristik, bei der die Entscheidung auf Basis des vorhersagekräftigsten Cues getroffen wird (sofern dieser zwischen den Optionen differenziert). Die Unterschiede zum PCS-Modell sind evident und werden von Glöckner (2006) ausführlich diskutiert: So wird im PCS-Modell eine automatische und kompensatorische Integration einer Vielzahl von Informationen angenommen, in der die letztliche Entscheidung auf der Gesamtinterpretation beruht. In zahlreichen Experimenten (z.B. Glöckner, 2006; Glöckner, 2008a; Glöckner und Betsch, 2008c) konnten Glöckner und Betsch zeigen, dass die auf PCS-Prozessen basierende automatische Informationsverarbeitung Personen tatsächlich dazu befähigt, eine Vielzahl an Informationen automatisch und kompensatorisch zu verarbeiten, und damit ihr Modell empirisch stützen.[8] Der Stellenwert der Konsistenz wird besonders in zwei Experimentalergebnissen deutlich: (1) Eine konsistente (im Sinne von eindeutige) Entscheidungsaufgabe geht einher mit hoher Konfidenz und niedriger Reaktionszeit (Glöckner, 2006; Glöckner und Betsch, 2008c). (2) Es konnten systematische Umwertungen von Cue-Validitäten nachgewiesen werden (Glöckner, 2006; Glöckner, Betsch & Schindler, im Druck). An dieser Stelle ist es sinnvoll, das PCS-Modell von zwei anderen Ansätzen abzugrenzen, die zwar Gemeinsamkeiten mit dem PCS-Modell haben, aber sich dennoch in grundlegenden Aspekten von diesem

[8] In weiteren Experimenten (Glöckner und Betsch, 2008a; Glöckner und Herbold, im Druck) konnte das PCS-Modell auf riskante Entscheidungen angewandt werden.

unterscheiden: Die *Decision Field Theory* (DFT) von Busemeyer und Townsend (1993) und das *Reconstruction After Feedback with Take the best* (RAFT)-Modell von Hoffrage, Hertwig und Gigerenzer (2000). Insbesondere zwei Gemeinsamkeiten verdeutlichen die Nähe von DFT und dem PCS-Modell. Erstens ist das Überschreiten einer Konfidenzschwelle ein zentraler Aspekt (Glöckner und Betsch, 2008b) und zweitens wählen beide Modelle einen konnektionistischen Ansatz (Busemeyer und Johnson, 2004). Glöckner und Betsch (2008b) weisen aber selbst darauf hin, dass es zwei wichtige Unterschiede zwischen beiden Ansätzen gibt: (1) In ihrem PCS-Modell erfolgt die Informationsintegration parallel und nicht seriell. (2) Im PCS-Modell bestehen bidirektionale Verbindungen zwischen Cues und Optionen, so dass diese selbst einer Interpretation und Bewertung unterzogen werden. Hinzuzufügen ist, dass Busemeyer und Townsend (1993) keinen automatischen, sondern einen deliberaten Prozess annehmen, und dass hierbei affektive Reaktionen auf Konsequenzen von Handlungen akkumuliert werden. Im PCS-Modell von Glöckner und Betsch (2008b) werden hingegen alle verfügbaren Informationen, unabhängig von ihrer objektiven Relevanz, integriert. Dies schließt gewiss nicht aus, dass einzelne Informationen affektiv konnotiert sind bzw. affektive Reaktionen eine Rolle spielen. Deutlich abzugrenzen ist das vorgestellte PCS-Modell vom RAFT-Modell, das Hoffrage et al. (2000) vorgeschlagen haben. Mit dem Verweis auf dieses RAFT-Modell widersprach Marewski (im Druck) der Interpretation von Glöckner et al. (im Druck), dass das Auftreten von *coherence shifts* als Evidenz gegen den von Gigerenzer et al. (1999) postulierten *adaptive toolbox*-Ansatz zu interpretieren ist, da dieser eine unidirektionale Informationsverarbeitung annehme. Hoffrage et al. (2000) schlagen mit RAFT ein Modell vor, in dem der sogenannte *hindsight bias* als Rekonstruktionsprozess aufgefasst wird, dem ein *Updating* von Wissen vorausgeht. Die Erforschung des *hindsight bias* nahm ihren Anfang bei den Untersuchungen von Fischhoff (1975) und kann als "tendency for people with outcome knowledge to believe falsely that they would have predicted the reported outcome of an event" (Hawkins und Hastie, 1990, S. 311) beschrieben werden. Der *hindsight bias* war Gegenstand zahlreicher Untersuchungen (Überblicke geben Christensen-Szalanski und Willham, 1991; Guilbault, Bryant, Brockway & Posavac, 2004; Hawkins und Hastie, 1990). In einer typischen Untersuchung gibt eine Versuchsperson Urteile ab, z.B. beantwortet sie Almanach-Fragen, und erhält dann Feedback. Zu einem späteren Zeitpunkt wird die Versuchsperson gebeten, sich an ihr initiales Urteil zu erinnern. Gemäß dem *hindsight bias* ist die Erinnerung der Versuchsperson dahingehend verzerrt, dass sie meint, initial die korrekte Antwort genannt zu haben.

Hoffrage et al. (2000) verwendeten in ihren Untersuchungen zum RAFT-Modell ein *memory design*, das heißt, dass die Versuchspersonen zunächst bestimmte Cue-Werte lernten, auf deren Basis später probabilistische Inferenzaufgaben gelöst werden mussten. Die Autoren gehen davon aus, dass diese Entscheidung gemäß der TTB-Heuristik getroffen wird. Auch die Konfidenz wurde erhoben. Zu einem späteren Zeitpunkt wurde den Versuchspersonen entweder Feedback gegeben oder nicht und sie sollten sich dann an ihre initialen Entscheidungen, ihre Konfidenzurteile und die gelernten Cue-Werte erinnern. Die Hypothesen der Autoren konnten bestätigt werden: (1) Feedback beeinflusst die Erinnerung systematisch, da das Wissen entsprechend der Rückmeldung aktualisiert wurde und dadurch *hindsight bias* auftrat. (2) Unter Zugrundelegung der erinnerten Cue-Werte konnte RAFT vorhersagen, ob *hindsight bias*, umgekehrter *hindsight bias* (Entscheidung und Konfidenz entgegen dem Feedback, bedingt durch unsystematische Verzerrungen) oder korrekte Erinnerung auftraten. (3) Der *hindsight bias* konnte reduziert werden, indem die Erinnerung der initialen Cue-Werte unterstützt wurde. Allerdings ist diese Art der Modifikation grundsätzlich verschieden von derjenigen, die Glöckner et al. (im Druck) als Resultat des Wirkens von PCS-Prozessen gefunden haben. Die entscheidenden Unterschiede erwähnen Glöckner und Betsch (im Druck): Erstens traten die Modifikationen vor der Entscheidung auf, zweitens wurde kein Feedback gegeben und drittens wurden die Cue-Werte nicht gelernt, sondern vorgegeben. Ein weiterer Punkt, den Glöckner und Betsch nennen, ist, dass bei ihnen Cue-Validitäten und nicht Cue-Werte modifiziert wurden. Allerdings können Hoffrage et al. (2000) nachweisen, dass mit der fälschlichen Erinnerung von Cue-Werten auch veränderte Cue-Validitäten einhergehen. Da sie die TTB-Heuristik als Entscheidungsregel annehmen, weisen sie allerdings nur eine generelle Aufwertung des validesten Cues nach, nicht eine systematische Reevaluation aller Cues. Folglich erscheint die Schlussfolgerung von Glöckner und Betsch (im Druck) durchaus berechtigt, nach der *coherence shifts* nicht durch den Ansatz der *adaptive toolbox* erklärt werden können.

Vor dem Hintergrund der bisherigen Entscheidungsforschung zeichnet sich das Modell von Glöckner und Betsch insbesondere dadurch aus, dass es ein integratives Rahmenmodell für die Entscheidungsfindung darstellt, das – gestützt durch zahlreiche empirische Befunde – die Stärke der menschlichen Informationsverarbeitung aufzeigt. Vor dem Hintergrund der Konsistenzforschung ist dieses Modell interessant, da es den PCS-Mechanismus als Kernmechanismus eines jeden Entscheidungsprozesses begreift und in ein Rahmenmodell einbettet.

2.4 BEFUNDE ZUM STREBEN NACH KONSISTENZ IN DER ENTSCHEIDUNGSFORSCHUNG

In den dargestellten Entscheidungstheorien und -modellen ist deutlich geworden, dass die Frage nach der Bedeutung von Konsistenz im Entscheidungsprozess eng verknüpft ist mit der Rekonstruktion von Informationen. Empirisch beobachten lässt sich dies als prädezisionale Informationsverzerrung. Die Reevaluation von Alternativen und Attributen erlaubt die Konstruktion einer konsistenten Repräsentation der Entscheidungssituation. In einigen Forschungsarbeiten wird ein direkter Zusammenhang zwischen dem Streben nach Konsistenz und der prädezisionalen Umwertung von Informationen hergestellt. Dies betrifft insbesondere Arbeiten, deren theoretische Basis der PCS-Ansatz darstellt (z.B. Glöckner et al., im Druck; Simon, Snow et al., 2004). Doch auch Forschergruppen, die nicht direkt zu den Vertretern des PCS-Ansatzes zu zählen sind, entdecken das Bedürfnis nach Konsistenz als Begründung für diesen *Bias* (Russo, Carlson, Meloy & Yong, 2008). Unabhängig davon, ob eine verzerrte Interpretation von Informationen mit dem Ziel der Konsistenz erklärt wird, geht sie doch damit einher. Aus diesem Grund sollen im Folgenden Befunde, die sich der Untersuchung der prädezisionalen Reevaluation von Informationen gewidmet haben, referiert werden. Zunächst wird in Kapitel 2.4.1 ein allgemeiner Überblick über dieses Phänomen sowie zwei besonders relevante Untersuchungsparadigmen gegeben. In den Kapiteln 2.4.2 und 2.4.3 werden dann aktuelle Befunde, getrennt für Charakteristika der Aufgabe und Charakteristika der Person, dargestellt. Abschließend werden die Befunde tabellarisch zusammengefasst (siehe Tabelle 1).

2.4.1 Prädezisionale Informationsverzerrung

Leon Festinger und seine Kollegen zeigten bereits früh, dass die Reduktion von Dissonanz bzw. das Erhöhen von Konsistenz eine wesentliche Rolle bei der Entscheidungsfindung einnimmt (z.B. Brehm, 1956; Festinger, 1957; Festinger, 1964). Festinger (1964) geht hierbei davon aus, dass Dissonanz nur auftreten kann, nachdem eine Entscheidung getroffen worden ist und wenn zudem *Commitment* besteht. Das Streben nach Konsistenz ist jedoch weder an diese Phase noch an die Voraussetzung des *Commitments* geknüpft. So konnte in einer Vielzahl von Forschungsarbeiten gezeigt werden, dass Informationsverzerrungen, die zu der Konstruktion einer konsistenten Entscheidungssituation führen, prädezisional und damit auch ohne *Commitment* auftreten (für einen Überblick siehe Brownstein, 2003).

Wichtige Erkenntnisse wurden durch frühe Arbeiten gewonnen, in denen bedeutsame Funktionen systematischer prädezisionaler Informationsumwertungen für den Entscheidungsprozess nachgewiesen werden konnten. Mann, Janis und Chaplin (1969) zeigten z.B., dass Reevaluationen dem Beenden des prädezisionalen Konflikts dienen, wenn keine weiteren Informationen mehr die Entscheidung beeinflussen können. Des Weiteren können sie dem Gewinnen von Sicherheit dienen, wenn wichtige Entscheidungen getroffen werden müssen (O'Neal, 1971). In der Folge wurden prädezisionale Informationsverzerrungen eher als zentraler und natürlicher Bestandteil des Entscheidungsprozesses interpretiert: "Coherence shifts (…) are the natural consequence of the normal mechanisms of cognitive processing" (Simon, 2004, S. 545 f.). Zu dieser Interpretation haben u.a. Befunde beigetragen, nach denen eine a priori vorliegende Präferenz bzw. Hypothese nicht nötig für das Auftreten der systematischen Umwertungen ist (Russo, Medvec & Meloy, 1996; Simon, Krawczyk et al., 2004). Dies erlaubt zudem eine deutliche Abgrenzung zu Phänomenen wie dem *confirmation bias* (Nickerson, 1998). In weiteren Studien konnte belegt werden, dass Informationen selbst dann verzerrt werden, wenn keine Entscheidung getroffen werden musste (z.B. Russo, Meloy & Medvec, 1998; Simon et al., 2001). Damit lässt sich festhalten, dass ein Streben nach Konsistenz besteht, das bereits durch die Konfrontation mit einer Entscheidungssituation aktiviert wird: "the pressure to achieve coherence guides the decision-making process itself" (Holyoak und Simon, 1999, S. 4). Die Vielzahl und Vielfalt an stützenden Befunden für prädezisionale Informationsverzerrung lässt sogar den Schluss zu, dass diese ein (nahezu) allgegenwärtiges Phänomen ist (Polman, 2010; Russo et al., 2008).

Die prädezisionale Verzerrung von Informationen umfasst sowohl die selektive Informationssuche als auch die Reevaluation von Alternativen (Brownstein, 2003). Die reichhaltige Evidenz zu beiden Phänomenen veranlasst Brownstein zu folgender Schlussfolgerung: "Therefore, it is time to replace the debate over whether biased predecision processing occurs with a more constructive analysis over when and why it is most likely to occur" (S. 566). Brownstein selbst nennt – auf Basis seines Reviews – einige Moderatoren, die vor allem Charakteristika der Situation darstellen. Demnach findet eine stärkere prädezisionale Informationsverzerrung statt, wenn (a) eine Entscheidung getroffen werden muss, (b) keine Erwartung besteht, diese Entscheidung zu rechtfertigen, (c) die Alternativen ähnlich attraktiv sind, (d) ein Kompromiss (*trade-off*) zwischen Attributen zu finden ist, (e) mehr als zwei Alternativen zur Wahl stehen, (f) der Entscheidungsprozess vorangeschritten

ist und (g) weniger Zeit bis zur Entscheidung verbleibt, (h) keine weiteren Informationen erwartet werden, (i) die Evaluationen privat vorzunehmen sind, (j) längere Deliberationszeit zur Verfügung steht, (k) Erregung (*arousal*) nicht auf eine bestimmte Ursache attribuiert werden kann, (l) die verfügbaren Informationen annehmbar sind, (m) die Wahl einer Alternative mehr Unterstützung voraussetzt und (n) im Falle einer Gruppenentscheidung diese Gruppe homogen ist. Brownstein nennt als weiteren Moderator die Konsequenz einer Entscheidung, weist aber auf die bis dato uneindeutige Evidenz hin. Nur wenige Moderatoren, die von Brownstein aufgelistet werden, betreffen Charakteristika der Person. So führen eine höhere Handlungsorientierung und eine Disposition zu guter Stimmung zu stärkerer prädezisionaler Informationsverzerrung.

Basierend auf diesen ersten Erkenntnissen fordert Brownstein (2003): "In the coming years, researchers are likely to continue to sharpen our understanding of biased predecision processing as an important part of the decision making process" (S. 566). In der Tat hat sich in den letzten Jahren eine Vielzahl an Forschungsarbeiten mit der prädezisionalen Informationsverzerrung befasst. In den folgenden beiden Kapiteln sollen daher – getrennt für Charakteristika der Aufgabe und Charakteristika der Person – wichtige Befunde vorgestellt werden. Die in dieser Arbeit vorgenommene Fokussierung auf den Entscheidungsprozess und auf die Konsistenz zwischen gleichzeitig vorliegenden Elementen schränkt die Betrachtung auf solche Studien ein, die sich mit der prädezisionalen Reevaluation von Alternativen bzw. Informationen befassen. Folglich werden Studien, die sich vornehmlich auf die postdezisionale Dissonanzreduktion (z.B. Brehm, 1956; Festinger, 1957; Festinger, 1964) bzw. Konsolidierung (z.B. Svenson, 1996; Svenson, Salo & Lindholm, 2009) oder auf die selektive Informationssuche beziehen, nicht berücksichtigt. Die Begriffe prädezisionale Informationsverzerrung und Reevaluation von Alternativen bzw. Informationen werden daher synonym verwendet. Vor allem die Forschergruppen um Russo (Bond, Carlson, Meloy, Russo & Tanner, 2007; Carlson und Russo, 2001; Meloy und Russo, 2004; Russo, Carlson & Meloy, 2006; Russo et al., 2008; Russo et al., 1996; Russo et al., 1998; Russo, Meloy & Wilks, 2000) und Simon (Holyoak und Simon, 1999; Simon, 2004; Simon und Holyoak, 2002; Simon et al., 2008; Simon, Krawczyk et al., 2004; Simon et al., 2001; Simon, Snow et al., 2004) haben sich intensiv mit der Reevaluation von Informationen befasst. Bevor daher konkrete Befunde dargestellt werden, sollen die beiden von Russo und Kollegen sowie Simon und Kollegen verwendeten Forschungsparadigmen in typischer Form dargestellt werden.

DAS FORSCHUNGSPARADIGMA VON RUSSO UND KOLLEGEN

Russo und Kollegen haben die Reevaluation von Informationen insbesondere für Präferenzentscheidungen (Meloy und Russo, 2004; Russo et al., 2006; Russo et al., 2008; Russo et al., 1996; Russo et al., 1998; Russo et al., 2000), aber auch für rechtliche Entscheidungen (Carlson und Russo, 2001) und für die Beurteilung einer einzelnen Option (Bond et al., 2007) untersucht. In einem typischen Experiment werden die Versuchspersonen vor eine Wahl zwischen zwei Alternativen A und B gestellt. Um den Entscheidungsprozess zu beobachten, verwenden sie die Methode der *Stepwise Evolution of Preference* (SEP). Hierbei werden die Attribute sequentiell dargeboten, wobei stets in einem kleinen Absatz beschrieben wird, wie ein Attribut bei beiden Alternativen ausgeprägt ist. Auf Basis im Vorhinein erhobener Ratings können diese Attribute als neutral oder als diagnostisch (eine Alternative favorisierend) eingestuft werden. Des Weiteren werden den Probanden bei dieser Methode nach jeder Darbietung eines Attributs drei Fragen gestellt. Erstens sollen sie auf einer 9-stufigen Skala angeben, wie sehr ihrer Meinung nach das Attribut für die Alternative A oder die Alternative B spricht. Zweitens sollen sie angeben, welche Alternative zu diesem Zeitpunkt präferiert wird. Um zu vermeiden, dass bereits eine Entscheidung getroffen wird, formulieren Russo und Kollegen diese Frage als Pferderennen. Die Versuchspersonen werden gebeten sich vorzustellen, dass die beiden Alternativen Pferde in einem Rennen sind, und dann benennen, welches Pferd gerade führend ist. Drittens sollen die Probanden Auskunft über ihre Konfidenz in diese Entscheidung geben, und zwar in Form eines Wetteinsatzes: So sollen sie einen festen Geldbetrag auf Alternative A und B aufteilen. Meistens können die Versuchspersonen den Entscheidungsprozess selbstständig beenden. Das Experiment endet dann mit der Angabe der finalen Entscheidung und der Konfidenz in dieselbe. Das Ausmaß an Informationsverzerrung wird im Vergleich zu einer Kontrollgruppe ermittelt. Diese Kontrollgruppe trifft keine Entscheidung. Um zudem Reevaluationen, die sich allein aus der Reflexion über die Entscheidungsaufgabe ergeben können (Russo et al., 1998; Simon et al., 2001), zu vermeiden, sind die einzelnen zu evaluierenden Attribute jeweils einem neuen Alternativenpaar zugeordnet. Die Differenz zwischen der Attributs-Bewertung der Kontrollgruppe und der Experimentalgruppe zeigt das Ausmaß der Umwertung an, das Vorzeichen die Richtung. Eine Umwertung ist positiv, wenn sie in Richtung der derzeitig präferierten Alternative erfolgt, und dementsprechend negativ, wenn das Gegenteil der Fall ist. Hieraus ergibt sich, dass eine Umwertung weder für das erste Attribut noch für Attribute, bei denen der Wetteinsatz gleichmäßig auf beide Alternativen

aufgeteilt und damit die Konfidenz in die Präferenz mit 50% beurteilt wurde, berechnet werden kann.

Das Forschungsprogramm von Simon und Kollegen

Simon und Kollegen haben die Reevaluation von Informationen hauptsächlich für rechtliche Entscheidungen (Holyoak und Simon, 1999; Simon, 2004; Simon und Holyoak, 2002; Simon et al., 2001; Simon, Snow et al., 2004), aber auch für Präferenzentscheidungen (Simon et al., 2008; Simon, Krawczyk et al., 2004) nachgewiesen. Im Gegensatz zu Russo und Kollegen untersuchen sie diese aber stets in einem *within subjects*-Design. Ein weiterer wichtiger Unterschied betrifft die Interpretation der Reevaluationen: Während Russo et al. davon ausgehen, dass sich bei der sequentiellen Beurteilung der Informationen eine Präferenz für eine Alternative herauskristallisiert, die durch die Umwertung der folgenden Informationen gestützt wird (z.B. Russo et al., 1998), nehmen Simon und Kollegen das Wirken bidirektionaler Prozesse zwischen Optionen und Informationen an, die zur Konstruktion eines stimmigen bzw. konsistenten Gesamtbildes führen (z.B. Holyoak und Simon, 1999). Die Umwertung der Informationen in Richtung der favorisierten Option bezeichnen sie als *coherence shifts* (siehe Kapitel 2.3.1). Simon und Kollegen haben ihr Paradigma im Laufe der Zeit modifiziert, insbesondere um postdezisionale Dissonanzreduktionsprozesse ausschließen zu können (vgl. Simon, Krawczyk et al., 2004; Simon et al., 2001; Simon, Snow et al., 2004). Ein derartiges typisches Experiment zu rechtlichen Entscheidungen beginnt mit einem Pretest. In diesem sind verschiedene Situationen in Form von Vignetten geschildert, gefolgt von ein bis zwei Aussagen zu dieser Situation. Diese können sich direkt auf den Sachverhalt beziehen oder eher allgemeine Überzeugungen betreffen. Die Aufgabe der Versuchspersonen ist es, auf einer 11-stufigen Skala anzugeben, wie sehr sie einer Aussage zustimmen. Nach einer Ablenkungsaufgabe folgt das Treatment: Den Versuchspersonen wird ein rechtlicher Fall präsentiert, und zwar sowohl die unbestrittene Evidenz als auch die Argumente der Anklage und der Verteidigung. Die Versuchspersonen sollen darüber reflektieren, ob sie den Angeklagten eher für schuldig oder für unschuldig halten. Sie geben dann eine vorläufige Tendenz (*leaning*) sowie ihre Konfidenz (auf einer 5-stufigen Skala) in diese an, treffen aber noch keine Entscheidung. Dies wird in verschiedener Weise sichergestellt, z.B. indem weitere Informationen angekündigt werden (z.B. Holyoak und Simon, 1999) oder indem die Versuchspersonen gar nicht erwarten, eine Entscheidung zu treffen (Simon et al., 2001). Dann wird den Probanden

ein Posttest vorgelegt, der ähnlich wie der Pretest aufgebaut ist, nur dass die Vignetten diesmal inhaltlich auf den dargebotenen Fall bezogen sind und die Aussagen die Ansichten von Anklage bzw. Verteidigung wiedergeben. Zum Teil wurde der Posttest auch vor der Mitteilung des *leanings* bearbeitet, was aber nicht zu anderen Ergebnissen führte (vgl. Simon et al., 2001). Eine finale Entscheidung, die Angabe der Konfidenz in diese sowie eine erneute Bearbeitung des Posttests bilden den Abschluss eines typischen Experimentes. Der Posttest wird ein zweites Mal ausgehändigt, um, wie bei Simon et al. (2001), die Reevaluation von Informationen nach dem Treffen einer Entscheidung zu verfolgen, oder um, wie bei Simon, Snow et al. (2004), den Einfluss einer weiteren Information auf das Urteil und die Bewertung der Informationen zu untersuchen. Der zentrale Auswertungsschritt besteht darin, die Einschätzungen der Aussagen sowie deren Interkorrelation im Pretest mit den Einschätzungen und Interkorrelationen im Posttest miteinander zu vergleichen. Um die Interkorrelationen interpretieren zu können, werden die Informationen vorher so skaliert, dass sie alle die Übereinstimmung mit einer der beiden Positionen wiedergeben. *Coherence shifts* werden daran deutlich, dass die Informationen, die für die Entscheidung sprechen, aufgewertet wurden, während die widersprechenden Informationen abgewertet wurden. Auch die Interkorrelation der Informationen ist dann im Posttest höher als im Pretest.

Experimente zu Präferenzentscheidungen haben Simon, Krawczyk et al. (2004) sowie Simon et al. (2008) durchgeführt. Hierbei handelte es sich um eine Entscheidung zwischen zwei Jobangeboten, die sich hinsichtlich verschiedener Attribute unterschieden. Diese Experimente zu Präferenzentscheidungen sind wie die Experimente zu rechtlichen Entscheidungen aufgebaut. Die Attribute werden als einzelne Aussagen dargeboten, deren Attraktivität und Wichtigkeit die Versuchspersonen sowohl in einem Pre- als auch in einem Posttest auf einer 10- bzw. 9-stufigen Skala einschätzen. Auch in diesen Experimenten ist der Vergleich der Einschätzungen der Attribute zum Pretest und zum Posttest entscheidend. Die Interpretation ist analog zu der oben dargestellten.

Nachdem eine kurze Einführung in das Phänomen der prädezisionalen Informationsverarbeitung gegeben worden ist, sollen im Folgenden zunächst Befunde vorgestellt werden, in denen Charakteristika der Aufgabe modifiziert wurden.

2.4.2 Charakteristika der Aufgabe

Eine Vielzahl an Studien hat sich der näheren Untersuchung der prädezisionalen Informationsverzerrung angenommen und hierbei verschiedene Charakteristika der Aufgabe manipuliert. Da Brownstein (2003) bereits einen umfassenden Überblick über entsprechende Studien gegeben hat (siehe Kapitel 2.4.1), werden im Folgenden hauptsächlich aktuelle Befunde vorgestellt, die sich mit dem Phänomen der prädezisionalen Reevaluationen als zentralen Bestandteil des Informationsverarbeitungsprozesses beschäftigen. Wenn es die Argumentation erfordert, wird jedoch auch Bezug zu den Studien, die bereits von Brownstein berücksichtigt wurden, genommen.

Für ein Verständnis und die Charakterisierung des Phänomens der prädezisionalen Reevaluation ist besonders die Frage relevant, ob als Aufgabe tatsächlich eine Entscheidung von Nöten ist. Brownstein (2003) interpretiert die Befunde dahingehend, dass das Treffen einer Entscheidung bzw. eine diesbezügliche Erwartung nicht notwendig für das Auftreten prädezisionaler Informationsverzerrungen ist, diese jedoch verstärkt. In zwei späteren Experimenten manipulierte er selbst unter anderem, ob eine Entscheidung erwartet wurde oder nicht, und konnte seine Einschätzung, dass die Erwartung einer Entscheidung einen moderierenden Effekt auf die Stärke des *Bias* hat, bestätigen (Brownstein, Read & Simon, 2004). Dieser Interpretation ist nichts entgegenzusetzen, doch für die Charakterisierung der Informationsverzerrung erscheint die Fokussierung auf den ersten Teil der Aussage besonders relevant: Für das Auftreten prädezisionaler Verzerrungen ist es *nicht* notwendig, dass das Treffen einer Entscheidung erwartet wird. Vielmehr resultieren sie bereits aus der Verarbeitung komplexer und widersprüchlicher Informationen. Diesbezüglich überzeugende Evidenz haben Russo et al. (1998), Simon et al. (2001) sowie Brownstein et al. (2004) erbracht. Simon et al. (2001) etwa instruierten Versuchspersonen, die Argumente eines rechtlichen Falls zu erinnern (Experiment 2) und fanden *coherence shifts*. In einem weiteren Experiment (Experiment 3) dachten die Versuchspersonen der ersten Bedingung, sie würden noch weitere Informationen erhalten (*reception condition*), diejenigen der zweiten Bedingung dachten, sie müssten die Argumente später kommunizieren können (*communication condition*), und diejenigen der dritten Bedingung wussten, dass es sich um eine Entscheidungsaufgabe handelt (*standard decision task*). In jeder Bedingung wurden deutliche *coherence shifts* gefunden. Dies veranschaulicht, dass die aus der Informationsverarbeitung resultierende Konstruktion von Konsistenz ein natürlicher Prozess ist. So folgern auch Brownstein et al. (2004) “that some degree of bias may be an integral part of

processing" (S. 902). Gestützt wird diese Interpretation zum einen durch Befunde, nach denen keine bestehende Präferenz nötig ist, um eine Reevaluation zu beobachten (Russo et al., 1996; Simon, Krawczyk et al., 2004), und zum anderen durch das Ergebnis, dass bereits bei der Beurteilung einer einzigen Option Informationsverzerrungen auftreten (Bond et al., 2007). Somit scheint die durch die Reevaluation erzielte Steigerung von Konsistenz nicht nur eine natürliche Folge des Entscheidungsprozesses, sondern allgemein der Informationsverarbeitung zu sein. Zu ergänzen ist, dass die Reevaluationen unbewusst stattfinden (z.B. Holyoak und Simon, 1999; Russo et al., 2006; Russo et al., 2000).

Doch wie generalisierbar sind diese Befunde? Die Tatsache, dass die prädezisionale Reevaluation von Informationen für verschiedene Entscheidungsaufgaben nachgewiesen werden konnte, spricht für die allgemeine Gültigkeit dieses Phänomens. Den Nachweis für Präferenzentscheidungen erbrachten insbesondere Russo und seine Kollegen (Meloy, 2000; Meloy und Russo, 2004; Russo et al., 2006; Russo et al., 2008; Russo et al., 1996; Russo et al., 1998; Russo et al., 2000), doch auch Simon und Kollegen (Simon et al., 2008; Simon, Krawczyk et al., 2004) sowie – unter weitestgehender Verwendung des typischen Forschungsparadigmas der Gruppe um Russo – Polman (2010). Betsch (2005b) macht auf das Auftreten von Informationsverzerrungen in Routineentscheidungen aufmerksam. Bei einer starken Routine können Informationen, die der Wahl dieser Routine widersprechen, abgewertet werden. Wie Betsch ausführt, wird dieser Effekt durch Zeitdruck verstärkt, und benennt damit eine weitere zu berücksichtigende Moderatorvariable. Eine Domäne, in der eigentlich möglichst objektive Entscheidungen erwartet werden, diente besonders häufig als Untersuchungsgegenstand: rechtliche Entscheidungen. Es konnte überzeugende Evidenz erbracht werden, dass Informationen in Richtung des später getroffenen Urteils verzerrt werden, und zwar sowohl hinsichtlich der Gewichtung (Pennington und Hastie, 1988) als auch hinsichtlich der inhaltlichen Interpretation (Carlson und Russo, 2001; Holyoak und Simon, 1999; Hope, Memon & McGeorge, 2004; Simon et al., 2001; Simon, Snow et al., 2004). Für den medizinischen Anwendungskontext wiesen Kostopoulou, Mousoulis und Delaney (2009) in einer Studie mit Experten Informationsverzerrungen nach. Je nach dem, welche Diagnose die Ärzte stellten, wurden die Informationen unterschiedlich gesucht und interpretiert. Bond et al. (2007) sowie DeKay, Patiño-Echeverri und Fischbeck (2009) konnten zeigen, dass selbst eindeutige Wahrscheinlichkeitsinformationen im Entscheidungsprozess systematisch verzerrt werden. Wie DeKay et al. (2009) herausstellen, konnte damit zugleich der wichtige Nachweis erbracht werden, dass Ambiguität für das Auftreten von

Informationsverzerrungen nicht notwendig ist. Sie ziehen aber eine moderierende Wirkung in Betracht (vgl. auch Simon, 2004). Dies steht in Einklang mit dem Ergebnis, dass mit steigender Diagnostizität der Information die Verzerrungen schwächer werden (Russo et al., 1998). Die jüngste Erweiterung stammt von Glöckner et al. (im Druck). Sie konnten unter Verwendung eines Pretest/Posttest-Designs in drei Experimenten zeigen, dass auch in probabilistischen Inferenzaufgaben eine systematische Reevaluation von Informationen, genauer gesagt von subjektiv eingeschätzten Cue-Validitäten, stattfindet. Je nach dem, welche Option später gewählt wurde, wurden in diesem Experiment die Validitäten der jeweiligen Cues auf- bzw. abgewertet. Die aktuellen Studien von Bond et al. (2007), DeKay et al. (2009) sowie Glöckner et al. (im Druck) weisen eine Gemeinsamkeit mit derjenigen von Kostopoulou et al. (2009) auf, die letztere für ihre Arbeit hervorheben: Es gibt eine objektiv richtige Antwort, so dass das Auftreten von Informationsverzerrungen auch auf diesen Bereich generalisiert werden kann. Somit scheint unabhängig von der Entscheidungsaufgabe zu gelten, dass Informationen systematisch verzerrt werden, ohne dass diese komplex oder widersprüchlich sind. Erneut wird deutlich, wie fundamental dieser Prozess zu sein scheint.

Wie bereits angesprochen, setzt das Auftreten von Informationsumwertungen im Entscheidungsprozess nicht voraus, dass vor der Evaluation der Informationen bereits eine Präferenz bzw. Favorisierung besteht. Dennoch konnte die Induktion einer Neigung interessante Erkenntnisse über den Prozess der prädezisionalen Reevaluation von Informationen liefern. So konnten Simon, Snow et al. (2004) in ihren Experimenten zu rechtlichen Entscheidungen zeigen, dass nicht nur die Manipulation eines Fakts das Urteil und damit auch die Einschätzung der übrigen Fakten beeinflusst, sondern dass sogar nur die Zuweisung zu einem Urteil diesen Effekt bewirkt. Damit sehen Simon, Snow et al. ihre Einschätzung bekräftigt, dass der Entscheidungsprozess durch das Wirken bidirektionaler Prozesse, die in der Erhöhung von Konsistenz resultieren, charakterisiert ist. Den für den rechtlichen Kontext äußerst beachtenswerten Befund, dass vor der Verhandlung veröffentlichte, für den Angeklagten nachteilige Informationen sowohl das Urteil als auch die Stärke der Informationsverzerrungen systematisch beeinflussen, konnten Hope et al. (2004) erbringen. Russo et al. (2006) konnten sogar nachweisen, dass durch eine gezielte Anordnung der Informationen eine zuvor als unterlegen eingeschätzte Alternative gewählt wird. Diese Wahl ging mit einer Verzerrung der Informationen zugunsten der gewählten Alternative einher. Wie die Autoren hervorheben, konnten sie zeigen, dass diese Verzerrung

inhaltlicher Art war. Das heißt, dass die Versuchspersonen die Informationen nicht nur abwerteten, sondern anders, und zwar im Sinne der favorisierten Alternative, interpretierten.

Wenn die prädezisionale Umwertung von Informationen sogar derart negative Folgen haben kann, stellt sich erst recht die Frage, welche Möglichkeiten des *Debiasing* es gibt. In zahlreichen Arbeiten wird dieses Problem erkannt und diskutiert (z.B. Bond et al., 2007; Carlson und Pearo, 2004; Carlson und Russo, 2001; Larson und King, 1996; Polman, 2010; Russo et al., 1998; Russo et al., 2000; Simon, 2004). Es werden verschiedene Vermutungen angestellt, wie die prädezisionale Verzerrung gemindert werden kann. Carlson und Russo (2001) sehen das Hauptproblem in dem fehlenden Bewusstsein, dass solche Umwertungen stattfinden können (vgl. Holyoak und Simon, 1999; Russo et al., 2006; Russo et al., 2000). Folglich schlagen sie vor, dieses Problem bewusst zu machen und davor zu warnen. Bond et al. (2007) halten die Konfrontation mit Informationen, die den unmittelbar vorher präsentierten Informationen widersprechen, für eine mögliche Maßnahme. Übertragen auf eine Entscheidung zwischen zwei Telefonen A und B würde dies bedeuten, dass nachdem z.B. die gute Gesprächsqualität von Telefon A begründet wurde, Beispiele für die schlechte Gesprächsqualität dieses Telefons aufgeführt werden. Letztlich veranlassen Bond und Kollegen damit die Versuchspersonen zur Anwendung der *consider-the-opposite*-Technik (Lord, Lepper & Preston, 1984) pro Attribut, indem sie eine gezielte Präsentation von Informationen vorschlagen. Eine unverzerrte Einschätzung von Attributen kann nach Russo et al. (1998) dadurch erreicht werden, dass die Attribute stets für ein neues Alternativenpaar eingeschätzt werden. Bei einer Entscheidung zwischen Alternativen halten sie eine Variation des Aggregationslevels der Attribute für sinnvoll, indem etwa aus zwei Attributen ein übergreifendes Attribut gebildet wird und umgekehrt. Dadurch sollen das Auftreten und das Ausmaß der Verzerrungen eingeschätzt werden können. Dieser Vorschlag zielt allerdings eher auf die Kontrolle von Umwertungen im Forschungskontext ab, nicht auf deren Linderung. Aus Polmans (2010) Befunden lässt sich ebenfalls eine Idee zur Reduzierung prädezisionaler Informationsverzerrungen ableiten. Polman unterscheidet zwischen Personen, die für sich entscheiden sollen (*personal decision maker*), und Personen, die für andere entscheiden sollen (*proxy decision maker*). Dass *personal decision maker* weniger prädezisionale Umwertungen zeigen als *proxy decision maker*, gibt laut Polman einen Anhaltspunkt für eine *Debiasing*-Strategie. Allerdings ist dadurch das Problem nur verlagert, da Polman ebenfalls zeigen konnte, dass *personal decision maker* stärkere postdezisionale Umwertungen vornehmen als *proxy decision maker*. Ein weiterer Vorschlag

stammt von Larson und King (1996). Sie diskutieren das Problem der Informationsverzerrungen für Organisationen und nehmen dabei eine systemische Perspektive ein. Um dieses Problem zu lindern, plädieren sie für eine Etablierung einer Lernkultur. So sollte nicht nur die Weitergabe guter Neuigkeiten belohnt werden, sondern ebenso diejenige schlechter Nachrichten. Damit werden die Ziele der Genauigkeit und des Lernens aus Fehlern verfolgt, anstatt den Fokus nur auf Ergebnisse und Erfolg zu legen. Inwiefern diese Idee, deren Realisierung – wie Larson und King einräumen – in Organisationen problematisch sein kann, aber auf eine Einzelperson zu übertragen ist, bleibt fraglich. Wenige Studien werden berichtet, in denen die Frage des *Debiasing* hinsichtlich prädezisionaler Informationsverzerrungen experimentell untersucht wurde. Simon (2004) berichtet Evidenz, nach der die Instruktion, nicht voreingenommen zu sein, keinen Effekt hatte, wohingegen die Instruktion, das Gegenteil zu berücksichtigen, zu einer Verringerung von *coherence shifts* führte. Carlson und Pearo (2004) konnten in zwei Experimenten zeigen, dass wenn Personen vor einer Entscheidungsaufgabe bereits die Komponenten der einzelnen Attribute evaluiert haben, die prädezisionale Reevaluation geringer ausfällt. Die Hypothese, dass systematische Verzerrungen reduziert werden können, wenn die Personen für ihre Entscheidungen verantwortlich gemacht werden, konnte hingegen nicht bestätigt werden (Russo et al., 2000). Zwar geben die dargestellten Moderatoren einen Anhaltspunkt dafür, wann diese Reevaluationen stärker und wann sie schwächer sind, doch konnten bisher nur wenige effektive Methoden des *Debiasing* ermittelt werden. Carlson und Pearo (2004) berichten von einigen – nicht publizierten – gescheiterten Versuchen, mittels direkter Warnungen oder Anreize für genaue Entscheidungen dem *Bias* entgegenzuwirken. Es bietet sich das Bild, dass es schwierig ist, geeignete *Debiasing*-Strategien zu finden. Dies erinnert an verwandte Phänomene wie den *hindsight bias*, bei dem die gleiche Schwierigkeit zu beobachten ist (Fischhoff, 1982).

In jüngerer Zeit konnten weitere Moderatoren identifiziert werden. Eine Variable, deren Einfluss von Brownstein (2003) – wie bereits angesprochen – nicht abschließend geklärt werden konnte, betrifft die Konsequenzen einer Entscheidung. Angesichts der uneinheitlichen Befundlage stellt er die Vermutung auf, dass bei einer Wahl zwischen Alternativen mit *unattraktiven* Konsequenzen höhere Konsequenzen zu einer stärkeren prädezisionalen Informationsverzerrung führen, während bei einer Wahl zwischen Alternativen mit *attraktiven* Konsequenzen niedrigere Konsequenzen zu einer stärkeren Umwertung führen. Der Frage nach dem Einfluss von Konsequenzen ist Brownstein dann

selbst weiter nachgegangen: Brownstein et al. (2004) haben manipuliert, ob den Versuchspersonen positive oder keine Konsequenzen einer Entscheidung in Aussicht gestellt wurden. Ein Kernbefund war, dass die Aussicht auf positive Konsequenzen im Vergleich zu keinen Konsequenzen zu einer stärkeren Favorisierung der gewählten Alternative im Entscheidungsprozess führte. Daraus folgerten die Autoren, dass die Erwartung von hohen Konsequenzen, unabhängig von ihrer Valenz, zu höheren prädezisionalen Informationsverzerrungen führt. Verstärkt werden können die Reevaluationen auch dann, wenn die Versuchspersonen in eine positive Stimmung versetzt werden (Meloy, 2000) und wenn eine Entscheidung für eine andere Person anstatt für sich selbst getroffen wird (Polman, 2010). Des Weiteren wiesen Meloy und Russo (2004) in drei Experimenten einen Zusammenhang zwischen der Valenz der Alternativen, der Wahlaufgabe und der Stärke der Umwertung nach: Wenn eine Entscheidung zwischen zwei positiven Alternativen getroffen werden soll, ist die Verzerrung stärker, wenn eine Alternative *ausgewählt* werden soll. Wird hingegen eine Entscheidung zwischen zwei negativen Alternativen getroffen, so finden stärkere Umwertungen statt, wenn eine Alternative *abgelehnt* werden soll.

Angesichts der Allgegenwärtigkeit prädezisionaler Informationsverzerrungen, den bisher beschränkten Möglichkeiten des *Debiasing* und der möglichen negativen Auswirkungen erscheint die Arbeit von Simon und Kollegen (2008) besonders relevant. Sie untersuchten die Dauer von *coherence shifts* und fanden, dass diese bereits nach 15 Minuten vollständig verschwunden sind. Daraus folgern die Autoren, dass keine Übertragung auf nachfolgende Entscheidungen zu befürchten ist. Allerdings räumen sie ein, dass unter bestimmten Bedingungen, nämlich wenn Situationen neuartig sind, wiederholt werden oder aber materielle Einsätze beinhalten, *coherence shifts* andauern könnten. Was allerdings nachgewiesen werden konnte, war ein Transfereffekt: War ein entsprechend dem Urteil umgedeutetes Argument Bestandteil eines weiteren Falls, wirkte sich die Umwertung auf das Urteil in diesem Fall sowie die Einschätzung der (neuen) Fakten aus (Holyoak und Simon, 1999).

Zusammenfassend ist festzuhalten, dass eine Vielzahl aktueller Studien, in denen Charakteristika der Aufgabe untersucht wurden, zu einem besseren Verständnis prädezisionaler Informationsverzerrungen beigetragen hat. So wird die Charakterisierung dieses Phänomens als (nahezu) allgegenwärtig und als natürliche Folge der Informationsverarbeitung gestützt. Darüber hinaus konnten weitere Variablen identifiziert werden, die einen moderierenden Effekt haben. Studien, in denen ein *Debiasing* erfolgreich war, sind bisher

eher rar, Ideen stehen jedoch zur Verfügung. Doch auch der Einfluss weiterer Moderatorvariablen oder die Frage nach der Persistenz von Informationsverzerrungen bedürfen der Klärung, um das Verständnis des Phänomens der prädezisionalen Informationsverzerrung und des damit verbundenen Strebens nach Konsistenz zu vertiefen.

2.4.3 Charakteristika der Person

Neben den Charakteristika der Aufgabe stellen auch bestimmte Charakteristika der Person Variablen dar, für die ein Zusammenhang mit dem Auftreten und der Stärke der prädezisionalen Informationsverzerrung vermutet werden kann. Der Begriff *Charakteristika der Person* wird hierbei im Folgenden breit gefasst, indem sowohl Persönlichkeitseigenschaften als auch demografische Variablen wie Alter, Bildung und Beruf hierunter subsumiert werden. Doch trotz dieser breiten Fassung lassen sich nur wenige Studien finden, die den Einfluss von Personvariablen auf die Reevaluation von Informationen im Entscheidungsprozess untersucht haben. Diese sollen im Folgenden kurz vorgestellt werden.

Die einzigen Personvariablen, für die Brownstein (2003) in seinem Review einen moderierenden Einfluss berichtet, sind Handlungsorientierung und die Disposition zu guter Stimmung. So konnten Beckmann und Kuhl (1984) zeigen, dass eine höhere Handlungsorientierung mit einer stärkeren Reevaluation der Alternativen einhergeht. Diesen Einfluss fanden sie speziell für eine Unterskala, nämlich der Handlungsorientierung bei der Handlungsplanung, was die Autoren auf die Tatsache, dass es sich im Experiment um eine Entscheidungsaufgabe handelte, zurückführen. Russo et al. (2000) wiesen einen positiven Zusammenhang zwischen der Disposition zu guter Stimmung und dem Ausmaß an Informationsumwertung bei Verkäufern nach.

Da die Reevaluation von Informationen zu einer Erhöhung von Konsistenz führt, lassen sich insbesondere Zusammenhänge zwischen der Stärke der Umwertungen und einem Bedürfnis nach Konsistenz bzw. einer Abneigung gegenüber Inkonsistenz annehmen (vgl. Simon, 2004). Simon vermutet einen moderierenden Effekt personaler Faktoren und weist insbesondere auf individuelle Unterschiede bezüglich der Toleranz für Inkonsistenz hin. Entsprechende Variablen hat Meloy (2000) in ihrem Experiment berücksichtigt, und zwar den *Need for Cognitive Closure* (Webster und Kruglanski, 1994), den *Personal Need for Simple Structure* (Neuberg und Newsom, 1993) sowie den *Need to Evaluate* (Jarvis und Petty, 1996). In ihrem Experiment konnte sie für keinen Faktor einen Einfluss nachweisen. Ob die *Preference For Consistency* (PFC, Cialdini, Trost & Newsom, 1995) die Stärke der

Informationsverzerrung moderiert, haben Russo et al. (1998) untersucht. Sie fanden jedoch keinen Zusammenhang zwischen diesem Faktor und der Reevaluation von Informationen. Demnach ist die Evidenz zu Faktoren, die einen Wunsch nach Konsistenz bzw. eine Intoleranz von Inkonsistenz erfassen, überraschend negativ.

Eine Variable, die relativ häufig untersucht wurde, ist die Expertise. So könnte man vermuten, dass Experten angesichts ihrer Fachkenntnis und ihrer Erfahrung vor diesem *Bias* geschützt sind. Die Empirie zeigt aber ein anderes Bild: Kostopoulou et al. (2009) berichten Ergebnisse, nach denen Mediziner bei der Diagnose einer Krankheit Informationen verzerren. Kostopoulou et al. betrachten den prädezisionalen Bias allerdings nur für die Ärzte, die ein falsches Diagnoseurteil abgegeben hatten, und hatten zudem keine Kontrollgruppe. Den Vergleich mit einer Kontrollgruppe, die keine Entscheidung treffen musste, haben hingegen Russo et al. (2000) in einer Untersuchung mit Verkäufern und Wirtschaftsprüfern ermöglicht. Die Experimentalgruppe musste eine berufsspezifische (für beide Gruppen unterschiedlich) und eine beruflich relevante (für beide Gruppen gleich) Entscheidung treffen. Beide Berufsgruppen zeigten die typischen Informationsumwertungen, wobei das Ausmaß an Berufserfahrung keinen Einfluss hatte. Brownstein et al. (2004) fragten Versuchspersonen nach ihrer Expertise hinsichtlich Pferderennen und fanden ebenfalls, dass „Experten“ stärkere Informationsverzerrungen zeigen, was sich in ihrer Studie als ein im Vergleich zu den Laien höheres Rating der gewählten Alternative ausdrückte. Eine Untersuchung mit angehenden Geschworenen führten Carlson und Russo (2001) durch. Im Vergleich zu Studenten werteten Geschworene die Informationen stärker um, ließen sich in ihrem Entscheidungs- und damit auch im Informationsbewertungsprozess stärker von vorherigen Einstellungen gegenüber der Verteidigung bzw. der Anklage leiten und waren sich schlussendlich sogar sicherer in ihrer Entscheidung. Da der Expertenstatus der Geschworenen aber angezweifelt werden kann, bleibt Raum für Alternativerklärungen. Carlson und Russo selbst halten einen Einfluss des Alters für denkbar, DeKay et al. (2009) vermuten angesichts ihrer eigenen Befunde (s.u.) darüber hinaus einen Einfluss der Bildung.

Als demografische Variablen wurden der Beruf, das Alter und die Bildung untersucht. In der bereits erwähnten Studie von Russo et al. (2000) untersuchten die Autoren auch, ob bestimmte Berufsgruppen zu einer stärkeren (Verkäufer) oder schwächeren (Wirtschaftsprüfer) Reevaluation neigen. Hierzu verglichen sie die Evaluation der Attribute, auf Basis derer eine Präferenzentscheidung zu treffen war, mit derjenigen einer studentischen Stichprobe, die diese Aufgabe in einer Vorstudie bearbeitet hatte. Die Berufsgruppen

unterschieden sich hinsichtlich des Ausmaßes an Umwertungen nicht von der studentischen Stichprobe, allerdings fanden die Autoren einen moderierenden Effekt des Berufs, wenn zusätzlich die beiden professionellen Entscheidungen berücksichtigt wurden: Verkäufer verzerrten die Informationen in einem stärkeren Ausmaß als Wirtschaftsprüfer. Wenn man berücksichtigt, dass eine Berufsentscheidung auch von Persönlichkeitseigenschaften wie Extraversion abhängt, lassen sich weitere Einflussfaktoren erahnen. DeKay et al. (2009) wiesen einen positiven Zusammenhang zwischen dem Alter und dem Ausmaß an Informationsverzerrungen sowie einen negativen Zusammenhang zwischen dem Bildungsniveau und diesem Ausmaß nach. Somit weisen vereinzelte Studien auf den Einfluss verschiedener Personvariablen hin, in anderen Studien bleiben erwartete Effekte aus. Die Suche nach personalen Einflussfaktoren widerspricht nicht der obigen Auffassung, nach der das Erhöhen von Konsistenz als eine natürliche Folge des Informationsverarbeitungsprozesses zu betrachten ist, da individuellen Unterschieden eher eine moderierende Funktion (vgl. Meloy, 2000; Russo et al., 1998; Simon, 2004) zugeschrieben werden kann. Folglich bietet sich hier ein spannendes Forschungsfeld.

Zum Abschluss dieses Kapitels 2.4 sollen die zentralen Forschungsbefunde noch einmal tabellarisch so zusammengefasst werden, dass die veränderte Interpretation prädezisionaler Informationsverzerrungen verdeutlicht wird (siehe Tabelle 1). Folglich erhebt die Tabelle nicht den Anspruch eines umfassenden Überblicks, sondern fokussiert auf die Befunde, die zur gegenwärtigen Charakterisierung prädezisionaler Informationsverzerrungen geführt haben.

Tabelle 1: Überblick über die veränderte Interpretation prädezisionaler Reevaluationen

INFORMATIONSUMWERTUNGEN FINDEN NUR NACH DER ENTSCHEIDUNG STATT.		
Schwerpunkte	*Zentrale Befunde*	*Studien*
Postdezisionale Dissonanz	Voraussetzungen der Dissonanzreduktion: Entscheidung, commitment; Konflikt vs. Dissonanz	Festinger (1957, 1964)
PRÄDEZISIONALE INFORMATIONSUMWERTUNGEN UNTERSTÜTZEN DEN ENTSCHEIDUNGSPROZESS.		
Schwerpunkte	*Zentrale Befunde*	*Studien*
Rekonstruktionen	Durch Umstrukturierung des Entscheidungsproblems werden passende Entscheidungsregeln gefunden.	Payne et al. (1988); Payne et al. (1992)
Inhaltliche Umwertungen	Umwertungen zur Vermeidung von Konflikt	Janis & Mann (1977); Mann et al. (1969)
	Umwertungen zur Erhöhung von Sicherheit	Mills (1968); Mills & O'Neal (1971); O'Neal (1971)
	Umwertungen zum Erlangen von Handlungsfähigkeit	Beckmann & Kuhl (1984)
PRÄDEZISIONALE INFORMATIONSUMWERTUNGEN SIND ZENTRAL IM INFORMATIONSVERARBEITUNGSPROZESS.		
Schwerpunkte	*Zentrale Befunde*	*Studien*
Voraussetzungen	Keine a priori Präferenz bzw. Hypothese nötig	Russo et al. (1998); Simon et al. (2001); Simon, Krawczyk et al. (2004)
	Keine Erwartung einer Entscheidung nötig	Simon et al. (2001); Brownstein et al. (2004)
	Unbewusster Prozess	Holyoak & Simon (1999); Russo et al. (2006); Russo et al. (2000)
	Auftreten auch bei einer Option	Bond et al. (2007)
	Ambigue Informationen nicht nötig	Bond et al. (2007); DeKay et al. (2009); Glöckner et al. (im Druck)
	Kein Zsh. mit versch. Personcharakteristika	Meloy (2000); Russo et al. (1998)
	Auftreten in versch. Entscheidungsaufgaben	u.a. Bond et al. (2007); DeKay et al. (2009); Glöckner et al. (im Druck); Holyoak & Simon (1999); Kostopoulou et al. (2009); Larson & King (1996); Russo et al. (1998); Simon, Krawzcyk et al. (2004)
Faktoren, die die Umwertungen verstärken	Treffen einer Entscheidung erwartet, Erwartung der Rechtfertigung, ähnliche Alternativen, weniger Zeit bis zur Entscheidung, keine Erwartung weiterer Informationen, längere Deliberation, unerklärtes arousal, Handlungsorientierung, Annehmbarkeit der Informationen, Instrumentalität der Informationen, dispositionell pos. Stimmung, homogene Gruppe	Review von Brownstein (2003)
	Ambigue Informationen	DeKay et al. (2009); Russo et al. (1998); Simon (2004); Simon, Krawczyk et al. (2004)
	Erwartung von Konsequenzen	Brownstein et al. (2004)
	Zeitdruck	Betsch (2005b)
	Proxy decision making	Polman (2010)
	Positive Stimmung	Meloy (2000)
	Wahl einer von zwei positiven Konsequenzen, Ablehnung einer von zwei negativen Konsequenzen	Meloy & Russo (2004)
	Hohes Alter, niedriges Bildungsniveau	DeKay et al. (2009)
Debiasing	Consider the opposite verringert Umwertungen	Simon (2004)
	Vorh. Evaluation von Attributskomponenten reduziert Bias	Carlson & Pearo (2004)
	Übernahme von Verantwortung hat keinen Effekt	Russo et al. (2000)
Dauer	Umwertungen sind temporär	Simon et al. (2008)

Anmerkungen: Fett hervorgehoben sind die jeweils vorherrschenden Leitbilder.

2.5 ZUSAMMENFASSUNG UND SCHLUSSFOLGERUNG

In diesem Kapitel werden die zentralen Aspekte der vorherigen Darstellungen zusammengefasst. Aufbauend auf diesen Ausführungen, die das in den letzten Jahrzehnten vertiefte Verständnis von Konsistenz ausdrücken, wird eine Arbeitsdefinition von Konsistenz vorgeschlagen. Des Weiteren lassen sich aus den theoretischen und empirischen Beiträgen Schlussfolgerungen für die Forschung ableiten. Diese Schlussfolgerungen münden abschließend in einer präzisen Fragestellung für die vorzustellenden empirischen Forschungsarbeiten.

Die Ausgangsbasis für die Thematisierung bzw. Ergründung von Konsistenz war die folgende Umschreibung: Konsistenz bezeichnet einen angenehmen Zustand der Einheitlichkeit, des Zusammenhangs und der Widerspruchsfreiheit von Elementen, der durch Restrukturierungen und Reinterpretationen dieser Elemente angestrebt wird. Es konnte herausgestellt werden, dass der Aspekt des Konsistenzstrebens bereits in der Gestaltpsychologie, insbesondere in dem zentralen Prinzip der Tendenz zur guten Gestalt (Wertheimer, 1923), zum Ausdruck kam. Ein damit verbundener Gesichtspunkt ist die Einnahme einer holistischen Perspektive, die schließlich für ein Erkennen von und Streben nach Konsistenz unabdingbar ist. Insbesondere methodische Schwächen hielten den gestaltpsychologischen Ansatz von einer Konkretisierung des Konsistenzstrebens ab. Temporäre Abhilfe schufen kognitive Konsistenztheorien (z.B. Abelson und Rosenberg, 1958; Festinger, 1957; McGuire, 1968), die auf der Gestaltpsychologie aufbauten. Das Konsistenzprinzip stellte in diesen Theorien den zentralen Untersuchungsgegenstand dar, obgleich sich die Theorien dahingehend unterschieden, ob die Minimierung von Inkonsistenz oder das Maximieren von Konsistenz im Vordergrund standen. Als ein Vorreiter für die spätere Konsistenzforschung sollte sich insbesondere McGuire (1968) erweisen, der auf die strukturelle Dynamik eines Systems fokussierte und hierbei das Streben nach Konsistenz als gegeben betrachtete. Allerdings mangelte es McGuires Ansatz wie den übrigen Konsistenztheorien an einem adäquaten Umgang mit Komplexität, der idealerweise in einem umfassenden formalen Modell hätte ausgedrückt werden sollen. Dies leisteten erst die sogenannten PCS-Modelle (z.B. McClelland und Rumelhart, 1981; Read und Miller, 1998a; Thagard, 1989). Die PCS-Ansätze stellten einen wichtigen Fortschritt in der Konsistenzforschung dar. Sie erlaubten eine formale Darstellung der gestaltpsychologischen Prinzipien und konnten dabei sowohl die Dynamik als auch die Komplexität der menschlichen Informationsverarbeitung abbilden.

Aufbauend auf den Arbeiten von McClelland, Rumelhart und der PDP Research Group (McClelland, Rumelhart & The PDP Research Group, 1986; Rumelhart, McClelland et al., 1986) trugen insbesondere die Forschergruppen um Paul Thagard und Stephen Read zur Verbreitung von PCS-Ansätzen bei. In den letzten drei Jahrzehnten wurde eine Vielzahl an Modellen entwickelt. Die breite Anwendbarkeit von PCS-Modellen und deren Mehrwert lassen die aktuelle Tendenz, nicht nur Modelle für bestimmte Phänomene, sondern für ganze Bereiche (z.B. Einstellungen, Persönlichkeit) zu entwickeln, als logische Konsequenz erscheinen. Auch in der Entscheidungsforschung werden immer häufiger PCS-Prozesse zur Erklärung des Entscheidungsprozesses erfolgreich herangezogen. Dieser Entwicklung ging zunächst die Durchsetzung einer konstruktiven Sichtweise des Entscheidungsprozesses voraus, die Anfang der 90er Jahre als neuer Trend in der Entscheidungsforschung bezeichnet wurde. Einige Entscheidungsansätze machten diesen konstruktiven Aspekt zum Kern ihrer Arbeiten, doch erst PCS-Ansätze ermöglichten eine formale und damit konkrete Erfassung der komplexen dynamischen Prozesse. Ein PCS-Modell für probabilistische Inferenzentscheidungen haben Glöckner und Betsch (2008b) entwickelt und getestet. In Abgrenzung zu anderen Entscheidungsmodellen erachten sie den Konsistenzmaximierungsprozess als Kernmechanismus jedes Entscheidungsprozesses. Dieser läuft automatisch ab. Weitere Entscheidungsstrategien kommen lediglich dann zum Einsatz, wenn der automatische Informationsverarbeitungsprozess nicht in einer ausreichend konsistenten Lösung mündet. Glöckner et al. (im Druck) konnten in mehreren Experimenten zeigen, dass eine systematische Reevaluation von Informationen im Entscheidungsprozess stattfindet, die das Maximieren von Konsistenz ausdrückt. Sie haben damit zum Verständnis der prädezisionalen Informationsverzerrung als Kennzeichen des Konsistenzstrebens beigetragen. Dass es prädezisionale Reevaluationen tatsächlich gibt, kann spätestens seit dem Review von Brownstein (2003) nicht mehr abgestritten werden. Die Studien der letzten Jahre haben sich der Untersuchung des zugrunde liegenden Prozesses weiter angenommen, und zwar unter Verwendung verschiedener methodischer Designs. Die Paradigmen von Russo und Kollegen sowie von Simon und Kollegen sind besonders populär. Die Befunde stützen die Interpretation, dass prädezisionale Informationsverzerrungen eine automatische Folge des Informationsverarbeitungsprozesses darstellen und unbewusst stattfinden. So bedarf es etwa weder der Erwartung, eine Entscheidung zu treffen, noch einer bestehenden Präferenz, damit Umwertungen auftreten. Auch die weite Verbreitung dieses Phänomens spricht dafür, dass es (nahezu) universell ist. Des Weiteren konnten weitere moderierende Faktoren identifiziert

werden. Personale Faktoren wurden selten und zumeist erfolglos untersucht. Angesichts möglicher negativer Konsequenzen der systematischen Umwertung von Informationen wird die Frage des Debiasing häufig diskutiert. Empirische Befunde, die auf eine erfolgreiche Methode verweisen, sind jedoch rar. Eine interessante neue Forschungsrichtung betrifft das Andauern prädezisionaler Informationsumwertungen. Entsprechende Ergebnisse von Simon und Kollegen (2008) deuten auf den temporären Charakter der systematischen Reevaluationen hin, wobei aber die Autoren selbst ein Andauern unter bestimmten Bedingungen für möglich halten.

Die entsprechenden Befunde wurden unter Verwendung verschiedener Paradigmen ermittelt, denen unterschiedlich akzentuierte Auffassungen der prädezisionalen Informationsverzerrung zugrunde liegen. Während etwa Russo und Kollegen (z.B. Carlson und Russo, 2001) eine aus der seriellen Darbietung von Informationen resultierende Verzerrung postulieren, nehmen Simon und Kollegen (z.B. Holyoak und Simon, 1999) als Vertreter des PCS-Ansatzes eine parallele Verarbeitung an. Auch in dem Modell von Glöckner und Betsch (2008b) wird davon ausgegangen, dass Informationen parallel und kompensatorisch verarbeitet werden. Zahlreiche empirische Befunde rechtfertigen diese Annahme (Glöckner, 2006; Glöckner, 2008a; Glöckner und Betsch, 2008c). Es besteht damit in dieser Hinsicht ein direkterer Bezug zu den Arbeiten von Simon und Kollegen.

Vor dem Hintergrund der theoretischen Ausführungen und den empirischen Befunden soll nun eine erweiterte bzw. präzisierte Arbeitsdefinition von Konsistenz vorgeschlagen werden, die die Grundlage für die in den folgenden Kapiteln vorzustellenden Experimente bildet. Das zentrale Kennzeichen ist, dass das Erreichen oder Erhöhen von Konsistenz eine nicht bewusste und automatische Folge des Informationsverarbeitungsprozesses ist. Dass die Erhöhung von Konsistenz als fester Bestandteil des Informationsverarbeitungsprozesses angesehen wird, weist auf einen anderen, für die vorliegende Arbeit besonders bedeutsamen Aspekt hin: Dies ist die Annahme von Dynamik. So wird Konsistenz durch die Restrukturierung und Reevaluation von Informationen hergestellt. In der vorliegenden Arbeit steht der Aspekt der Reevaluation im Vordergrund.

Auf Basis dieser Ausführungen wird folgende Arbeitsdefinition von Konsistenz vorgeschlagen: *Konsistenz bezeichnet einen Zustand der Einheitlichkeit, des Zusammenhangs und der Widerspruchsfreiheit von Elementen in einer bestimmten Situation, der im Zuge des Informationsverarbeitungsprozesses automatisch durch Reevaluationen hergestellt wird.*

Nachdem die zentralen Aspekte der vorherigen Kapitel zusammengefasst und eine erweiterte Arbeitsdefinition von Konsistenz vorgeschlagen worden ist, soll nun die Fragestellung, die den vorzustellenden empirischen Forschungsarbeiten zugrunde liegt, präzisiert werden. Wie einleitend in Kapitel 2.1 bereits erwähnt wurde, soll die vorliegende Arbeit einen Beitrag zu der Erforschung des Strebens nach Konsistenz in der Entscheidungsforschung leisten, indem das Verständnis des Prozesses der Konsistenzmaximierung vertieft wird. Das Modell von Glöckner und Betsch (2008b) stellt die theoretische Basis dar, wobei lediglich das primäre Netzwerk fokussiert wird.[9] Als Entscheidungstyp wird die probabilistische Inferenzaufgabe betrachtet, für die die Autoren ein spezifiziertes PCS-Netzwerk vorgestellt haben. Der Bezug zu diesem PCS-Modell ist vor allem damit zu begründen, dass sich PCS-Modelle im Allgemeinen und dieses Modell im Speziellen als geeignet zur Erfassung der komplexen und dynamischen Informationsverarbeitung erwiesen haben. Dem angesprochenen Modell liegt eine konsistenzorientierte Sichtweise zugrunde, die formal dargestellt und durch zahlreiche Experimente und Simulationsstudien fundiert wurde. Zudem bezieht sich das Modell auf probabilistische Inferenzentscheidungen, die dadurch, dass es eine objektiv richtige Lösung gibt, weniger Raum als etwa Präferenzentscheidungen für individuelle Interpretationen lassen. Befunde zu systematischen Informationsverzerrungen belegen daher umso eindrücklicher das Wirken eines fundamentalen Konsistenzmaximierungsprozesses. Um das Verständnis des Konsistenzmaximierungsprozesses zu vertiefen, wird vor dem Hintergrund des dargestellten Forschungsstandes zunächst eine Frage hinsichtlich der zu verwendenden Methode geklärt, bevor auf inhaltlicher Ebene Charakteristika der Aufgabe und der Person als moderierende Faktoren untersucht werden. Die methodische Frage betrifft das Untersuchungsparadigma: Glöckner et al. (im Druck) haben in ihren Experimenten ein an das Design von Simon und Kollegen (siehe Kapitel 2.4.1) angelehntes Pretest-Posttest-Design gewählt. Ungeklärt ist jedoch, ob die Bearbeitung des Pretests einen Einfluss auf die nachfolgenden Einschätzungen hat. Dies gilt es zu klären, um die Verwendung des Paradigmas in weiteren Experimenten zu rechtfertigen. Um ein vertieftes Verständnis des Konsistenzmaximierungsprozesses zu erlangen, erscheint eine Ausweitung des Forschungsdesigns auf verschiedene Informationslagen angebracht. Dadurch kann außerdem ein möglicher Einflussfaktor bezüglich der Stärke der Informationsumwertung untersucht werden: Wie bereits dargestellt, konnte die Ambiguität einer Information als Moderator identifiziert werden. Offen ist jedoch, inwiefern eine ambigue, im

[9] Im Folgenden wird der Einfachheit halber auch kurz vom PCS-Modell gesprochen.

Sinne einer ausgeglichenen Informationslage in probabilistischen Inferenzaufgaben moderierend wirkt. Als weiteres Charakteristikum der Situation soll der bisher nicht ausreichend geklärte Einfluss der Wiederholung einer Entscheidungssituation untersucht werden. Auf der Seite der Personcharakteristika sollen verschiedene Eigenschaften, für die ein Zusammenhang mit der Stärke der Informationsumwertung vermutet werden kann oder in anderen Entscheidungs- und Aufgabentypen gefunden wurde, betrachtet werden.

Zusammengefasst lässt sich die Fragestellung dieser Arbeit folgendermaßen präzisieren: Das Ziel besteht darin, das Verständnis der Konsistenzmaximierung im Entscheidungsprozess zu vertiefen, indem insbesondere sowohl Charakteristika der Situation als auch der Person hinsichtlich ihres Einflusses auf die Stärke prädezisionaler Informationsverzerrungen in probabilistischen Inferenzentscheidungen untersucht werden.

3. EMPIRISCHER TEIL

3.1 ÜBERBLICK ÜBER DEN EMPIRISCHEN TEIL

Um Charakteristika der Aufgabe und der Person hinsichtlich ihres Einflusses auf prädezisionale Informationsverzerrungen in probabilistischen Inferenzentscheidungen zu untersuchen, wurden empirische Arbeiten durchgeführt, die verschiedenen, aber aufeinander aufbauenden thematischen Blöcken zuzuordnen sind. Abbildung 2 gibt dies schematisch wieder.

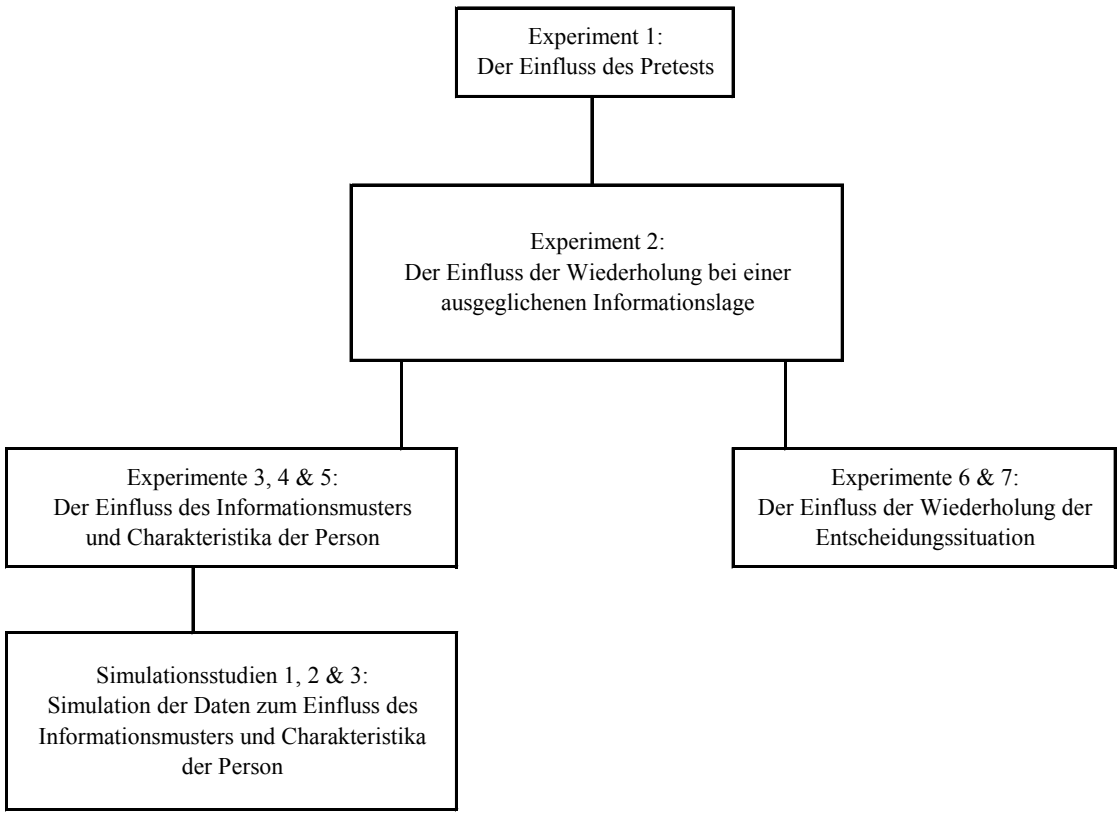


Abbildung 2: Schematischer Überblick über die empirischen Arbeiten

In Experiment 1 wird zunächst der wichtigen Frage nachgegangen, ob die zu beobachtende systematische Umwertung von Cue-Validitäten tatsächlich auf das Treatment, nämlich die Reflexion über eine Entscheidungsaufgabe, oder aber auf Pretest-Effekte oder die Interaktion zwischen Pretest und Treatment zurückzuführen ist. Da in den von Glöckner et al. (im Druck) berichteten Experimenten zur Umwertung von Cue-Validitäten in

probabilistischen Inferenzentscheidungen stets eine unausgeglichene Informationslage verwendet wurde, wird in Experiment 2 der Versuch unternommen, diese Befunde auf Entscheidungsaufgaben mit einer ausgeglichenen Informationslage auszuweiten. Darüber hinaus wird der Einfluss der Wiederholung auf diese prädezisionale Informationsverzerrung untersucht. Im anschließenden Experimentalteil wird auf die Bedeutung des Informationsmusters sowie auf die Relevanz von Charakteristika der Person fokussiert. Aufbauend auf den Ergebnissen von Experiment 2 wird in den Experimenten 3 und 4 der Einfluss der Ausgeglichenheit der Informationslage systematisch ergründet. In Experiment 4 wird bereits eine Erweiterung um personale Faktoren vorgenommen, deren Untersuchung in Experiment 5 weiter verfolgt wird. Die nachfolgenden Simulationsstudien dienen der Überprüfung, wie gut die in den Experimenten 3, 4 und 5 gewonnenen Ergebnisse mit dem PCS-Modell vorhergesagt werden können. Unter Berücksichtigung der vorherigen Befunde wird in den Experimenten 6 und 7 erneut der Einfluss der Wiederholung aufgegriffen und untersucht.

3.2 METHODISCHE VORÜBERLEGUNGEN

Um das Verständnis des Konsistenzmaximierungsprozesses zu vertiefen, werden prädezisionale Informationsverzerrungen in probabilistischen Inferenzentscheidungen untersucht. Bevor die entsprechenden Experimente dargestellt werden, sollen einige methodische Vorüberlegungen angestellt werden, die das verwendete Forschungsparadigma begründen und das Verständnis der Vorgehensweise erleichtern. So wird zunächst dem Problem der Messung subjektiver Cue-Validitäten begegnet (Kapitel 3.2.1). Die mangelnde Konvergenz zwischen impliziten und expliziten Einschätzungen wurde vielfach diskutiert (z.B. Evans, Clibbens, Cattani, Harris & Dennis, 2003; Nisbett und Wilson, 1977). Vor diesem Hintergrund ist zu erörtern, welches Messverfahren zur Erfassung von Cue-Umwertungen ratsam ist und welche Aspekte hierbei zu beachten sind. Dann wird diskutiert, welche Auswertungsschritte bei der Erfassung prädezisionaler Umwertungen nötig sind, und das entsprechende Vorgehen erläutert (Kapitel 3.2.2).

3.2.1 Die Messung subjektiver Cue-Validitäten

Das PCS-Modell für probabilistische Inferenzentscheidungen von Glöckner und Betsch (2008b) stellt die Basis für die darzustellenden Experimente dar. Glöckner et al. (im Druck) schlagen ein Paradigma zur Erfassung der prädezisionalen Umwertungen vor, das in

ihren drei Experimenten in leicht modifizierter Form zum Einsatz kommt. Wesentliche Kennzeichen des Paradigmas sind (1) die explizite Abfrage der subjektiven Cue-Validität, (2) die Verwendung eines Pretest-Posttest-Designs und (3) die Verhinderung einer Entscheidung. Diese Grundbausteine werden auch in der vorliegenden Arbeit verwendet, weshalb ihre Notwendigkeit erläutert und ihre Umsetzung dargestellt werden sollen.

Das erste Kennzeichen, die explizite Abfrage der subjektiven Cue-Validität, erfordert die Erörterung des Konzeptes der Cue-Validität und die Begründung für eine explizite Erfassung derselben. *Validität* ist definiert als die bedingte Wahrscheinlichkeit, dass eine aus einer Referenzklasse R entnommene Option A eine andere Option B bezüglich eines bestimmten Kriteriums übertrifft, wenn ein Merkmal (Cue) ausschließlich bei Option A vorliegt (Gigerenzer, Hoffrage & Kleinbölting, 1991). Weiterhin unterscheiden Gigerenzer et al. in ihrer Theorie der Probabilistischen Mentalen Modelle die ökologische Validität, die der in der realen Umwelt anzutreffenden Validität entspricht, von dem Konzept der Cue-Validität, die der Einschätzung der Versuchsperson entspricht und die im Folgenden als *subjektive Cue-Validität* bezeichnet wird. Während Gigerenzer et al. eine hohe Übereinstimmung von ökologischen und subjektiven Cue-Validitäten postulieren, ist diese Annahme, wie Glöckner (2006) überzeugend darstellt, eher unwahrscheinlich. Um zu ermitteln, welche Vorhersagekraft und damit welche subjektive Bedeutung einem Cue beigemessen wird und wie sich diese Einschätzung eventuell verändert, ist folglich die Fokussierung auf subjektive Cue-Validitäten erforderlich. Aufgrund der Schwerpunktsetzung in dieser Arbeit wird auf einen Vergleich mit ökologischen Validitäten verzichtet. Glöckner et al. (im Druck) fragen sowohl nach der subjektiven Einschätzung objektiver Cue-Validitäten (Experimente 3) als auch im Sinne Brunswiks (1955) nach der Cue-Verwendung und damit nach dem Vertrauen in einzelne Cues (Experiment 1 und 2). Letzteres Vorgehen wird von den Autoren explizit bevorzugt. Um allerdings die Annahme zu entkräften, dass in probabilistischen Inferenzentscheidungen die Schlussfolgerung von Cues auf Optionen unidirektional ist (Gigerenzer et al., 1999), und folglich die Annahme bidirektionaler Prozesse zu stärken, haben sie in einem Experiment die subjektive Einschätzung objektiver Cue-Validitäten erfasst. In allen Experimenten konnten sie systematische Umwertungen zeigen, obgleich die Evidenz im dritten Experiment schwächer war. Um aufzuzeigen, dass entgegen der verbreiteten Annahme Cue-Validitäten nicht stabil sind, wird in den darzustellenden Experimenten die Erfassung der subjektiven Einschätzung der objektiven Cue-Validitäten bevorzugt. Hierbei gilt es zu beachten, das Validitätskonzept zu definieren

und gleichzeitig deutlich zu machen, dass *subjektive* Cue-Validitäten erfasst werden. So werden ein Verständnis des Konzeptes und eine valide Messung ermöglicht. Die explizite Messung wurde realisiert, indem die Versuchspersonen auf einer Skala, deren Endpunkte mit 50% und 100% bezeichnet waren, die Validität einzelner Cues einschätzten. Ersichtlich wird, dass in den Experimenten von Glöckner et al. (im Druck) sowie in den im Folgenden darzustellenden die prädezisionalen Verzerrungen als systematische Umwertung von subjektiven Cue-Validitäten verstanden werden. Folglich wird in dieser Arbeit der in der Arbeitsdefinition genannte Aspekt der *Reevaluation* (siehe Kapitel 2.5) in einem spezifischen Sinne verstanden. Nun stellt sich die konkrete Frage, *wie* subjektive Cue-Validitäten gemessen werden können. Unter Berücksichtigung der Befunde zur Übereinstimmung implizit und explizit erhobener Einschätzungen (z.B. Nisbett und Wilson, 1977) kommt Glöckner (2006) zu dem Schluss, dass eine explizite Messung bei der Erfassung von subjektiven Cue-Validitäten kritisch zu sehen ist. Er schlägt eine Methode vor, die jedoch nur eine ordinale Erfassung der impliziten Cue-Rangfolge erlaubt. Folglich ist sie – wie Glöckner selbst anmerkt – für eine präzise Erfassung der subjektiven Cue-Validitäten und damit für die vorliegenden Forschungsabsichten unzureichend. Geeignete andere Maße wurden bisher nicht entwickelt, weshalb in den Arbeiten zu prädezisionalen Informationsverzerrungen auf explizite Abfragen zurückgegriffen wird (z.B. Glöckner et al., im Druck; Russo et al., 1998; Simon, Krawczyk et al., 2004). Auch in dieser Arbeit werden die subjektiven Cue-Validitäten daher explizit abgefragt. Zudem sei auf die Aussage von Fazio und Olson (2003), dass hinsichtlich der Diskrepanz zwischen expliziten und impliziten Messungen der Aufgabentyp zu beachten ist, hingewiesen. Sie verweisen auf Studien, in denen eine hohe Übereinstimmung gefunden wurde, wenn es sich um "mundane, socially noncontroversial objects" (S. 304) handelte. Diese Bedingung ist in den vorzustellenden Experimenten erfüllt, so dass die explizite Angabe von subjektiven Cue-Validitäten als vertretbar erachtet wird. Ein Problem stellt diese explizite Abfrage aber dann dar, wenn sie erinnert wird und damit eventuell in einem zweiten Test verfälscht wiedergegeben wird. Derartige Effekte gilt es auszuschließen, so dass ein entsprechendes Experiment (Kapitel 3.3) durchgeführt wurde.

Die Messung der Veränderung subjektiver Cue-Validitäten in einem Pretest-Posttest-Design ist aus mehreren Gründen angemessen. So wird auf diesem Wege der Tatsache Rechnung getragen, dass die Einschätzung der Cue-Validitäten subjektiv ist und sich damit zwischen Personen deutlich unterscheiden kann. Zudem kann auf diesem Wege die

Umwertung in Abhängigkeit von der individuellen (tendenziellen oder finalen) Entscheidung betrachtet werden, so dass Nulleffekte durch Aggregation ausgeschlossen werden können. Obgleich die Betrachtung der Veränderung pro Person vorteilhaft ist, birgt das Design doch die Gefahr, dass sich die Bearbeitung des Pretests auf die Einschätzungen im Posttest auswirkt. Zwar sollen durch die Bearbeitung eines thematisch unverbundenen Experimentes nach dem Pretest Erinnerungseffekte ausgeschlossen werden, doch es kann streng genommen bisher nicht entschieden werden, ob die Veränderung der subjektiven Cue-Validitäten auf Pretesteffekte, Treatmenteffekte oder die Interaktion zwischen Pretest und Treatment zurückzuführen ist. Dieser wichtigen Frage wird im ersten Experiment (Kapitel 3.3) nachgegangen.

Die Notwendigkeit des dritten Kennzeichens, des Verhinderns einer Entscheidung, resultiert aus der vielfach diskutierten Frage, ob es sich bei der Umwertung von Informationen oder Alternativen um ein prä- oder ein postdezisionales Phänomen handelt. In Kapitel 2.4 konnte anhand einer Vielzahl von Befunden aufgezeigt werden, dass neben postdezisionalen Umwertungen auch prädezisionale Umwertungen stattfinden können. Der Fokus der vorliegenden Arbeit liegt auf dem Entscheidungsprozess und damit auf der prädezisionalen Phase. Folglich galt es bei der Konzeption der Experimente darauf zu achten, die Umwertungen in dieser Phase zu erfassen und eine endgültige Entscheidung zu verhindern. So wurden die Versuchspersonen instruiert, lediglich darüber zu reflektieren, für welche Option sie sich auf Basis der vorliegenden Informationen entscheiden würden. Um das Treffen einer endgültigen Entscheidung hinauszuzögern, wurde den Versuchspersonen eine weitere relevante Information in Aussicht gestellt. Dieses Vorgehen hat sich in früheren Studien bewährt (z.B. Glöckner et al., im Druck, Experiment 3; Simon, Snow et al., 2004, Experiment 3). Zudem gaben in den vorzustellenden Experimenten die Versuchspersonen – falls überhaupt – ihre Entscheidung erst dann an, wenn der Posttest bereits bearbeitet worden war. Gewiss kann nicht ausgeschlossen werden, dass die Versuchspersonen bereits vor der Bearbeitung des Posttests eine Entscheidung getroffen haben. Allerdings konnte diese nicht endgültig sein und war nicht bindend, so dass postdezisionale Dissonanzreduktionsprozesse nach Festinger (1964) ausgeschlossen sind (vgl. auch Glöckner et al., im Druck; Simon et al., 2001).

Der aus diesen Kennzeichen resultierende allgemeine Grundaufbau der Experimente ist in Anhang A-1 schematisch dargestellt.

3.2.2 Die Auswertung der Veränderung subjektiver Cue-Validitäten

Hinsichtlich der Erfassung der Veränderung subjektiver Cue-Validitäten als Ausdruck des Konsistenzstrebens ist zu unterscheiden, auf welcher Ebene diese Umwertungen erfasst werden. Auf der Ebene der Alternativen bzw. Optionen konnte bereits Brehm (1956) den sogenannten *spreading apart*-Effekt nachweisen: Die gewählte Alternative wird in ihrer Attraktivität aufgewertet, wohingegen die nicht gewählte Alternative in ihrer Attraktivität abgewertet wird. Während Brehm diesen Effekt für die postdezisionale Phase nachwies, konnte er auch für die prädezisionale Phase gezeigt werden. So berichten etwa Beckmann und Kuhl (1984) Ergebnisse, dass der Unterschied zwischen der Attraktivität zweier Alternativen zum zweiten Messzeitpunkt, zu dem die Versuchspersonen eine tendenzielle, aber keine endgültige Entscheidung trafen, als größer eingeschätzt wurde als zum ersten Messzeitpunkt. In den in Kapitel 2.4 dargestellten Experimenten wurde zudem die Umwertung einzelner Informationen bzw. Cues erfasst. Dass dieser Auswertungsschritt einen Mehrwert bedeutet, macht etwa Svenson (1996) deutlich: Er gibt zu bedenken, dass Umwertungen auf der Attribut-Ebene gefunden werden könnten, die auf der globalen Alternativen-Ebene nicht gefunden werden könnten. Außerdem erhalte man durch die Analyse auf der Ebene der Attribute Einblicke in mögliche Veränderungen im Alternativen-Profil.

Um nun auf der einen Seite den Einfluss eines oder weniger Cues auf die globale Einschätzung der Optionen und auf der anderen Seite die Veränderung jedes einzelnen Cues zu erfassen, werden sowohl die Ebene der Optionen als auch die Ebene der Cues betrachtet.

Bezüglich der Ebene der Optionen tritt zunächst die Schwierigkeit auf, dass – wie bereits in Kapitel 3.2.1 angesprochen – im vorliegenden Paradigma Validitätseinschätzungen einzelner Cues erhoben werden. Eine Einschätzung auf der Ebene von Optionen erfolgt nicht. Da aber Glöckner (2008a) zeigen konnte, dass die Entscheidungen von Versuchspersonen die von einer *weighted additive*-Strategie (Payne et al., 1988) vorhergesagten Entscheidungen approximieren, erscheinen die gewichteten Summen ein plausibles Maß für globale Einschätzungen zu sein. Unter zusätzlicher Berücksichtigung der Befunde von Glöckner und Betsch (2008c), dass Versuchspersonen bei ihren Entscheidungen bedenken, dass Cues mit einer Validität v von 50% keine Vorhersagekraft haben, und folglich die Cue-Validitäten korrigieren ($p_{korrigiert} = v_{cue} - 50$), werden die *korrigierten* gewichteten Summen berechnet und als Maß für die Optionen-Ebene verwendet. Hierbei werden die korrigierten

subjektiven Validitäten mit der Cue-Information, die bei Vorliegen des Cues +1 und bei Fehlen des Cues -1 beträgt, multipliziert und die jeweiligen Produkte aufsummiert:

$$WADD_{korrigiert_i}(t) = \sum_{j=1}^{k} p_{korrigiert_j} * c_j(i)$$

$WADD_{korrigiert_i}(t)$ korrigierte gewichtete Summe für Option *i* zum Zeitpunkt *t*

$p_{korrigiert_j}$ korrigierte subjektive Validität von Cue *j*

$c_j(i)$ Information des Cue *j* über Option *i*

k Anzahl der Cues

Für jede Option kann nun die korrigierte gewichtete Summe für den Pretest und für den Posttest ermittelt werden, indem die jeweiligen Berechnungen entweder auf den Validitätsangaben aus dem Pretest oder aus dem Posttest beruhen. Um nun zu ermitteln, ob auch bei probabilistischen Inferenzentscheidungen ein *spreading apart* der Optionen zu beobachten ist, werden die Differenzen *D* zwischen den korrigierten gewichteten Summen pro Zeitpunkt *t* gebildet, das heißt separat für den Pretest (t1) und für den Posttest (t2) nach der folgenden Formel:

$$D(t) = WADD_{korrigiert_i}(t) - WADD_{korrigiert_j}(t)$$

$WADD_{korrigiert_i}(t)$ korrigierte gewichtete Summe für Option *i* zum Zeitpunkt *t*

$WADD_{korrigiert_j}(t)$ korrigierte gewichtete Summe für Option *j* zum Zeitpunkt *t*

i = gewählte Option, wenn eine Entscheidung getroffen wurde

$WADD_{korrigiert_i}(t1) > WADD_{korrigiert_j}(t1)$, wenn keine Entscheidung getroffen wurde

Wenn in den Experimenten eine finale Entscheidung getroffen wird, basiert die Bildung der Differenz auf derselben. Wurde etwa Option B gewählt, so wird sowohl für die Berechnung der Differenz im Pretest als auch für diejenige im Posttest die korrigierte gewichtete Summe von Option A von der korrigierten gewichteten Summe von Option B subtrahiert. Wenn keine finale Entscheidung getroffen wurde, wird gemäß den Befunden von Glöckner (2008a) angenommen, dass die Personen die Option wählen, die im Pretest die

höhere korrigierte gewichtete Summe aufweist. Entsprechend wird für die Differenzbildung im Pretest und im Posttest die niedrigere korrigierte gewichtete Summe von der höheren subtrahiert. Da gemäß dem Konsistenzmaximierungsprozess zu erwarten ist, dass Cues, die für die später gewählte Option sprechen, in ihrer Validität aufgewertet werden, während Cues, die gegen die später gewählte Option sprechen, in ihrer Validität abgewertet werden (Glöckner et al., im Druck), sollte sich ein *spreading apart* der Optionen in einer Vergrößerung der Differenz zwischen den korrigierten gewichteten Optionen äußern.

Darüber hinaus werden Umwertungen auf der Ebene der Cues untersucht. Diese Analyse entspricht der Untersuchung von Validitäts-Umwertungen einzelner Cues bei probabilistischen Inferenzen, wie sie Glöckner et al. (im Druck) vorgenommen haben. Da die Richtung der Veränderung von der später getroffenen Entscheidung abhängt, werden im Folgenden die Cues entsprechend der angenommenen oder tatsächlichen Entscheidung in aufzuwertende und abzuwertende Cues eingeteilt und die Veränderung der Validitätseinschätzung vom Pretest zum Posttest analysiert. Cues, die für die angenommene oder tatsächliche Entscheidung sprechen, werden als aufzuwertend klassifiziert, während Cues, die gegen diese Entscheidung sprechen, als abzuwertend klassifiziert werden.

Die in der vorliegenden Arbeit vorgenommenen Analysen auf der Ebene der Optionen sowie auf der Ebene der Cues erlauben eine umfassende Untersuchung der Umwertung von Cue-Validitäten, die Ausdruck des Konsistenzmaximierungsprozesses ist.

Nachdem die Vorgehensweise bezüglich der Erfassung subjektiver Cue-Validitäten und die Auswertung möglicher Veränderungen erläutert worden ist, werden im Folgenden die empirischen Arbeiten vorgestellt.

3.3 Experiment 1: Der Einfluss des Pretests

Zielsetzung

In bisherigen Experimenten zur Veränderung von Cue-Validitäten in probabilistischen Inferenzentscheidungen (Glöckner et al., im Druck) wurde stets ein Pretest-Posttest-Design verwendet. Da der Einfluss des Pretests auf die Umwertung, wie in Kapitel 3.2.1 beschrieben, noch unklar ist, soll das vorliegende Experiment einen wichtigen Beitrag zur Klärung dieser Frage liefern. Die Zielsetzung dieses ersten Experiments war es folglich, zu

untersuchen, ob die Veränderung der Cue-Validitäten auf Pretesteffekte, Treatmenteffekte oder die Interaktion zwischen Pretest und Treatment zurückzuführen ist.

METHODE

Design. Um die Frage zu beantworten, ob Pretesteffekte, Treatmenteffekte oder die Interaktion zwischen Pretest und Treatment für die Umwertung subjektiver Cue-Validitäten verantwortlich sind, wurde das Untersuchungsschema des Solomon-Viergruppenplans (Solomon, 1949) verwendet (Abbildung 3).

	Pretest	Treatment	Posttest
Bedingung 1	x	x	x
Bedingung 2	x		x
Bedingung 3		x	x
Bedingung 4			x

Abbildung 3: Schema des Solomon-Viergruppenplans. Ein x zeigt die Bearbeitung des Experimentalteils an.

Versuchspersonen der ersten Bedingung bearbeiteten zunächst den Pretest (Einschätzung subjektiver Cue-Validitäten), erhielten daraufhin das Treatment (Szenario und anschließende Reflexion über die Entscheidungssituation) und bearbeiteten zuletzt den Posttest (Einschätzung subjektiver Cue-Validitäten analog zum Pretest). Versuchspersonen der zweiten Bedingung bearbeiteten lediglich den Pre- und den Posttest, erhielten aber kein Treatment. Versuchspersonen der dritten Bedingung erhielten das Treatment, ohne zuvor einen Pretest durchgeführt zu haben, und bearbeiteten anschließend den Posttest. Versuchspersonen der vierten Bedingung beantworteten lediglich die Fragen des Posttests.

Hypothesen. Gemäß dem von Glöckner und Betsch (2008b) postulierten PCS-Modell für probabilistische Inferenzentscheidungen und dem diesem zugrunde liegenden Prinzip der automatischen Konsistenzmaximierung wurde erwartet, dass die Veränderung subjektiver Cue-Validitäten auf das Treatment, nicht aber auf den Pretest bzw. die Interaktion zwischen Pretest und Treatment zurückzuführen ist.

Material. Die in dem vorliegenden Experiment verwendete Entscheidungssituation basiert auf der Arbeit von Glöckner et al. (im Druck, Experiment 2). Im Gegensatz zu dem Vorgehen der Autoren wurde nach subjektiven Validitäten und nicht nach der Cue-Verwendung gefragt. Des Weiteren wurde die Entscheidungssituation umformuliert, so dass sie eindeutig als probabilistische Inferenzaufgabe zu betrachten war. Sowohl im Pre- als auch im Posttest wurden die Versuchspersonen gebeten, die Validität der Wetterberichte

(Cues) von „www.wetteronline.de“, „Sat.1“, „Bild-Zeitung“ und „ZDF“ für das Wetter an einem Ferienort in sieben Tagen einzuschätzen. Als Treatment wurde den Versuchspersonen eine Matrix, in deren Zeilen die Cues (Wetterberichte) und in deren Spalten die Optionen (Ferienort A, Ferienort B) abgetragen waren, dargeboten. Diese Matrix gab Auskunft darüber, ob die jeweiligen Wetterberichte für zwei Ferienorte A und B sonniges Wetter vorhersagen (+) oder nicht (-). Ob die Versuchspersonen einen Pretest bearbeiteten bzw. ein Treatment erhielten, hing von der jeweiligen Versuchsbedingung ab. Den Posttest sowie die anschließenden demografischen Fragen beantworteten hingegen alle Versuchspersonen. Die vollständigen Instruktionen sind in Anhang A-2 zu finden.

Stichprobe. An dem Experiment nahmen insgesamt 199 Studenten (138 weiblich, 61 männlich; Durchschnittsalter 22.53 Jahre, *SD* = 3.62) teil. Die Daten wurden in drei Erhebungen an den Universitäten Erfurt (47 und 83 Versuchspersonen) und Bonn (69 Versuchspersonen) erhoben. Das Experiment wurde in *Visual Basic* programmiert und war stets Teil einer Experimentalbatterie. Diese Batterien dauerten etwa 60-80 Minuten (Erhebung 1, Erfurt), 50 Minuten (Erhebung 2, Erfurt) und 90 Minuten (Erhebung 3, Bonn), wobei circa 5-10 Minuten auf das hier zu berichtende Experiment entfielen. Die Versuchspersonen wurden über das Online-Rekrutierungssystem *Orsee* (Greiner, 2004) angeworben und erhielten als Aufwandsentschädigung durchschnittlich 11 € (Erhebung 1), 6 € (Erhebung 2) bzw. 13 € (Erhebung 3). Die Versuchspersonen wurden den Bedingungen zufällig, aber mit dem Ziel einer Gleichverteilung zugeordnet. So wurden 52 Versuchspersonen der ersten Bedingung, 45 Versuchspersonen der zweiten Bedingung, 52 Versuchspersonen der dritten Bedingung und 50 Versuchspersonen der vierten Bedingung zugeteilt.

Durchführung. Für die Versuchspersonen der ersten beiden Bedingungen begann das Experiment mit dem Pretest. Nachdem ihnen das Konzept der subjektiven Cue-Validität anhand eines Beispiels erklärt worden war, schätzten sie die Validität (Vorhersagekraft) von Wetterberichten (Cues) verschiedener Quellen für das Wetter an einem Ferienort in sieben Tagen ein. Diese Einschätzung nahmen sie anhand eines Schiebereglers auf einer Skala von 50% bis 100% vor. Es folgte ein thematisch unverbundenes Experiment (etwa 30 Minuten), das in allen drei Erhebungen identisch war. Die Versuchspersonen der ersten Bedingung erhielten daraufhin das Treatment: Sie wurden gebeten sich vorzustellen, dass sie in den Urlaub fahren möchten und bei der Wahl zwischen zwei Ferienorten A und B an den Ort reisen möchten, an dem sie eher sonniges Wetter erwarten. Aus diesem Grund bezögen sie Informationen aus Wetterberichten verschiedener Quellen, die ihnen vorgestellt wurden. Des

Weiteren wurden die Versuchspersonen instruiert, noch keine endgültige Entscheidung zu treffen, da sie noch die Vorhersage eines weiteren Wetterberichtes (ARD) abwarten möchten. Die Vorhersagen der Wetterberichte für das Wetter an den Ferienorten A und B wurden anhand einer Matrix dargeboten. Während nur „Sat.1“ für Ferienort A sonniges Wetter vorhersagte, für Ort B aber nicht, war bei den Wetterberichten von „www.wetteronline.de“, „Bild-Zeitung“ und „ZDF“ das Gegenteil der Fall. Die Versuchspersonen wurden nun gebeten zu überlegen, an welchem Ferienort sie aufgrund der vorliegenden Informationen tendenziell eher sonniges Wetter erwarteten. Im Anschluss schätzten sie in dem analog zum Pretest aufgebauten Posttest erneut die subjektiven Validitäten der Wetterberichte auf einer Skala von 50% bis 100% ein. Die Versuchspersonen der zweiten Bedingung bearbeiteten ebenfalls nach dem Pretest das thematisch unverbundene Experiment, erhielten im Anschluss daran aber kein Treatment. Stattdessen beantworteten sie direkt die Fragen des Posttests. Die Versuchspersonen der dritten Bedingung führten zunächst das thematisch unverbundene Experiment durch, erhielten dann das Treatment und bearbeiteten schließlich den Posttest. Da ihnen das Konzept der Validität noch nicht erklärt worden war, erfolgte dies im Posttest. Die Versuchspersonen der vierten Bedingung führten ebenfalls zunächst das thematisch unverbundene Experiment durch, erhielten dann aber kein Treatment, sondern bearbeiteten direkt den Posttest. Auch ihnen wurde zuvor das Konzept der Validität erklärt. Im Anschluss an den Posttest beantworteten alle Versuchspersonen demografische Fragen.

ERGEBNISSE

Vorausgehende Analysen. Drei Versuchspersonen haben sowohl im Pretest als auch im Posttest durchgängig die Cue-Validitäten auf 50% geschätzt, was auf ein mangelndes Verständnis des dahinterstehenden Konzeptes oder der verwendeten Skala schließen lässt.[10] Diese Versuchspersonen wurden aus den folgenden Analysen ausgeschlossen. Von den verbleibenden 196 Versuchspersonen (137 weiblich, 59 männlich; Durchschnittsalter: 22.52 Jahre, $SD = 3.61$) hatten 47 an der ersten Erhebung und 80 an der zweiten Erhebung in Erfurt teilgenommen. Die übrigen 69 Personen hatten an der dritten Erhebung in Bonn teilgenom-

[10] Da nur Versuchspersonen ausgeschlossen wurden, die die Validitäten stets auf 50% oder 100% geschätzt haben, wird von einem *weichen* Kriterium gesprochen. Es können auch Versuchspersonen ausgeschlossen werden, die entweder im Pretest oder im Posttest durchgängig 50% oder 100% geschätzt haben. Dies war bei 5 Versuchspersonen der Fall. Entsprechend wird der Ausschluss dann unter einem *harten* Kriterium vorgenommen.

men. Damit verblieben 50 Personen in der ersten Bedingung, 44 in der zweiten Bedingung, 52 in der dritten Bedingung und 50 in der vierten Bedingung.

Wie in Kapitel 3.2.2 beschrieben, wurden die Differenzen zwischen den korrigierten gewichteten Summen der Optionen zum Zeitpunkt des Pretests und des Posttests ermittelt: Die korrigierte gewichtete Summe der Option, die im Pretest niedriger war, wurde von der korrigierten gewichteten Summe der Option, die im Pretest höher war, subtrahiert. Diese Berechnung erfolgte separat für den Pretest (Differenz_pre) und für den Posttest (Differenz_post). Eine explorative Analyse mittels Boxplots ergab keine extremen Ausreißer.[11] Deskriptive Daten zu den Validitätseinschätzungen in den einzelnen Bedingungen sind in Tabelle B-1 des Anhangs B aufgeführt.

Unterschiede zwischen den Erhebungen. In einem ersten Schritt wurde überprüft, ob die Durchführung des Experimentes zu verschiedenen Zeitpunkten und an verschiedenen Universitäten einen Einfluss auf die Ergebnisse hatte. Hierzu wurden zwei einfaktorielle ANOVAs mit „Erhebung“ als *between subjects*-Faktor und Differenz_pre bzw. Differenz_post als abhängiger Variable berechnet. Der Haupteffekt „Erhebung“ wurde weder bezüglich der Differenz zwischen den korrigierten gewichteten Summen im Pretest noch bezüglich der Differenz zwischen den korrigierten gewichteten Summen im Posttest signifikant, $F(2, 91) < 1$, $p = .91$, $\eta^2 = .00$ bzw. $F(2, 193) < 1$, $p = .46$, $\eta^2 = .01$.[12]

Der Einfluss des Pretests. Gemäß dem Vorschlag von Campbell und Stanley (1966) wurde eine zweifaktorielle Varianzanalyse mit den zweistufigen Faktoren „Pretest“ (ja vs. nein) und „Treatment“ (ja vs. nein) als *between subjects*-Faktoren durchgeführt. Die abhängige Variable war Differenz_post, d.h. die Differenz zwischen den korrigierten gewichteten Summen der Optionen im Posttest. Eine post-hoc durchgeführte Teststärkenanalyse für den Haupteffekt „Treatment“, den Haupteffekt „Pretest“ und die Interaktion „Treatment x Pretest“ mittels GPower (Faul, Erdfelder, Lang & Buchner, 2007)[13] ergab, dass das Anlegen des konventionellen α-Niveaus von .05 bei $N = 196$ das Entdecken mittlerer Effekte ($f = .25$, Cohen, 1988) mit einer Teststärke von .94 erlaubt. Es zeigten sich nicht signifikante Ergebnisse bezüglich des Treatments, $F(1, 192) < 1$, $p = .47$, $\eta_p^2 = .00$[14], des Pretests, $F(1, 192) = 2.35$, $p = .13$, $\eta_p^2 = .01$, sowie bezüglich der Interaktion Treatment x Pretest, $F(1, 192) < 1$, $p = .57$, $\eta_p^2 = .00$. Der nicht signifikante Haupteffekt „Pretest“ besagt,

[11] Dies gilt auch unter dem harten Kriterium.
[12] Unter Anlegen des harten Kriteriums zeigen sich ebenfalls keine signifikanten Ergebnisse.
[13] Sämtliche Teststärkenanalysen wurden mit der Version 3.1.2 durchgeführt.
[14] Die Abkürzung η_p^2 steht hier wie im Folgenden für das partielle Eta-Quadrat.

dass sich die Gruppen in Abhängigkeit davon, ob sie einen Pretest bearbeitet haben oder nicht, nicht bedeutend unterscheiden. Dass die Interaktion nicht signifikant wurde, lässt laut Walton Braver und Braver (1988) darauf schließen, dass keine Sensitivierung durch den Pretest vorliegt. Dass der Haupteffekt „Treatment“ nicht signifikant wurde, lässt laut Autoren aber nicht die Schlussfolgerung zu, dass das Treatment keinen Einfluss hat, da die Informationen des Pretests in den Bedingungen 1 und 2 nicht berücksichtigt wurden. Walton Braver und Braver schlagen daher eine separate Analyse für die ersten beiden Gruppen vor, und zwar eine Kovarianzanalyse, einen *t*-Test für unabhängige Stichproben oder eine Varianzanalyse mit Messwiederholung. Ein signifikantes Ergebnis in einem der Tests lässt laut Walton Braver und Braver (1988) den Schluss zu, dass das Treatment einen Effekt hat, der Pretest hingegen keinen Einfluss hat. Es wurde eine Kovarianzanalyse mit „Treatment“ als *between subjects*-Faktor, Differenz_post als abhängiger Variable und Differenz_pre als Kovariate berechnet. Wie die Autoren selbst anmerken ist aufgrund des Ignorierens der Gruppen 3 und 4 die Teststärke relativ gering. Eine Berechnung mittels GPower (Faul et al., 2007) zeigt etwa, dass unter den vorliegenden Voraussetzungen ein mittlerer Effekt der Größe $f = .25$ (Cohen, 1988) unter Anlegen eines Signifikanzniveaus von $\alpha = .05$ nur mit einer Teststärke von 67% gefunden werden kann. Durch Erhöhen des Niveaus auf $\alpha = .10$ kann diese auf 78% gesteigert werden, obgleich auch dieser Wert noch nicht zufriedenstellend ist (vgl. Rasch, Friese, Hofmann & Naumann, 2006a). Dennoch konnte der interessierende Effekt gefunden werden. Neben einem signifikanten Einfluss der Kovariate, $F(1, 91) = 71.73$, $p < .001$, $\eta_p^2 = .44$, wurde nämlich der interessierende Haupteffekt „Treatment“ signifikant, $F(1, 91) = 3.86$, $p < .10$, $\eta_p^2 = .04$. Gemäß Walton Braver und Braver (1988) kann der Einfluss des Pretests ausgeschlossen, der des Treatments hingegen angenommen werden: “If (…) the test is significant, evidence of treatment effects unqualified by pretest sensitization is obtained, and no further testing is necessary“ (S. 151).[15]

Die Umwertung subjektiver Cue-Validitäten. Nachdem gezeigt worden ist, dass kein Einfluss des Pretests, sondern lediglich des Treatments vorliegt, stellt sich natürlich die Frage, ob dieser Einfluss wirklich in einer systematischen Umwertung subjektiver Cue-Validitäten besteht. Da die Experimentalanordnung in Bedingung 1 dem bisher verwendeten Design zur Untersuchung der Veränderung von Cue-Validitäten (Glöckner et al., im Druck) weitestgehend entspricht, konnte überprüft werden, ob sich die Veränderung von Cue-

[15] Berechnet man sämtliche Analysen unter Anlegen des harten Kriteriums, so zeigt sich stets ein ähnliches Ergebnismuster.

Validitäten mit dem vorliegenden modifizierten Paradigma replizieren ließ. Wie in Kapitel 3.2.2 beschrieben, wurde die Auswertung auf der Ebene der Optionen und auf der Ebene der Cues vorgenommen. Um zu ermitteln, ob ein *spreading apart* der Optionen aufgetreten ist, wurde eine messwiederholte ANOVA mit „Messzeitpunkt“ (Pretest und Posttest) als *within subjects*-Faktor berechnet. Wenn $\alpha = .05$ gesetzt wird, kann bei $n = 50$ ein mittlerer Effekt ($f = .25$, Cohen, 1988) mit einer Teststärke von .91 gefunden werden (Faul et al., 2007). Auf der Ebene der Optionen zeigte sich ein signifikantes *spreading apart*, $F(1, 49) = 5.94$, $p < .05$, $\eta_p^2 = .11$.[16] Auch für die Analyse auf der Ebene der Cues wurde zunächst überprüft, ob es zu extremen Validitätseinschätzungen kam. Eine Ausreißer-Analyse anhand von Boxplots pro Bedingung zeigte keinen extremen Ausreißer.[17] Anschließend wurde eine messwiederholte ANOVA mit den Innersubjektfaktoren „Messzeitpunkt“ und „Cue“ berechnet, wobei die Interaktion „Messzeitpunkt x Cue“ von Interesse war. Eine diesbezügliche post-hoc durchgeführte Teststärkenanalyse mittels GPower (Faul et al., 2007) ergab, dass ein mittlerer Effekt ($f = .25$, Cohen, 1988) bei $\alpha = .05$ und $n = 50$ mit einer Teststärke von .96 gefunden werden kann. Auf der Ebene der Cues zeigten sich deskriptiv systematische Umwertungen für jeden Cue (siehe Abbildung 4).

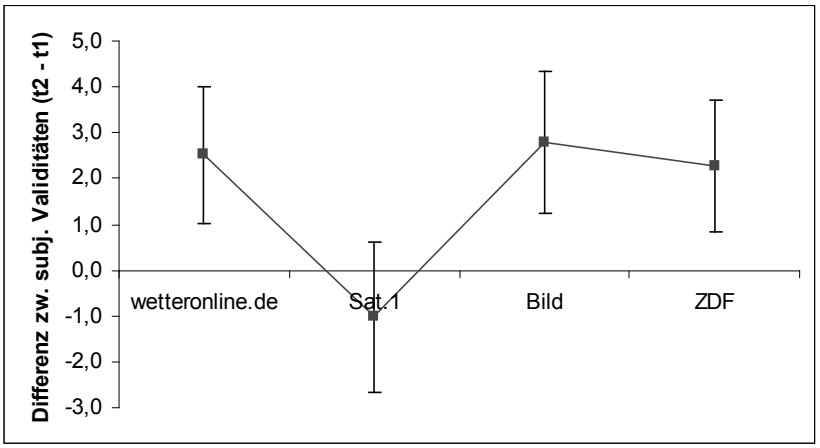


Abbildung 4: Differenz der subjektiven Cue-Validitäten zwischen den beiden Messzeitpunkten t1 (Pretest) und t2 (Posttest). Es wurde die Validitätseinschätzung des Pretests von derjenigen des Posttests subtrahiert. Werte > 0 kennzeichnen folglich eine Aufwertung, Werte < 0 eine Abwertung. Die Fehlerbalken geben zu beiden Seiten jeweils den Standardfehler der Differenzwerte an.

[16] Unter Anlegen des harten Kriteriums zeigt sich ebenfalls ein signifikantes Ergebnis.
[17] Unter Anlegen des harten Kriteriums gab es ebenfalls keine extremen Ausreißer.

Jedoch wurde die Interaktion „Messzeitpunkt x Cue“ knapp nicht signifikant, $F(2.45, 119.95) = 2.60$, $p = .07$, $\eta_p^2 = .05$.[18] Geplante Vergleiche mit gepaarten *t*-Tests (einseitig) zeigten, dass zwei der vier Cues signifikant umgewertet wurden, und zwar „www.wetteronline.de“, $t(49) = 1.70$, $p < .05$, $d_z = .24$[19], und „Bild“, $t(49) = 1.80$, $p < .05$, $d_z = .25$.[20]

Weitere Ergebnisse. Es konnte gezeigt werden, dass die Cues systematisch umgewertet werden, obgleich dies lediglich zwei von vier Cues betraf. Daher könnte der Einwand erhoben werden, dass lediglich Cues, deren Einschätzung der Validität unreliabel ist, einer Veränderung unterliegen.[21] Dieses Experiment erlaubt eine Überprüfung dieser Vermutung. Die entsprechenden Cues wurden nämlich auch von Versuchspersonen der zweiten Bedingung eingeschätzt, mit dem Unterschied, dass in dieser Bedingung kein Treatment erfolgt war. Die Höhe der Korrelationen zwischen der Pretest-Einschätzung und der Posttest-Einschätzung kann folglich Hinweise auf die Reliabilität der Cue-Einschätzungen geben. Bestehen deutliche Unterschiede zwischen den Korrelationskoeffizienten der umgewerteten und der nicht umgewerteten Cues, deutet dies auf die Relevanz der Reliabilität hin. Es ergaben sich Korrelationen von .80 („www.wetteronline.de“), .88 („Sat.1“), .91 („Bild“) und .94 („ZDF“). Bereits die deskriptive Betrachtung der Korrelationskoeffizienten verdeutlicht, dass die signifikant umgewerteten Cues nicht weniger reliable Validitätseinschätzungen aufweisen. Einschränkend ist darauf hinzuweisen, dass die Schlussfolgerung, dass in diesem Experiment die Reliabilität der Cues nicht für das Auftreten von Umwertungen ausschlaggebend ist, nur indirekter Art ist, da die entsprechenden Ergebnisse an zwei unterschiedlichen Stichproben gewonnen wurden. Es ist jedoch gerade aufgrund der randomisierten Zuweisung nicht plausibel anzunehmen, dass für die Versuchspersonen der ersten Bedingung die Validitätseinschätzungen weniger reliabel waren.

DISKUSSION

Im vorliegenden Experiment wurde untersucht, ob die systematische Veränderung von Cue-Validitäten auf Pretesteffekte, Treatmenteffekte oder die Interaktion zwischen

[18] Da die Sphärizitätsannahme verletzt war, wurden die Freiheitsgrade korrigiert. Der Empfehlung von Rasch et al. (2006b), bei $\varepsilon > .75$ die Korrektur nach Huynh-Feldt vorzunehmen, wurde hier wie auch im Folgenden gefolgt. Bei der Berechnung der Teststärke wurde die Verletzung der Sphärizitätsannahme berücksichtigt.
[19] Die Berechnung der Effektgröße für einen abhängigen *t*-Test erfolgte hier wie auch im Folgenden mittels GPower (Faul et al., 2007). d_z ist das entsprechende Maß für abhängige Stichproben.
[20] Unter Anlegen des harten Kriteriums gehen die Ergebnisse zwar in die gleiche Richtung, sind aber nicht mehr signifikant.
[21] Ich danke Frank Renkewitz für diesen Hinweis.

Pretest und Treatment zurückzuführen ist. Die Ergebnisse zeigten, dass die Umwertung lediglich durch das Treatment zu begründen ist, während ein Einfluss des Pretests sowie eine Sensitivierung durch den Pretest nicht gefunden werden konnten. Dieses Ergebnis ist aus zwei Gründen von zentraler Bedeutung für die vorzustellenden Forschungsarbeiten. Der erste Grund betrifft das Design der Experimente: So rechtfertigen die Ergebnisse das im Folgenden verwendete Pretest-Posttest-Untersuchungsparadigma, da Urteile über Veränderungen vom Pretest zum Posttest möglich sind, ohne dass Pretesteffekte einschränkend berücksichtigt werden müssten. Der zweite Grund betrifft die theoretische Basis des Forschungsprogramms: Die Resultate stützen die aus dem PCS-Modell abgeleitete Hypothese, dass systematische Veränderungen von Cue-Validitäten auf die Reflexion über eine Entscheidungssituation und den damit einhergehenden Versuch, eine konsistente Interpretation bzw. eine gute Gestalt zu konstruieren, zurückzuführen sind. Damit stellen sie eine wichtige Unterstützung des PCS-Ansatzes dar. Die Evidenz für systematische Veränderungen selbst war zwar auf der Ebene der Optionen überzeugend, auf der Ebene der Cues jedoch nicht. Es wurden lediglich zwei von vier Cues signifikant umgewertet, wobei unter dem harten Ausschlusskriterium sogar kein Cue signifikant reevaluiert wurde. Die Tendenz geht zwar stets in die vorhergesagte Richtung, allerdings nicht (immer) signifikant. Dass einige Cues signifikant umgewertet werden und andere nicht, scheint allerdings nicht auf die Reliabilität der Cues zurückzuführen zu sein, auch wenn der entsprechende Nachweis nur indirekter Art ist. An diese Ergebnisse und ihre Interpretation schließt sich die Frage an, welche Faktoren einen Einfluss auf die Stärke dieses Prozesses haben. Dies stellt die zentrale Frage der Forschungsarbeiten dar, die in den folgenden Kapiteln vorgestellt werden.

3.4 Experiment 2: Der Einfluss der Wiederholung bei einer ausgeglichenen Informationslage

Zielsetzung

Das vorrangige Ziel des vorzustellenden Experimentes bestand darin, die zur Untersuchung der systematischen Umwertung subjektiver Cue-Validitäten verwendete Entscheidungssituation zu modifizieren, um sowohl die Anwendung auf eine neue Entscheidungssituation zu testen als auch das Verständnis des Konsistenzmaximierungsprozesses zu erweitern. Dass Informationsverzerrungen bei einer ausgeglichenen Informationslage auftreten, wurde bisher zwar für verschiedene Entscheidungsaufgaben nachgewiesen

(z.B. Simon, Krawczyk et al., 2004; Simon, Snow et al., 2004). In den von Glöckner et al. (im Druck) berichteten Experimenten zur prädezisionalen Informationsumwertung in probabilistischen Inferenzentscheidungen wurde jedoch stets eine unausgeglichene Informationslage verwendet. Folglich sollte erstens spezifisch für derartige probabilistische Inferenzaufgaben überprüft werden, ob eine systematische Umwertung von subjektiven Validitäten auch bei einer ausgeglichenen Informationslage auftritt. Des Weiteren wurde in bisherigen Experimenten zur prädezisionalen Informationsumwertung der Einfluss der Wiederholung einer Entscheidungssituation nicht untersucht. Hoeffler und Ariely (1999) konnten hingegen für den Bereich der Präferenzentscheidungen bereits zeigen, dass die Wiederholung der Entscheidungssituation zu einer Stabilisierung der Präferenz führt. Aus diesem Grund sollte zweitens untersucht werden, inwiefern die wiederholte im Vergleich zur einmaligen Darbietung einer Entscheidungssituation eine Auswirkung auf die Größe der Veränderung subjektiver Cue-Validitäten hat.

METHODE

Design. Die Veränderung subjektiver Cue-Validitäten wurde *within subjects* in einem Pre-/Post-Design gemessen. *Between subjects* wurde hierbei manipuliert, ob den Versuchspersonen im Treatment eine Entscheidungssituation oder – in Anlehnung an Hoeffler und Ariely (1999) – zwölf Entscheidungssituationen präsentiert wurden.

Hypothesen. Es wurde entsprechend dem im PCS-Modell von Glöckner und Betsch (2008b) postulierten Prozess der automatischen Konsistenzmaximierung und dessen Charakterisierung als Kernmechanismus jedes Entscheidungsprozesses erwartet, dass in der ersten Bedingung systematische Umwertungen auftreten. Genauer gesagt wurde davon ausgegangen, dass (a) im Sinne eines *spreading aparts* die Differenz zwischen den korrigierten gewichteten Summen der beiden Optionen vom Pre- zum Posttest zunimmt und dass (b) der Cue, der für die zu erwartende Entscheidung spricht, aufgewertet wird, während der Cue, der gegen diese Entscheidung spricht, abgewertet wird. Im PCS-Modell wird der Konsistenzmaximierungsprozess für *eine* Entscheidungssituation beschrieben, weshalb keine eindeutige Hypothese bezüglich des Einflusses der Wiederholung abgeleitet werden kann. Es lassen sich aber Plausibilitätsüberlegungen anstellen: So kann angenommen werden, dass die wiederholte Darbietung einer *identisch konzipierten* Entscheidungssituation als wiederholtes Wirken des Konsistenzmaximierungsprozesses verstanden werden kann und damit ebenfalls zu systematischen Umwertungen führen sollte. Explorativ wurde untersucht, ob die

wiederholte Darbietung von Entscheidungssituationen einen Einfluss auf die Stärke der systematischen Veränderung hat.

Material. Die Versuchspersonen wurden sowohl im Pre- als auch im Posttest instruiert, die Validität der beiden Cues „Universität“ und „Flughafen“ für die Einwohnerzahl realer brasilianischer Städte mit mehr als 100000 Einwohnern auf jeweils einem Schieberegler, der von 50% bis 100% reichte, einzuschätzen. Im Rahmen des Treatments wurde den Versuchspersonen als Coverstory mitgeteilt, dass sie eine Reise nach Brasilien planen. Während die Versuchspersonen der ersten Bedingung gebeten wurden sich vorzustellen, dass sie sich nur noch entscheiden müssten, in welche von zwei brasilianischen Städten sie fahren möchten, wurde den Versuchspersonen der zweiten Bedingung gesagt, sie wollten eine Tour durch zwölf brasilianische Städte planen. Hierbei stünden pro Zielstadt zwei Städte zur Auswahl, so dass sich die Versuchspersonen zwölf Mal zwischen zwei Städten entscheiden müssten. Die Matrix bzw. die Matrizen, anhand der bzw. anhand derer die Versuchspersonen zwischen den Städten entscheiden sollten, enthielten Informationen darüber, ob zwei brasilianische Städte eine Universität oder einen Flughafen haben (+) oder nicht haben (-). Hierbei wurden die Städte so gepaart, dass eine Stadt eine Universität, aber keinen Flughafen hat und umgekehrt. Das Material und die Instruktionen sind Anhang A-3 zu entnehmen.

Stichprobe. An dem Experiment nahmen insgesamt 121 Studenten der Universität Erfurt (108 weiblich, 13 männlich; Durchschnittsalter: 21.06 Jahre, SD = 2.64) teil. Sie wurden über Listen und E-Mails rekrutiert und erhielten als Aufwandsentschädigung entweder 6 € oder eine Versuchspersonenstunde. Die Teilnehmer wurden den Bedingungen zufällig, aber mit dem Ziel einer Gleichverteilung zugeordnet. So wurden 61 Versuchspersonen der ersten und 60 Versuchspersonen der zweiten Bedingung zugeteilt. Das Experiment wurde im Rahmen einer Experimentalbatterie, die circa 50 bis 60 Minuten dauerte, durchgeführt. Auf das hier zu berichtende Experiment entfielen etwa 20 bis 30 Minuten. Das Experiment wurde mit der Software *z-Tree* (Fischbacher, 2007) programmiert.

Durchführung. Nach einer anschaulichen Erläuterung des Validitätskonzeptes wurden die Versuchspersonen im Pretest instruiert, möglichst gut und differenziert die subjektive Validität der Merkmale „Universität“ und „Flughafen“ für die Einwohnerzahl brasilianischer Städte mit mehr als 100000 Einwohnern einzuschätzen. Es folgte ein thematisch unverbundenes Experiment (ca. 18 Min). Im zweiten Teil des Experimentes erfolgte das Treatment: Den Versuchspersonen wurde das Szenario „Reise nach Brasilien“,

das an die jeweilige Versuchsbedingung, d.h. „einmalige Präsentation der Entscheidungssituation“ bzw. „wiederholte Präsentation der Entscheidungssituation“, angepasst wurde, präsentiert. So sollten sich die Versuchspersonen der ersten Bedingung vorstellen, sie wollten gemeinsam mit einem Freund in den Urlaub fahren. Als Reiseziel hätten sie sich bereits für Brasilien entschieden. Nun stünden sie vor der Wahl zwischen zwei Städten, für die ein identisches Sonderangebot vorläge. In der Hoffnung auf bessere Freizeitangebote wollten sie gerne in die größere der beiden Städte fahren. Sie besäßen diese spezielle Information aber nicht, sondern könnten aus der ihnen zur Verfügung stehenden Landkarte lediglich schließen, ob die jeweilige Stadt eine Universität oder einen Flughafen hat. Nach einer Erläuterung der Matrix anhand eines Beispiels wurden die Versuchspersonen instruiert, sich zu überlegen, welche der nachfolgend dargebotenen Städte sie aufgrund der vorliegenden Informationen tendenziell für größer halten würden. Sie wurden explizit darauf hingewiesen, noch keine Entscheidung zu treffen. Für die Versuchspersonen der Bedingung „wiederholte Präsentation der Entscheidungsmatrix“ wurde das Szenario lediglich dahingehend abgeändert, dass sich die Versuchspersonen nun vorstellen sollten, eine Tour durch zwölf brasilianische Städte zu planen. Die jeweiligen realen Städtepaare wurden in einer *fixed random*-Reihenfolge dargeboten. Im Anschluss sollten die Versuchspersonen in beiden Bedingungen im Posttest erneut die Validität der Cues „Universität“ und „Flughafen“ für die Einwohnerzahl von brasilianischen Städten mit mehr als 100000 Einwohnern einschätzen. Abschließend wurden demografische Angaben erhoben sowie die vorherige Kenntnis der brasilianischen Städte abgefragt.

ERGEBNISSE

Vorausgehende Analysen. Da vier Versuchspersonen nach dem Pretest nicht das thematisch unverbundene Experiment bearbeiteten und folglich Erinnerungseffekte nicht ausgeschlossen werden können, wurden sie aus den nachfolgenden Analysen ausgeschlossen. In diesem Experiment schätzte keine Versuchsperson in einem oder in beiden Tests die Validität aller Cues mit 50% oder 100% ein, so dass wegen mangelnden Konzept- oder Skalenverständnisses keine weiteren Versuchspersonen ausgeschlossen werden mussten. Von den verbleibenden 117 Versuchspersonen (105 weiblich, 12 männlich; Durchschnittsalter: 21.03 Jahre, SD = 2.63) waren 60 Teilnehmer zufällig der ersten und 57 Teilnehmer zufällig der zweiten Bedingung zugeteilt worden. Die Validitätseinschätzungen der Cues pro Bedingung und pro Messzeitpunkt sind in Tabelle B-2 des Anhangs B aufgeführt.

Umwertung subjektiver Cue-Validitäten. Ob die subjektiven Cue-Validitäten systematisch umgewertet wurden, wurde sowohl auf der Ebene der Optionen als auch auf der Ebene der Cues untersucht. Für die Analyse auf der Ebene der Optionen wurde wie in Kapitel 3.2.2 beschrieben vorgegangen: Es wurden die korrigierten gewichteten Summen für jede Option berechnet, und zwar für den Pretest und für den Posttest. Anschließend wurde die Differenz zwischen den korrigierten gewichteten Summen im Pretest und im Posttest berechnet. Hierbei war die Höhe der korrigierten gewichteten Summen im Pretest ausschlaggebend. So wurde für beide Messzeitpunkte die korrigierte gewichtete Summe der Option, die im Pretest niedriger war, von der korrigierten gewichteten Summe der Option, die im Pretest höher war, subtrahiert. Eine pro Bedingung berechnete explorative Datenanalyse mittels Boxplots zeigte einen extremen Ausreißer in der zweiten Bedingung. Versuchspersonen, die vom Pre- zum Posttest ihre Hierarchie wechselten und bei denen die Differenz zwischen den korrigierten gewichteten Summen im Posttest folglich negativ war, werden in Anlehnung an Holyoak und Simon (1999) als *Switcher* bezeichnet. Im vorliegenden Datensatz wurden 25 Switcher identifiziert, was 21.37% aller Versuchspersonen entspricht. Da keine plausible Erklärung dafür gefunden werden kann, dass Versuchspersonen die Rangfolge der Cue-Validitäten wechselten, wurden diese Versuchspersonen ebenfalls aus den nachfolgenden Analysen ausgeschlossen. Allerdings betrifft dies eine hohe Anzahl der Versuchspersonen, weshalb in der Diskussion gesondert auf diesen Aspekt eingegangen wird. Zur Untersuchung der Veränderung der Differenzen zwischen den korrigierten gewichteten Summen wurde eine messwiederholte ANOVA mit „Messzeitpunkt“ (Pretest und Posttest) als *within subjects*-Faktor und „Bedingung“ (einmalige und wiederholte Darbietung) als *between subjects*-Faktor berechnet. Post-hoc durchgeführte Teststärkenanalysen mit GPower (Faul et al., 2007) führten zu dem Ergebnis, dass das Anlegen eines α-Niveaus von .05 bei $N = 92$ das Entdecken eines mittleren Effektes ($f = .25$, Cohen, 1988) mit Teststärken von .76 (Haupteffekt „Bedingung“) und von > .99 (Haupteffekt „Messzeitpunkt“, Interaktion „Messzeitpunkt x Bedingung“) erlaubt. Signifikant wurden weder der Haupteffekt „Messzeitpunkt“, $F(1, 90) < 1$, $p = .82$, $\eta_p^2 = .00$, noch der Haupteffekt „Bedingung“, $F(1, 90) < 1$, $p = .36$, $\eta_p^2 = .01$, noch die Interaktion „Messzeitpunkt x Bedingung“, $F(1, 90) = 1.07$, $p = .30$, $\eta_p^2 = .01$.[22] Folglich konnte die Hypothese,

[22] Das Ergebnismuster bleibt bestehen, wenn der extreme Ausreißer ausgeschlossen wird.

nach der in beiden Bedingungen systematische Umwertungen auftreten sollten, nicht bestätigt werden. Zudem konnte kein Einfluss der Wiederholung gefunden werden.[23]

Die Auswertung auf der Ebene der Cues erfolgte, indem zunächst die Cues gemäß der Hierarchie der korrigierten gewichteten Summen im Pretest in einen aufzuwertenden und einen abzuwertenden Cue eingeteilt wurden (für einen deskriptiven Überblick siehe Tabelle B-3 in Anhang B). Eine pro Bedingung durchgeführte explorative Analyse mittels Boxplots zeigte keinen extremen Ausreißer. Es wurde eine messwiederholte ANOVA mit „Messzeitpunkt“ sowie „Cue“ (aufzuwertend und abzuwertend) als *within subjects*-Faktoren und „Bedingung“ als *between subjects*-Faktor berechnet. Die für die vorliegende Fragestellung interessierenden Interaktionen „Messzeitpunkt x Cue“ sowie „Messzeitpunkt x Cue x Bedingung“ wurden nicht signifikant, $F(1, 90) < 1$, $p = .82$, $\eta_p^2 = .00$ bzw. $F(1, 90) = 1.07$, $p = .30$, $\eta_p^2 = .01$. Post-hoc durchgeführte Teststärkenanalysen mit dem Programm GPower (Faul et al., 2007) ergaben für diese beiden Interaktionen, dass bei $\alpha = .05$ und $N = 92$ ein mittlerer Effekt ($f = .25$, Cohen, 1988) mit einer Teststärke von jeweils > .99 entdeckt werden kann. Signifikant wurden lediglich die Interaktion „Messzeitpunkt x Bedingung“, $F(1, 90) = 5.53$, $p < .05$, $\eta_p^2 = .06$, sowie der Haupteffekt „Cue“, $F(1, 90) = 195.96$, $p < .001$, $\eta_p^2 = .69$. Der Haupteffekt „Cue“ besagt, dass sich die Validitätseinschätzungen je nach Cue unterscheiden. Eine Inspektion der Mittelwerte verdeutlicht, dass die Validitätseinschätzung des aufzuwertenden Cues signifikant höher ist als die des abzuwertenden Cues, was eine logische Folge der vorgenommenen Klassifikation ist. Getrennt für beide Bedingungen berechnete ANOVAs mit Messwiederholung spezifizieren, dass die signifikante Interaktion „Messzeitpunkt x Bedingung“ darauf zurückzuführen ist, dass in der zweiten Bedingung die Validitätseinschätzungen im Posttest signifikant höher sind als im Pretest (Haupteffekt „Messzeitpunkt“), $F(1, 45) = 4.82$, $p < .05$, $\eta_p^2 = .10$. Dieser Effekt war in der ersten Bedingung nicht signifikant, $F(1, 45) = 1.22$, $p = .28$, $\eta_p^2 = .03$. Ergebnisse von *t*-Tests zeigen, dass in der zweiten Bedingung zwar beide Cues im Posttest als valider eingeschätzt werden als im Pretest, diese Aufwertung überraschenderweise aber lediglich beim abzuwertenden Cue signifikant ist, $t(45) = 2.09$, $p < .05$, $d_z = .31$. Beim aufzuwertenden Cue war keine signifikante Veränderung zu beobachten, $t(45) = 1.31$, $p = .20$, $d_z = .19$.[24]

[23] 34 Versuchspersonen kannten eine der Städte und 58 kannten keine Stadt. Es wurde eine Kovarianzanalyse mit der Stadtkenntnis als Kovariate und der Veränderung der Differenz zwischen den korrigierten gewichteten Summen als abhängiger Variable berechnet. Die Berücksichtigung der Kovariate führte zum gleichen Ergebnis.

[24] Um den Einfluss der Kenntnis der Städte zu ermitteln, wurde erneut eine Kovarianzanalyse berechnet. Als abhängige Variable wurde die mittlere Cue-Umwertung verwendet. Die Berücksichtigung der Kovariate veränderte das Ergebnismuster nicht.

Diskussion

Zur Untersuchung der Fragestellung, ob subjektive Cue-Validitäten auch bei einer ausgeglichenen Informationslage verändert werden und ob die Wiederholung der Entscheidungssituation einen Einfluss auf die Richtung und Stärke der Umwertung hat, wurden Analysen auf der Ebene der Optionen und der Ebene der Cues durchgeführt. Auf der Ebene der Optionen konnten weder ein signifikantes *spreading apart* der Optionen noch ein Einfluss der Wiederholung gezeigt werden. Auch auf der Ebene der Cues konnten keine systematischen Umwertungen gefunden werden. Die in der zweiten Bedingung vergleichsweise höheren Validitätseinschätzungen im Posttest waren nach dem PCS-Modell nicht zu erwarten und können auf Basis der vorliegenden Daten nicht plausibel erklärt werden. Folglich konnte die aufgestellte Hypothese, dass – wie nach dem PCS-Modell zu erwarten war – systematische Umwertungen subjektiver Cue-Validitäten bei einer ausgeglichenen Informationslage auftreten, nicht bestätigt werden. Ob tatsächlich die Ausgeglichenheit der Informationslage das Auftreten systematischer Umwertungen verhindert hat, ist in folgenden Experimenten zu prüfen. Auffallend ist zudem die hohe Anzahl an Switchern. Inwiefern dies der Schwierigkeit der Entscheidungssituation oder dem Material geschuldet ist, bleibt offen. Bereits Glöckner et al. (im Druck, Experiment 3) hatten auf das Städteparadigma zurückgegriffen und systematische Umwertungen gefunden, weshalb es unwahrscheinlich ist, dass die Ergebnisse nur auf das Material zurückzuführen sind. Da bereits in der ersten Bedingung keine systematischen Umwertungen gefunden werden konnten, konnte der Einfluss der Wiederholung auf die Umwertung subjektiver Cue-Validitäten auf Basis der vorliegenden Daten nicht bestimmt werden.

Die Modifikation der Entscheidungssituation war damit im Sinne des PCS-Modells zwar nicht erfolgreich, hat aber hinsichtlich des Anspruchs, das Verständnis des Konsistenzmaximierungsprozesses in probabilistischen Inferenzentscheidungen zu vertiefen, auf einen möglichen Einflussfaktor hingewiesen: die Ausgeglichenheit der Informationslage. Den Einfluss dieser und weiterer (personaler) Faktoren wird in den folgenden Experimenten überprüft. Darauf aufbauend wird der Einfluss der Wiederholung als weiterer Einflussfaktor untersucht.

3.5 DER EINFLUSS DES INFORMATIONSMUSTERS UND CHARAKTERISTIKA DER PERSON AUF DIE VERÄNDERUNG SUBJEKTIVER CUE-VALIDITÄTEN

Anhand des ersten Experimentes konnte Evidenz dafür geliefert werden, dass die systematische Umwertung von Cue-Validitäten auf die Reflexion über eine Entscheidungssituation zurückgeführt werden kann. Der nächste Schritt bei der Erforschung des Prozesses der automatischen Konsistenzmaximierung besteht in der Identifikation und systematischen Untersuchung potentieller Einflussfaktoren. Einen Hinweis auf einen möglichen Einflussfaktor hat Experiment 2 gegeben: So fand keine systematische Umwertung von Cues statt, wenn die Informationslage ausgeglichen war. Im Rahmen der folgenden Experimentalserie wurde die Bedeutung der Ausgeglichenheit der Informationslage für die prädezisionale Informationsumwertung gezielt untersucht. Unter Berücksichtigung von Experimenten zum Zusammenhang zwischen der prä- oder postdezisionalen Informationsumwertung und Persönlichkeitsmaßen (z.B. Beckmann und Kuhl, 1984; Cialdini et al., 1995; Russo et al., 1998; Russo et al., 2000) wurden des Weiteren einige ausgewählte personale Faktoren im Rahmen der folgenden Experimentalserie berücksichtigt.

3.5.1 Experiment 3: Die Veränderung subjektiver Cue-Validitäten bei ausgeglichenen und unausgeglichenen Informationslagen

ZIELSETZUNG

Das vornehmliche Ziel des zu berichtenden Experimentes bestand darin, den Einfluss der Anzahl und der Verteilung der Cues auf das Auftreten und die Stärke prädezisionaler Informationsumwertungen zu untersuchen. Diese Frage stellte sich, da (a) der Einfluss der Ausgeglichenheit der Informationsauflage auf die Umwertung von Cues in probabilistischen Inferenzentscheidungen bisher nicht untersucht worden ist und (b) die Ergebnisse des zweiten Experimentes einen Hinweis darauf liefern, dass keine oder eine schwächere Umwertung von Cues bei ausgeglichenen Informationslagen auftritt.

METHODE

Design. Die Veränderung der subjektiven Cue-Validitäten wurde auch in diesem Experiment *within subjects* in einem Pre-/Post-Design gemessen. So wurde die subjektive Einschätzung der Validität verschiedener Cues sowohl im Pretest als auch im Posttest

erhoben. *Between subjects* wurden die Anzahl und Verteilung der Cues variiert. Alle Versuchspersonen beantworteten schließlich, wie häufig sie mit den Cues Erfahrungen sammeln, und demografische Fragen.

Hypothesen. Gemäß dem Prinzip der automatischen Konsistenzmaximierung wurde eine systematische Umwertung von Cue-Validitäten unabhängig von der Informationslage erwartet. Aus dem PCS-Modell lassen sich jedoch keine konkreten Hypothesen bezüglich möglicher Unterschiede in der Stärke der Umwertungen ableiten. Angesichts der Ergebnisse des zweiten Experimentes wurde aber erwartet, dass eine Interaktion zwischen der Stärke der Veränderung der Cue-Validitäten und dem dargebotenen Muster besteht. So wurde angenommen, dass prädezisionale Informationsumwertungen stärker sind, wenn die Informationslage unausgeglichen ist, als wenn sie ausgeglichen ist. Explorativ wurde untersucht, ob in ausgeglichenen Informationslagen die Anzahl der Cues einen Einfluss auf die Stärke der Umwertung hat. Zudem wurde explorativ der Frage nachgegangen, ob die Erfahrung, im Sinne der Häufigkeit der Konfrontation mit einem Cue, mit der Veränderbarkeit dieses Cues zusammenhängt.

Material. In dem vorliegenden Experiment wurde die gleiche Coverstory wie im ersten Experiment verwendet. So wurden die Versuchspersonen im Pre- wie auch im Posttest angewiesen, die Validität verschiedener Wetterberichte für das Wetter an einem Ferienort in sieben Tagen einzuschätzen. Die Versuchspersonen der Bedingung 1 schätzten die Validität der Cues „ZDF“ und „www.wetteronline.de“ ein, die Versuchspersonen der Bedingung 2 die Validität der Cues „www.wetteronline.de“, „Sat.1“, „Bild-Zeitung“ und „ZDF“ und die Versuchspersonen der Bedingung 3 gaben eine Validitätseinschätzung für die Cues „Süddeutsche Zeitung“, „ZDF“, „Bild“, „www.N24.de“, „Sat.1“ und „www.wetteronline.de“ ab. Die entsprechende Skala reichte von 50% bis 100%, wobei lediglich die Endpunkte der Skala mit diesen prozentualen Angaben bezeichnet waren. Zwischen dem Pre- und dem Posttest wurde den Versuchspersonen jeweils eine Matrix dargeboten, die Informationen darüber enthielt, ob Wetterberichte verschiedener Quellen für zwei Ferienorte A und B sonniges Wetter vorhersagen (+) oder nicht vorhersagen (-). Während in den Bedingungen 1 und 3 eine ausgeglichene Informationslage präsentiert wurde, da jeweils ein Cue gegen einen anderen sprach (Bedingung 1) bzw. jeweils drei Cues gegen drei andere Cues sprachen (Bedingung 3), wurde in der Bedingung 2 eine unausgeglichene Informationslage dargeboten, da ein Cue gegen drei andere Cues sprach.[25] Im Anschluss an den Posttest

[25] Bedingung 2 entsprach dem in Experiment 1, Bedingung 1 verwendeten Material.

wurden die Versuchspersonen gebeten, auf einer 5-stufigen Skala anzugeben, wie häufig sie die Wetterberichte der jeweils präsentierten Quellen sehen bzw. lesen. Demografische Angaben bildeten den Abschluss des Fragebogens. Die Instruktionen sind in Anhang A-2 und A-4 zu finden.

Stichprobe. Es nahmen insgesamt 119 Studenten (78 weiblich, 37 männlich, 4 ohne Angabe; Durchschnittsalter: 22.18 Jahre, *SD* = 2.77) der Universität Erfurt an dem Experiment teil. Diese Versuchspersonen waren über das Onlinerekrutierungssystem *ORSEE* (Greiner, 2004) angeworben worden. Das Experiment, das Bestandteil einer Experimentalbatterie war, wurde als Papierfragebogen durchgeführt. Die Erhebung dauerte insgesamt 40-50 Minuten, wobei etwa 5-10 Minuten auf das hier zu berichtende Experiment entfielen. Die Versuchspersonen erhielten als Aufwandsentschädigung entweder 6 € oder eine Versuchspersonenstunde. Die Versuchspersonen wurden zufällig, aber mit dem Ziel einer Gleichverteilung den drei Bedingungen zugeteilt.

Durchführung. Zunächst wurde den Versuchspersonen wie in Experiment 1 das Konzept der Validität am Beispiel der Vorhersagekraft von Wetterberichten anschaulich erklärt. Daraufhin wurden die Versuchspersonen instruiert, eine möglichst gute und differenzierte Einschätzung der Validität verschiedener Wetterberichte (Cues) für das Wetter an einem Ferienort in sieben Tagen abzugeben. Je nach dem, welcher Bedingung die Versuchsperson zufällig zugeteilt worden war, wurden ihr zwei (Bedingung 1), vier (Bedingung 2) oder sechs (Bedingung 3) Quellen präsentiert. Die Einschätzung wurde auf einer Skala von 50%-100% vorgenommen. Nach einem thematisch unverbundenen Experiment, das 20-25 Minuten dauerte, sollten sich die Versuchspersonen im zweiten Teil des Experimentes vorstellen, dass sie in den Urlaub fahren möchten und dass sie bei der Wahl zwischen zwei möglichen Ferienorten A und B an den Ort reisen möchten, an dem sie eher sonniges Wetter erwarten. Den Versuchspersonen wurde mitgeteilt, aus welchen Quellen sie die Wetterberichte bezögen, wobei diese Quellen den im Pretest präsentierten Quellen entsprachen. Des Weiteren wurden sie darauf hingewiesen, dass sie noch keine endgültige Entscheidung treffen sollen, da sie noch einen weiteren Wetterbericht (der ARD) abwarteten. Je nach Bedingung sahen die Versuchspersonen nun verschiedene Matrizen: In Bedingung 1 sagte „www.wetteronline.de“ für den Ferienort A sonniges Wetter vorher und „ZDF“ für den Ferienort B. In Bedingung 2 sprach der Wetterbericht von „Sat.1“ für sonniges Wetter an Ort A, während diejenigen von „www.wetteronline.de“, „Bild“ und „ZDF“ für den Ort B sprachen. In Bedingung 3 sagten die „Süddeutsche Zeitung“, „ZDF“ und

„www.wetteronline.de“ sonniges Wetter an Ferienort A vorher, „Bild“, „www.N24.de“ und „Sat.1“ aber für Ort B (siehe Anhang A-4). Die Versuchspersonen sollten nun überlegen, an welchem Ort sie aufgrund der vorliegenden Informationen tendenziell eher sonniges Wetter erwarteten. Eine tendenzielle oder finale Entscheidung wurde nicht getroffen. Daraufhin schätzten die Versuchspersonen wie im Pretest die Validität der in ihrer jeweiligen Bedingung präsentierten Cues auf einer Skala von 50% bis 100% ein und gaben auf einer 5-stufigen Skala an, wie häufig sie die entsprechenden Wetterberichte sehen bzw. lesen. Abschließend wurden demografische Angaben erbeten.

ERGEBNISSE

Vorausgehende Analysen. Vier Versuchspersonen schätzten die Validität eines jeden Cue mit 50% oder 100% ein. Da dies auf ein mangelndes Verständnis der Validitätsskala oder des -konzepts schließen lässt, wurden diese Versuchspersonen aus den folgenden Analysen ausgeschlossen.[26] Von den verbleibenden 115 Versuchspersonen (37 männlich, 75 weiblich, 3 ohne Angabe; Durchschnittsalter: 22.18 Jahre, *SD* = 2.78) waren 37 zufällig der ersten Bedingung, 38 der zweiten Bedingung und 40 der dritten Bedingung zugeteilt worden. Die Validitätseinschätzungen sind Tabelle B-4 in Anhang B zu entnehmen.

Umwertung von Cue-Validitäten. Sowohl auf der Ebene der Optionen als auch auf der Ebene der Cues wurde analysiert, ob eine signifikante systematische Umwertung von subjektiven Validitäten stattgefunden hat. Zunächst wurden wie in den vorherigen Experimenten die korrigierten gewichteten Summen für jede Option berechnet, und zwar sowohl für den Pretest als auch für den Posttest. Basierend auf der gemäß den korrigierten gewichteten Summen im Pretest zu erwartenden Entscheidung wurde für beide Messzeitpunkte die Differenz zwischen den korrigierten gewichteten Summen ermittelt. Pro Bedingung durchgeführte explorative Analysen mittels Boxplots zeigten keine extremen Ausreißer. Switcher (n = 6) wurden erneut aus den folgenden Analysen ausgeschlossen.

Um die Veränderung der Differenzen zwischen den korrigierten gewichteten Summen der Optionen zu untersuchen, wurde eine messwiederholte ANOVA mit „Messzeitpunkt“ als *within subjects*-Faktor und „Muster“ als *between subjects*-Faktor berechnet. Eine post-hoc durchgeführte Teststärkenanalyse mit GPower (Faul et al., 2007) ergab, dass es das Anlegen eines α-Niveaus von .05 bei N = 109 erlaubt, mittlere Effekte (f = .25, Cohen, 1988)

[26] Da es nur Versuchspersonen gab, die sowohl im Pre- als auch im Posttest die Validitäten stets mit 50% oder 100% eingeschätzt haben, ist eine Differenzierung zwischen weichem und hartem Kriterium nicht erforderlich.

mit Teststärken von > .99 (Haupteffekt „Messzeitpunkt“, Interaktion „Messzeitpunkt x Muster“) und von .66 (Haupteffekt „Bedingung“) zu finden. Es zeigten sich signifikante Effekte des Messzeitpunktes, $F(1, 106) = 17.57$, $p < .001$, $\eta_p^2 = .14$, des Musters, $F(2, 106) = 68.01$, $p < .001$, $\eta_p^2 = .56$, und der Interaktion des Messzeitpunktes und des Musters, $F(2, 106) = 3.68$, $p < .05$, $\eta_p^2 = .07$. Der Haupteffekt „Messzeitpunkt“ ist mit der – gemittelt über alle Bedingungen – im Posttest höheren Differenz zwischen den korrigierten gewichteten Summen zu begründen, der Haupteffekt „Muster“ mit der – gemittelt über beide Messzeitpunkte – höheren Differenz für das Muster der zweiten Bedingung. Die Interaktion ist für die vorliegende Fragestellung besonders interessant.[27] Um diese näher zu untersuchen, wurden mittels geplanter gepaarter *t*-Tests die Veränderungen der Differenzen zwischen den korrigierten gewichteten Summen pro Muster geprüft. Während sich für jedes Muster die Tendenz abzeichnete, dass die Differenz zwischen den korrigierten gewichteten Summen im Posttest größer ist als im Pretest (siehe Abbildung 5), zeigte sich nur für das zweite Muster ein signifikantes Ergebnis, $t(37) = 3.65$, $p < .01$, $d_z = .59$.

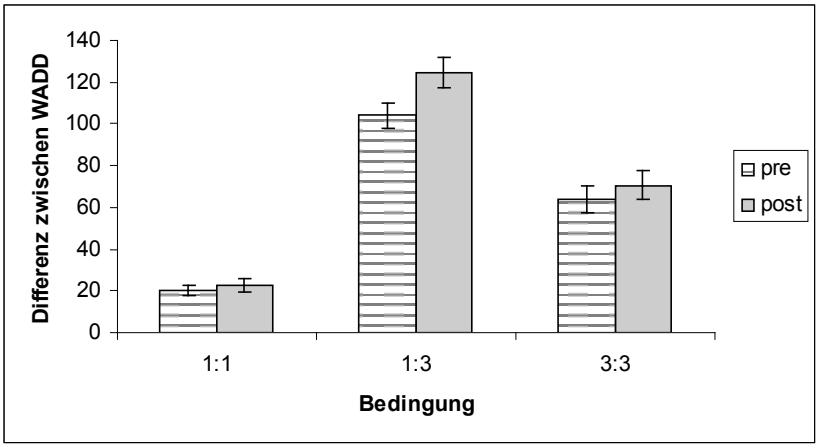


Abbildung 5: Differenz zwischen den korrigierten gewichteten Summen (WADD) der Optionen zum Pretest (pre) und zum Posttest (post). Die Bedingungsbezeichnungen ergeben sich aus der Art des präsentierten Cue-Musters. Fehlerbalken zeigen Standardfehler an.

Um eine mögliche Umwertung von subjektiven Cue-Validitäten auf der Ebene der Cues zu analysieren, wurden die Cues zunächst gemäß der Entscheidung, die nach der Höhe der korrigierten gewichteten Summen im Pretest zu erwarten ist, in aufzuwertende und

[27] Um dem möglichen Einwand zu begegnen, dass die Ergebnisse aufgrund einer Konfundierung von Cue-Anzahl und Höhe des *spreading apart* verzerrt sind, wurde zudem eine univariate ANOVA mit der mittleren Cue-Umwertung (= Höhe des *spreading apart* dividiert durch die doppelte Anzahl der Cues) als abhängiger Variable und „Bedingung“ als Zwischensubjektfaktor berechnet. Auch hier zeigte sich der Bedingungsunterschied.

abzuwertende Cues eingeteilt (für einen deskriptiven Überblick siehe Tabelle B-5 in Anhang B). Explorative Analysen mittels Boxplots zeigten in keiner Bedingung wiederholte extreme Ausreißer. Im Anschluss wurde zur Untersuchung der Veränderung der Cue-Validitäten für jede Bedingung eine messwiederholte ANOVA mit „Messzeitpunkt“ und „Cue“ als *within subjects*-Faktoren berechnet. Für die interessierende Interaktion „Messzeitpunkt x Cue“ post-hoc durchgeführte Teststärkenanalysen mit GPower (Faul et al., 2007) ergaben, dass bei $\alpha = .05$ ein mittlerer Effekt ($f = .25$, Cohen, 1988) in der ersten Bedingung ($n = 33$) mit einer Teststärke von .97, in der zweiten Bedingung ($n = 38$) von .87 und in der dritten Bedingung ($n = 37$) von .95 gefunden werden kann. Zwar zeigte sich für das Muster 1 auf der deskriptiven Ebene eine systematische Umwertung eines Cues, jedoch wurde die Interaktion „Messzeitpunkt x Cue“ nicht signifikant, $F(1, 32) = 2.81$, $p = .10$, $\eta_p^2 = .08$.[28] Für die Muster 2 und 3 wurde diese Interaktion hingegen signifikant, $F(1.83, 67.86) = 7.78$, $p < .01$, $\eta_p^2 = .17$[29] bzw. $F(4.23, 152.40) = 2.72$, $p < .05$, $\eta_p^2 = .07$[30]. Für alle Muster wurden für jeden Cue gepaarte *t*-Tests (einseitig) berechnet: Nur in Muster 2 wurden alle vier Cues systematisch und signifikant umgewertet (siehe Abbildung 6).

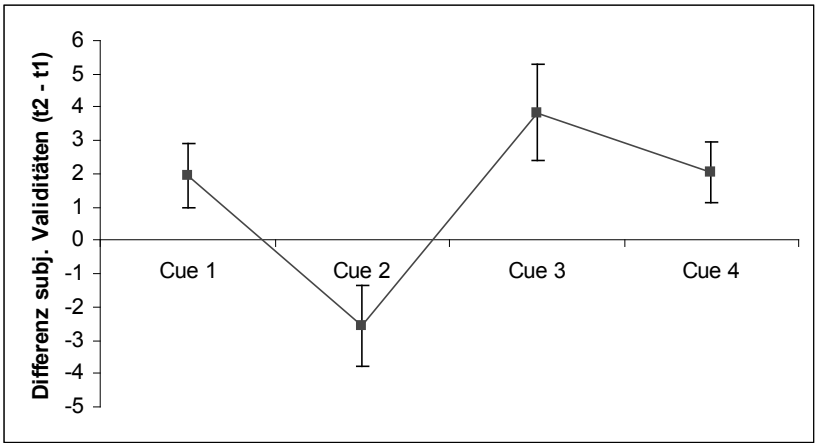


Abbildung 6: Differenz der subjektiven Cue-Validitäten zwischen den beiden Messzeitpunkten t1 (Pretest) und t2 (Posttest). Es wurde die Validitätseinschätzung des Pretests von derjenigen des Posttests subtrahiert. Werte > 0 kennzeichnen folglich eine Aufwertung, Werte < 0 eine Abwertung. Cue 1, 3 und 4 sind als aufzuwertend klassifizierte Cues, Cue 2 ist ein abzuwertender Cue. Die Fehlerbalken stellen Standardfehler der Differenzwerte dar.

[28] Eine weitere Versuchsperson musste ausgeschlossen werden, da die gewichteten Summen im Pretest gleich hoch waren und folglich keine Klassifikation der Cues vorgenommen werden konnte.

[29] Wegen einer Verletzung der Sphärizitätsannahme wurden die Freiheitsgrade korrigiert. Der Empfehlung von Rasch et al. (2006b), bei $\varepsilon < .75$ eine Korrektur nach Greenhouse-Geisser (gemäß SPSS) vorzunehmen, wurde hier wie auch im Folgenden gefolgt. Diese Verletzung wurde bei der Teststärkenberechnung berücksichtigt.

[30] Da eine Verletzung der Sphärizitätsannahme vorlag und $\varepsilon > .75$, wurde eine Korrektur der Freiheitsgrade nach Huynh-Feldt vorgenommen. Diese Verletzung wurde bei der Teststärkenberechnung berücksichtigt.

In Muster 3 wurden nur drei Cues systematisch umgewertet und diese Umwertung war lediglich bei einem Cue signifikant. In Muster 1 wurde ein Cue systematisch und signifikant abgewertet. Tabelle 2 gibt einen Überblick über die Ergebnisse der *t*-Tests pro Bedingung.

Tabelle 2: Ergebnisse der *t*-Tests pro Bedingung

	Cue	*t*-Wert	*df*	*p* (einseitig)	d_z
Bedingung 1 (1:1)	aufzuwertend	-0.24	32	.41	-.04
	abzuwertend	-2.15	32	< .05	-.37
Bedingung 2 (1:3)	aufzuwertend_1	2.04	37	< .05	.33
	aufzuwertend_2	2.66	37	< .01	.43
	aufzuwertend_3	2.25	37	< .05	.36
	abzuwertend	-2.12	37	< .05	-.34
Bedingung 3 (3:3)	aufzuwertend_1	0.66	36	.26	.11
	aufzuwertend_2	-1.28	36	.10	-.21
	aufzuwertend_3	-0.02	36	.49	-.00
	abzuwertend_1	0.64	36	.26	.10
	abzuwertend_2	-2.35	36	< .05	-.39
	abzuwertend_3	-1.11	36	.14	-.18

Anmerkungen: In Bedingung 1 wurde ein 1:1-Muster dargeboten, in Bedingung 2 ein 1:3-Muster und in Bedingung 3 ein 3:3-Muster. Die Cues wurden entsprechend der zu erwartenden Entscheidung in aufzuwertende und abzuwertende Cues eingeteilt. Ein Cue ist aufzuwertend, wenn er für die zu erwartende Entscheidung spricht, und abzuwertend, wenn er dieser widerspricht. Die Indizes dienen der Unterscheidung verschiedener Cues.

Weitere Ergebnisse: Der Einfluss der Erfahrung. Um zu ermitteln, ob ein häufigerer Konsum der einzelnen Quellen von Wetterberichten zu einer verminderten Umwertung von Cue-Validitäten führt, wurden pro Bedingung Korrelationen zwischen der absoluten Differenz zwischen der Pretest- und der Posttest-Einschätzung und der Häufigkeit des Konsums berechnet. Die Verwendung der absoluten Differenz erlaubt es, die Richtung der Umwertung, die abhängig von der tendenziellen Entscheidung ist, zu vernachlässigen und gibt Aufschluss über die generelle Veränderbarkeit eines Cues. Die Korrelationen sind in Tabelle 3 aufgeführt.

Tabelle 3: Korrelationen zwischen der Cue-Umwertung (absolut) und der Häufigkeit der Erfahrung

	Bedingung 1 (1:1)	Bedingung 2 (1:3)	Bedingung 3 (3:3)
ZDF	.110	-.114	-.107
Wetteronline.de	.106	.138	-.127
Sat.1		.030	.190
Bild		.445**	.054
Süddeutsche Zeitung			-.084
N24.de			.040
n	32	38	37

Anmerkungen: In Bedingung 1 wurde ein 1:1-Muster dargeboten, in Bedingung 2 ein 1:3-Muster und in Bedingung 3 ein 3:3-Muster. **: $p < .01$.

Mit Ausnahme der Korrelation zwischen der absoluten Differenz der Pretest- und Posttest-Einschätzungen der Validität der Bild-Zeitung und der Häufigkeit des Konsums der Bild-Zeitung sind alle Korrelationen sehr niedrig und nicht signifikant. Für diese Cues kann die Hypothese, dass zunehmende Erfahrung zu einer stabileren Validitätseinschätzung führt, nicht bestätigt werden. Die signifikant positive Korrelation zwischen der absoluten Differenz der Validität des Cues „Bild“ und der Häufigkeit des Konsums dieser Zeitung spricht dafür, dass eine zunehmende Erfahrung mit einer erhöhten Veränderbarkeit einhergeht. Nach Plausibilitätsüberlegungen wäre ein negativer Zusammenhang zu erwarten gewesen. Es ist allerdings zu berücksichtigen, dass der Konsum von Bild hauptsächlich mit „nie“ und „selten“ eingeschätzt wurde, vereinzelt mit „manchmal“. Insofern ist eher davon zu sprechen, dass ein *relativ* häufiger Konsum mit einer stärkeren Veränderbarkeit einhergeht. Wichtiger bzw. problematischer ist aber die eingeschränkte Varianz. Des Weiteren ist zu beachten, dass die Korrelation lediglich in der zweiten Bedingung hoch und signifikant ausfällt. In dieser Bedingung hatte der Cue „Bild“ eine erhöhte Chance auf Veränderbarkeit, da er mit der Mehrheit der Cues für eine Option sprach und damit als aufzuwertender Cue klassifiziert wurde und zudem relativ niedrige mittlere Pretest-Validitätswerte von 64.67, *SD* = 10.96 (Bedingung 2) aufwies. Diese Ausführungen deuten an, dass diese Korrelation nicht überbewertet werden sollte. Dennoch ist ein inhaltlicher Effekt des Cues möglich. Dies gilt es zu diskutieren.

DISKUSSION

Hinsichtlich der Umwertung von subjektiven Cue-Validitäten haben sich sowohl auf der Ebene der Optionen als auch auf der Ebene der Cues deutliche Unterschiede zwischen den einzelnen Mustern gezeigt. So ist zwar in jeder Bedingung ein tendenzielles *spreading apart* der Optionen zu erkennen, doch verweist die signifikante Interaktion zwischen Messzeitpunkt und Muster auf einen Einfluss der Informationslage. Nachfolgende Tests zeigten, dass nur in dem zweiten Muster, in dem ein Cue gegen drei andere sprach, ein signifikantes *spreading apart* zu verzeichnen war. Auf der Ebene der Cues zeigte sich auch für das erste Muster eine teilweise systematische Umwertung, allerdings wurde die Interaktion zwischen Messzeitpunkt und Cue nur für die Muster zwei und drei signifikant. Gepaarte *t*-Tests offenbarten weitere Unterschiede: Während beim dritten Muster nur drei Cues systematisch umgewertet wurden und hiervon nur ein einziger signifikant, wurden beim zweiten Muster alle vier Cues systematisch und signifikant umgewertet. Beim ersten Muster wurde ein Cue systematisch und signifikant abgewertet. Aus diesen Analysen lässt sich schließen, dass die Ausgeglichenheit der Informationslage einen Einfluss auf die Veränderung von Cue-Validitäten hat, und zwar derart, dass eine unausgeglichene Verteilung eher zu einer (stärkeren) Umwertung führt als eine ausgeglichene. Ob bei gleicher Verteilung die Anzahl an Cues einen Einfluss ausübt, kann auf Grundlage der vorliegenden Ergebnisse nicht abschließend beantwortet werden, da sich auf der Ebene der Cues zwar eine signifikante Interaktion nur beim dritten Muster, in dem die Cue-Anzahl hoch war, finden ließ, in beiden Mustern aber nur ein Cue signifikant abgewertet wurde.

Damit konnten die Hypothesen nur teilweise bestätigt werden: Die nach dem PCS-Modell zu erwartende systematische Umwertung subjektiver Cue-Validitäten in allen Mustern konnte nicht beobachtet werden. Ein Einfluss der Informationslage konnte wie erwartet bestätigt werden. Diese Vorhersage wurde aber auf Basis der Ergebnisse vorheriger Experimente generiert. Vor dem Hintergrund der bisherigen Befunde zur prädezisionalen Informationswertung hätte auch erwartet werden können, dass gerade eine ausgeglichene Informationslage stärkere Verzerrungen bewirkt, wenn man nämlich Ausgeglichenheit im Sinne von Ambiguität interpretiert. So wird vor dem Hintergrund bisheriger Erkenntnisse vermutet, dass ambigue Informationen zu einer stärkeren Reevaluation führen (DeKay et al., 2009; Simon, 2004). Angesichts dieser Ausführungen erscheint es ratsam, zunächst die Ergebnisse zu replizieren, bevor Vermutungen über die Gründe angestellt werden. Dies ist vor allem aufgrund einer methodischen Schwäche ratsam: So beruhen die Ergebnisse auf der

Annahme, dass die Personen die Option wählen, die die höhere gewichtete Summe im Pretest hatte. Obgleich diese Annahme angesichts der Befunde von Glöckner (2008a) durchaus plausibel ist, wäre der Vergleich mit einer tatsächlich getroffenen Entscheidung überzeugender.

Als weiteres Ergebnis kann festgehalten werden, dass mehrheitlich kein Zusammenhang zwischen der Häufigkeit des Konsums und der Stabilität der Einschätzung zu finden war. Der positive Zusammenhang, der bei dem Cue „Bild“ gefunden wurde, ist sowohl aus Plausibilitäts- als auch aus statistischen Gründen eher mit Vorsicht zu interpretieren. Dennoch ist zu bedenken, dass die inhaltliche Bedeutung eines Cues dessen Veränderbarkeit beeinflussen kann. Glöckner (2006) hat derartige Unterschiede bezüglich der Veränderbarkeit von Cues diskutiert. Denkbar ist, dass wenn man einen Cue als besonders positiv bzw. negativ einschätzt, die Resistenz zur Abwertung bzw. Aufwertung größer ist. Inwiefern dies zutrifft, kann auf Basis der vorliegenden Daten natürlich nicht beantwortet werden. Die Untersuchung spezifischer Cues wird zudem in dieser Arbeit nicht verfolgt, bietet aber eine Anregung für weitere Forschungsarbeiten.

Zusammenfassend betrachtet deuten die vorliegenden Ergebnisse darauf hin, dass die Ausgeglichenheit der Informationslage einen möglichen Einflussfaktor bei der Umwertung von Cue-Validitäten darstellt. Ziel der nachfolgenden Studien ist es, diesen Einfluss zu replizieren und weitere Faktoren zu identifizieren.

3.5.2 Experiment 4: Die Veränderung subjektiver Cue-Validitäten in Abhängigkeit von dem Informationsmuster und Charakteristika der Person

ZIELSETZUNG

In dem vorliegenden Experiment wurde erstens der Frage nachgegangen, ob sich das Ergebnis, dass die Ausgeglichenheit der Informationslage, d.h. die Verteilung der Cues, einen Einfluss auf die Stärke der prädezisionalen Umwertung hat, mit anderem Material und methodischen Verbesserungen replizieren lässt. Zweitens sollte neben dem Einfluss der Informationslage auch der Einfluss von Persönlichkeitsfaktoren auf die Stärke der prädezisionalen Umwertung von Cue-Validitäten ermittelt werden. Es wurden Maße gewählt, die in verschiedener Weise das Bedürfnis nach Konsistenz bzw. die Abneigung gegenüber Inkonsistenz erfassen.

METHODE

Design. Erneut wurden die subjektiven Cue-Validitäten *within subjects* zu zwei Messzeitpunkten erfasst. Als Treatment wurden *between subjects* die Anzahl und die Verteilung der Cues variiert. Im Gegensatz zum vorherigen Experiment trafen die Versuchspersonen nach dem Posttest eine endgültige Entscheidung und gaben die Konfidenz in diese Entscheidung an. Darüber hinaus bearbeiteten alle Versuchspersonen drei Persönlichkeitsfragebögen, die in einer der 3! = 6 möglichen Reihenfolgen den einzelnen Versuchspersonen zufällig, aber mit dem Ziel einer Gleichverteilung ausgehändigt wurden.

Hypothesen. Es wurde gemäß der Ergebnisse des vorherigen Experimentes eine Interaktion zwischen dem Messzeitpunkt und dem dargebotenen Muster erwartet. So wurde angenommen, dass die Veränderung der Validitäten dann am stärksten ist, wenn ein Cue gegen drei andere spricht. Bezüglich des generellen Auftretens systematischer Umwertungen wurde die nach dem PCS-Modell abzuleitende Hypothese, dass diese in allen Mustern zu beobachten ist, beibehalten.

Bezüglich der Persönlichkeitsvariablen wurde aus Plausibilitätsgründen (siehe Kapitel 2.4.3) erwartet, dass diese einen Zusammenhang mit der Stärke der prädezisionalen Umwertung von Cue-Validitäten aufweisen. Es wurden neben der deutschen Kurzskala zur Erfassung des Bedürfnisses nach kognitiver Geschlossenheit (16-NCCS, Schlink und Walther, 2007) die Kurzform der *Maximization Scale* (MAX, Nenkov, Morrin, Ward, Schwartz & Hulland, 2008) und die Kurzskala zur Erfassung der Präferenz für Konsistenz (PFC-B, Cialdini et al., 1995) eingesetzt, wobei die letzten beiden Skalen ins Deutsche übersetzt wurden. Die Instrumente werden später näher erläutert. Bezüglich des *Need for Cognitive Closure* (NCC) wurde erwartet, dass ein positiver Zusammenhang mit der Stärke an prädezisionaler Umwertung besteht, da durch die Umwertung Ambiguität verringert wird. Entsprechend der von Schwartz et al. (2002) vorgenommenen Unterscheidung zwischen *satisficern* und *maximizern* wurde angenommen, dass eine positive Korrelation zwischen dem in der *Maximization Scale* erreichten Score und der Stärke der prädezisionalen Umwertung besteht. Bezüglich der *Preference for Consistency* (PFC) wurde erwartet, dass Personen, die eine hohe Präferenz für Konsistenz haben, auch eine stärkere Umwertung von Cue-Validitäten zeigen. Hinsichtlich der Konfidenzurteile wurde angenommen, dass diese in einem positiven Zusammenhang mit der Größe der Differenz zwischen den gewichteten Summen stehen (vgl. Glöckner und Betsch, 2008c).

Material[31]. In einem Vortest schätzten 39 Versuchspersonen (20 weiblich, 19 männlich; Durchschnittsalter: 25 Jahre, *SD* = 5.91) die Validität von 16 Cues ein.[32] Basierend auf diesen Einschätzungen wurden sechs Cues aus dem mittleren Validitätsspektrum für das vorliegende Experiment ausgewählt: Geruch (*MW* = 84.89, *SD* = 14.38), Lagerdauer (*MW* = 77.63, *SD* = 16.10), Lagerklima (*MW* = 75.78, *SD* = 16.62), Reifeperiode (*MW* = 74.68, *SD* = 17.38), Transport (*MW* = 73.28, *SD* = 15.95) und Färbung (*MW* = 65.45, *SD* = 18.23). Im Pre- und im Posttest des vorliegenden Experiments wurden die Versuchspersonen instruiert, die Validität von – je nach Bedingung – zwei, vier oder sechs Kennzeichen einer Orange für die Qualität einer Orangensorte auf einer Skala von 50% bis 100% einzuschätzen. Die Versuchspersonen der ersten Bedingung schätzten die Validität der Kennzeichen „Lagerdauer“ und „Reifeperiode“ ein, die Versuchspersonen der zweiten Bedingung die Validität der Kennzeichen „Lagerdauer“, „Reifeperiode“, „Färbung“ und „Lagerklima“ und die Versuchspersonen der dritten Bedingung schätzten die Validität der Kennzeichen „Lagerdauer“, „Reifeperiode“, „Färbung“, „Lagerklima“, „Geruch“ und „Transport“ ein. Als Treatment wurde den Versuchspersonen eine Matrix präsentiert, die Informationen darüber enthielt, ob ein bestimmtes – zuvor eingeschätztes – Kennzeichen bei einer von zwei Orangensorten A und B vorhanden ist (+) oder nicht (-). Im Anschluss an den Posttest sollten die Versuchspersonen entscheiden, von welcher Orangensorte sie eine höhere Qualität erwarten. Die Sicherheit in diese Entscheidung gaben sie auf einer Skala an, deren Endpunkte mit „sehr unsicher“ und „sehr sicher“ bezeichnet waren. Nach der Beantwortung demografischer Fragen bearbeiteten die Versuchspersonen drei Fragebögen in randomisierter Reihenfolge, und zwar die deutsche Kurzskala zur Erfassung des Bedürfnisses nach kognitiver Geschlossenheit (16-NCCS, Schlink und Walther, 2007), eine Kurzform der *Maximization Scale* (MAX, Nenkov et al., 2008) sowie eine Kurzform der *Preference for Consistency Scale* (PFC-B, Cialdini et al., 1995). Der *Need for Cognitive Closure* (NCC) ist nach Webster und Kruglanski (1994) durch verschiedene Aspekte gekennzeichnet, und zwar “desire for predictability, preference for order and structure, discomfort with ambiguity, decisiveness, and closed-mindedness” (S. 1049). Er lässt sich laut Autoren sowohl als personale als auch als situationale Variable auffassen. Schlink und Walther (2007) entwickelten eine deutsche Kurzskala (16 Items) zur Erfassung des Bedürfnisses nach

[31] Ich danke Susann Fiedler sehr herzlich für ihre Mitarbeit bei der Materialerstellung.

[32] Der Vortest wurde im Rahmen einer Experimentalbatterie, die etwa eine Stunde dauerte, im *MPI DecisionLab* durchgeführt. Die Versuchspersonen waren via *ORSEE* (Greiner, 2004) rekrutiert worden und wurden mit 12 € pro Stunde entlohnt.

kognitiver Geschlossenheit, die eine ökonomische, reliable und valide Messung gestattet. Auf einer 6-stufigen Skala geben die Versuchspersonen an, wie stark einzelne Aussagen ihre persönliche Einstellung, Meinung und Erfahrungen ausdrücken. Die Endpunkte sind mit „stimme gar nicht zu“ (1) und „stimme völlig zu“ (6) bezeichnet, während die anderen Skalenpunkte nicht näher benannt werden. Schwartz et al. (2002) differenzieren zwischen *maximizern*, die stets die beste Wahl treffen möchten, und *satisficern*, die lediglich eine zufriedenstellende Wahl treffen möchten. Nenkov und Kollegen (2008) haben die zur Erfassung dieser dispositionalen Variablen entwickelte *Maximization Scale* von Schwartz et al. (2002) analysiert und eine kürzere, 6 Items umfassende Skala erstellt, die sich laut Autoren durch Validität, Reliabilität und eine stabile dimensionale Struktur auszeichnet.[33] Die Versuchspersonen geben auf einer 7-stufigen Skala an, wie stark sie einzelnen Aussagen zustimmen. Die Endpunkte der Skala wurden mit „stimme gar nicht zu“ (1) und „stimme völlig zu“ (7) übersetzt, während die dazwischen liegenden Punkte nicht näher benannt werden. Wie in der Langform werden in der Kurzform drei Dimensionen der Maximierung erfasst, und zwar „Suche nach Alternativen“, „Entscheidungsschwierigkeit“ und „Hohe Standards“. Mit dem 9 Items umfassenden PFC-Fragebogen von Cialdini et al. (1995) wird Konsistenz als “tendency to base one’s responses to incoming stimuli on the implications of existing (prior entry) variables, such as previous expectancies, commitments and choices” (S. 318) erfasst. Auf einer 9-stufigen Skala geben die Versuchspersonen an, ob sie einzelnen Aussagen widersprechen oder zustimmen. Die einzelnen Skalenpunkte wurden folgendermaßen übersetzt: (1) ich widerspreche stark, (2) ich widerspreche, (3) ich widerspreche etwas, (4) ich widerspreche leicht, (5) weder widerspreche ich noch stimme ich zu, (6) ich stimme leicht zu, (7) ich stimme etwas zu, (8) ich stimme zu, (9) ich stimme stark zu. Das Material und die Instruktionen können Anhang A-5 entnommen werden.

Stichprobe. Es nahmen insgesamt 110 Versuchspersonen (81 weiblich, 28 männlich, 1 ohne Angabe; Durchschnittsalter: 23.64 Jahre, *SD* = 6.57), hauptsächlich Studenten, aus Bonn und Umgebung teil. Die Versuchspersonen waren über das Online-Rekrutierungssystem *ORSEE* (Greiner, 2004) angeworben worden. Das Experiment wurde als Teil einer Experimentalbatterie papierbasiert erhoben. Die Erhebung dauerte insgesamt ca. 60 Minuten, wobei etwa 15 Minuten auf das hier zu berichtende Experiment entfielen. Die Versuchspersonen erhielten als Aufwandsentschädigung im Durchschnitt 12 €.

[33] Herzlicher Dank an Gergana Y. Nenkov für die Zusendung der englischsprachigen Instruktionen.

Durchführung. Zu Beginn des Experimentes wurde den Versuchspersonen das Konzept der Cue-Validität am Beispiel der Vorhersagekraft von Orangen-Kennzeichen für die Qualität einer Orangensorte anschaulich erläutert. Daraufhin wurden die Versuchspersonen gebeten, möglichst gute und differenzierte Einschätzungen der Validitäten verschiedener Orangen-Kennzeichen für die Qualität einer Orangensorte abzugeben. Je nach dem, welcher Bedingung eine Versuchsperson zufällig zugeordnet worden war, waren zwei (Bedingung 1), vier (Bedingung 2) oder sechs (Bedingung 3) Kennzeichen einzuschätzen. Die Einschätzung der Validitäten wurde erneut auf einer Skala von 50% bis 100% vorgenommen. Nach einem thematisch unverbundenen Experiment, das etwa 15-20 Minuten dauerte, sollten sich die Versuchspersonen im anschließenden Treatment vorstellen, dass sie für die Zubereitung einer exklusiven Nachspeise Orangen einer hochwertigen Qualität benötigen und dass sie im Supermarkt die Auswahl zwischen zwei Sorten haben, die sich im Preis nicht unterscheiden. Den Versuchspersonen wurde mitgeteilt, dass sie Informationen darüber erhalten, ob Kennzeichen, die für die Qualität einer Orangensorte sprechen, vorliegen (+) oder nicht vorliegen (-). Sie wurden weiterhin darauf hingewiesen, dass sie noch keine endgültige Entscheidung treffen sollen, da sie noch die Meinung eines Mitarbeiters einholen möchten. Die Versuchspersonen sahen dann je nach Bedingung verschiedene Matrizen: In Bedingung 1 lag „sehr lange Reifeperiode“ nur bei Sorte A vor, während „optimale Lagerdauer“ nur bei Sorte B vorlag (1:1-Muster). In Bedingung 2 lag „optimale Lagerdauer“ nur bei Sorte A vor, während „kräftige/gleichmäßige Färbung“, „sehr lange Reifeperiode“ und „optimales Lagerklima“ nur bei Sorte B vorlagen (1:3-Muster). In Bedingung 3 lagen „optimale Lagerdauer“, „intensiver fruchtiger Geruch“ und „sorgfältiger Transport“ nur bei Sorte A vor, während „sehr lange Reifeperiode“, „kräftige/gleichmäßige Färbung“ und „optimales Lagerklima“ nur bei Sorte B vorlagen (3:3-Muster). Die Versuchspersonen sollten nun überlegen, von welcher Sorte sie aufgrund der vorliegenden Informationen tendenziell eher eine höhere Qualität erwarten. Im Anschluss schätzten die Versuchspersonen erneut die Validitäten der präsentierten Cues ein. Dieser Posttest war analog zum Pretest aufgebaut. Nach dem Posttest wurde dann allen Versuchspersonen mitgeteilt, dass sie die Meinung des Mitarbeiters nicht mehr einholen können und sie daher auf Basis der vorliegenden Informationen entscheiden müssen, welche Orangensorte ihrer Meinung nach eine höhere Qualität hat. Neben der Entscheidung selbst gaben die Versuchsteilnehmer auch die Sicherheit in diese Entscheidung an. Anschließend machten die Teilnehmer demografische Angaben. Zuletzt bearbeiteten die Versuchspersonen die drei Fragebögen NCCS-16, MAX

und PFC-B. Die Reihenfolge wurde individuell, aber mit dem Ziel einer Gleichverteilung randomisiert, so dass jede Reihenfolge pro Session etwa zwei Mal realisiert wurde.

ERGEBNISSE

Vorausgehende Analysen. Eine erste Analyse der Validitätseinschätzungen zeigte, dass vier Versuchspersonen stets 50% oder 100% ankreuzten, was auf ein mangelndes Verständnis der Skala zurückgeführt werden könnte. Diese vier Versuchspersonen wurden aus den folgenden Analysen ausgeschlossen.[34] Von den verbleibenden 106 (28 männlich, 77 weiblich, 1 ohne Angabe; Durchschnittsalter: 23.75 Jahre, SD = 6.65) Versuchspersonen waren 37 der ersten, 37 der zweiten und 32 der dritten Bedingung zufällig zugeteilt worden. Die deskriptiven Validitätseinschätzungen sind Tabelle B-6 in Anhang B zu entnehmen.

Umwertung subjektiver Cue-Validitäten. Eine Analyse der auf Basis der getroffenen Entscheidung berechneten Differenzen zwischen den gewichteten Summen der Optionen zum Pretest und zum Posttest zeigte, dass zwölf Versuchspersonen als Switcher zu klassifizieren sind, da sie – gemessen an den gewichteten Summen – im Pretest eine andere Option favorisierten als im Posttest. Diese Switcher wurden aus allen folgenden Analysen ausgeschlossen. Eine explorative Analyse mittels Boxplots zeigte keine extremen Ausreißer. Zur Auswertung der Veränderung der Differenzen zwischen den korrigierten gewichteten Summen der Optionen wurde eine zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor gerechnet. Der Innersubjektfaktor war „Messzeitpunkt“ mit den Stufen Differenz_pre (Differenz zwischen den gewichteten Summen im Pretest) und Differenz_post (Differenz zwischen den gewichteten Summen im Posttest). Der Zwischensubjektfaktor war „Cue-Muster“. Eine post-hoc durchgeführte Teststärkenanalyse mittels GPower (Faul et al., 2007) zeigte, dass unter Anlegen eines α-Niveaus von .05 ein Haupteffekt „Messzeitpunkt“ sowie eine Interaktion „Cue-Muster x Messzeitpunkt“ mittlerer Größe (f = .25, Cohen, 1988) mit einer Teststärke > .99 gefunden werden können. Die Teststärke zur Entdeckung des Haupteffekts „Cue-Muster“ dieser Größe beträgt aufgrund der hohen Korrelation zwischen den Stufen des Faktors „Messzeitpunkt“ lediglich .57. Der Haupteffekt „Cue-Muster“ wurde signifikant, $F(2, 91) = 23.15$, $p < .001$, $\eta_p^2 = .34$. Dies lässt sich darauf zurückführen, dass die Differenz zwischen den gewichteten Summen aufgrund der jeweiligen Informationslage unterschiedlich und in der zweiten Bedingung besonders hoch war. Während der Haupteffekt

[34] Gemäß der bisherigen Notation entspricht dies dem Anlegen eines weichen Kriteriums. Unter Anlegen eines harten Kriteriums müssen sieben Versuchspersonen ausgeschlossen werden.

„Messzeitpunkt“, der das *spreading apart* über die Bedingungen hinweg anzeigt, nicht signifikant wurde, zeigte sich eine signifikante Interaktion „Cue-Muster x Messzeitpunkt“, $F(2, 91) = 3.33$, $p < .05$, $\eta_p^2 = .07$. Diese Interaktion ist für die zugrunde liegende Fragestellung besonders interessant, da sie darauf hinweist, dass sich die Veränderung der Cue-Validitäten (gemessen am *spreading apart* der Alternativen) zwischen den Bedingungen unterscheidet.[35] Eine Inspektion der Mittelwerte zeigt, dass das *spreading apart* in der zweiten Bedingung am größten ist (siehe Abbildung 7).

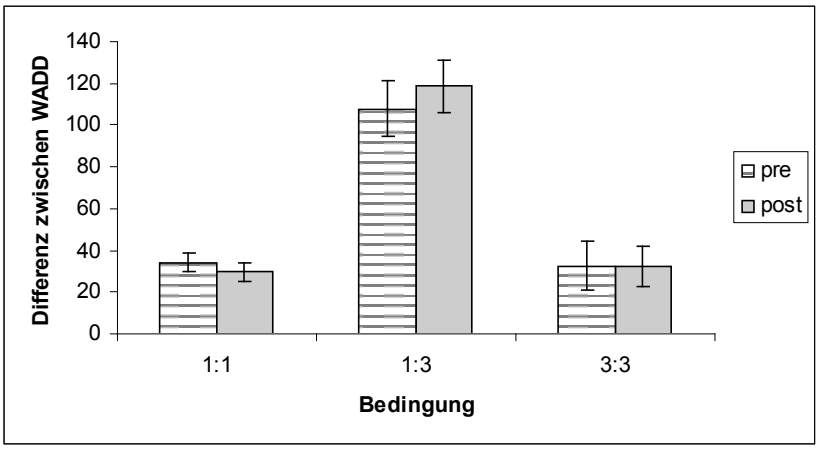


Abbildung 7: Differenz zwischen den korrigierten gewichteten Summen (WADD) der Optionen zum Pretest (pre) und zum Posttest (post). Die Bedingungsbezeichnungen ergeben sich aus der Art des präsentierten Cue-Musters. Fehlerbalken geben Standardfehler an.

Um die beobachtete Interaktion zu spezifizieren, wurden pro Bedingung gepaarte *t*-Tests (einseitig) mit den Variablen Differenz_pre und Differenz_post berechnet. Der Test wurde lediglich für die zweite Bedingung, d.h. das 1:3-Muster signifikant, $t(36) = 2.16$, $p < .05$, $d_z = .35$. Die Teststatistiken für die erste und dritte Bedingung lauten $t(31) = -1.33$, $p = .10$, $d_z = -0.24$ und $t(24) = -0.06$, $p = .48$, $d_z = -0.01$.[36]

Die Umwertung von Cue-Validitäten lässt sich nicht nur anhand des *spreading apart* von Alternativen, sondern auch anhand der Umwertung der einzelnen Cues untersuchen. Aus diesem Grund wurden die Cues auf Basis der getroffenen Entscheidung in aufzuwertende und abzuwertende Cues eingeteilt (für einen deskriptiven Überblick siehe Tabelle B-7 in

[35] Erneut wurde eine univariate ANOVA mit den mittleren Cue-Umwertungen als abhängiger Variable und „Bedingung“ als Zwischensubjektfaktor berechnet (siehe Fußnote 27). Es zeigte sich auch hier ein signifikanter Bedingungsunterschied.

[36] Legt man das harte Kriterium an, bleibt das Ergebnismuster bestehen. Lediglich die Interaktion scheitert mit $p = .07$ knapp am konventionellen α-Niveau von .05.

Anhang B).[37] Eine pro Bedingung vorgenommene explorative Analyse der Pretest- und Posttest-Einschätzungen der Cues ergab keine Extremwerte. Pro Bedingung wurden dann zweifaktorielle Varianzanalysen mit Messwiederholung auf den beiden Faktoren „Messzeitpunkt“ und „Cue“ berechnet. Die interessierende Interaktion „Messzeitpunkt x Cue“ wurde in keiner Bedingung signifikant (Tabelle 4).[38]

Tabelle 4: Interaktionseffekte der messwiederholten Varianzanalysen pro Bedingung

Bedingung	*F*	*df*	*p*	η_p^2
1 (1:1)	1.78	1, 31	.19	.05
2 (1:3)	1.48	1.44, 46.01[a]	.24	.04
3 (3:3)	1.30	5, 120	.27	.05

Anmerkungen: Die Bedingungsbezeichnungen ergeben sich aus der Art des präsentierten Cue-Musters. [a]Aufgrund einer Verletzung der Freiheitsgrade wurde eine Korrektur der Freiheitsgrade, und zwar nach Greenhouse-Geisser, vorgenommen.

Eine Umwertung ließ sich somit auf der Ebene der Cues unter Anlegen des konventionellen α-Niveaus von .05 nicht feststellen. Es gilt allerdings zu bedenken, dass eine post-hoc durchgeführte Teststärkenanalyse mittels GPower (Faul et al., 2007) ergab, dass bei α = .05 ein entsprechender Effekt mittlerer Größe (*f* = .25, Cohen, 1988) lediglich mit Teststärken von .77 (Bedingung 1, *n* = 32), .54 (Bedingung 2, *n* = 33) und .75 (Bedingung 3, *n* = 25) zu entdecken war.[39]

Der Einfluss von Persönlichkeitsvariablen. Als Persönlichkeitsvariablen, die in einem Zusammenhang mit der Umwertung subjektiver Cue-Validitäten stehen könnten, wurden der *Need for Cognitive Closure* (NCC), die Tendenz zu maximieren oder zu satisfizieren und die *Preference for Consistency* (PFC) in Betracht gezogen. Der NCC wurde mit der 16-NCC-Skala (Schlink und Walther, 2007) erfasst. Die interne Konsistenz der Gesamtskala war mit α = .80 zufriedenstellend. Die aus sechs Items bestehende Kurzform der *Maximization Scale* (Nenkov et al., 2008) wurde eingesetzt, um Maximierer von Satisfizierern zu unterscheiden. Aufgrund der Kürze der Gesamtskala und insbesondere der drei Subskalen waren die internen Konsistenzen – wie schon bei Nenkov et al. (2008) –

[37] In der zweiten Bedingung wählten vier Versuchspersonen Option A, was eine einheitliche Klassifikation in aufzuwertende und abzuwertende Cues nicht möglich macht. Folglich wurden diese Versuchspersonen aus dieser Analyse ausgeschlossen.
[38] Das gleiche Ergebnis zeigte sich auch unter Anlegen des harten Kriteriums.
[39] Für die Berechnung der Teststärke in der zweiten Bedingung wurde die Korrektur der Freiheitsgrade aufgrund einer Verletzung der Sphärizitätsannahme berücksichtigt.

weniger hoch. Für die Gesamtskala war Cronbachs α = .59, für die Subskala *Alternative Search* betrug es .34, für die Subskala *Decision Difficulty* .21 und für die Subskala *High Standards* .71. Die Präferenz für Konsistenz wurde mit dem entsprechenden Fragebogen von Cialdini et al. (1995), der PFC-B, erfasst. Eine Reliabilitätsanalyse ergab eine interne Konsistenz von α = .76. Die Mediane der einzelnen Skalen betrugen 3.19 (16-NCCS), 3.50 (*Alternative Search*), 4.50 (*Decision Difficulty*), 5.00 (*High Standards*) und 4.67 (PFC-B).

Um einen möglichen Einfluss dieser Persönlichkeitsvariablen auf die Umwertung subjektiver Cue-Validitäten zu ermitteln, wurde zunächst die mittlere Umwertung berechnet, indem die Veränderung der Differenz zwischen den korrigierten gewichteten Summen durch die doppelte Anzahl der Cues dividiert wurde. Dadurch wurde ein von der Anzahl der Cues unabhängiges Maß generiert. Es wurden dann bivariate Korrelationen zwischen dem jeweiligen Persönlichkeitsmaß und der mittleren Umwertung berechnet. Unter Zugrundelegung eines α-Niveaus von .05 (einseitige Testung) kann bei N = 94 ein mittlerer Effekt (r = .3, Cohen, 1988) mit einer Teststärke > .92 entdeckt werden. Die Ergebnisse sind Tabelle 5 zu entnehmen. Wie in der Tabelle ersichtlich ist, konnte lediglich ein signifikanter Zusammenhang mit der Präferenz für Konsistenz gefunden werden.

Tabelle 5: Korrelationen der Persönlichkeitsmaße mit der mittleren Cue-Umwertung

Skala	r	r^2	$t(92)$
NCC	.06	.00	0.55
Maximization Scale	.06	.00	0.56
- Alternative Search	-.01	.00	-0.13
- Decision Difficulty	.03	.00	0.32
- High Standards	.11	.01	1.02
PFC	.19	.04	1.88* [a]

Anmerkungen: *: $p < .05$. [a] Der Test wurde mit N = 93 durchgeführt, da eine Versuchsperson einen Großteil der Items nicht beantwortete und daher nicht berücksichtigt wurde.

Weitere Ergebnisse: Entscheidungen und Konfidenzurteile. In der ersten Bedingung entschieden sich 17 Versuchspersonen für die Option bzw. Orangensorte A und 15 für B. In der zweiten Bedingung entschied sich die Mehrheit der Versuchpersonen, nämlich 33, für Option B, während lediglich 4 Versuchspersonen Option A wählten. In der dritten Bedingung wählten 15 Versuchspersonen Option A und 10 Personen wählten Option B. Von

besonderem Interesse war die Überprüfung der Hypothese, dass die Konfidenzurteile in einem positiven Zusammenhang mit der Größe der Differenz zwischen den gewichteten Summen stehen. Eine Inspektion der mittleren Differenzen im Posttest zeigt, dass die Differenz in der zweiten Bedingung deutlich höher ist (*MW* = 118.63, *SD* = 74.44) als in den Bedingungen 1 (*MW* = 32.25, *SD* = 47.98) und 3 (*MW* = 29.69, *SD* = 23.76), die sich wiederum nur geringfügig unterscheiden. Folglich wurde ein im Vergleich höheres Konfidenzurteil für die zweite Bedingung erwartet. Um dies zu überprüfen, wurde eine univariate Varianzanalyse mit Konfidenz als abhängiger Variable und „Bedingung" als Zwischensubjektfaktor berechnet. Der Haupteffekt „Bedingung" wurde signifikant, $F(2, 91) = 4.11$, $p < .05$, $\eta^2 = .08$ (siehe Abbildung 8).

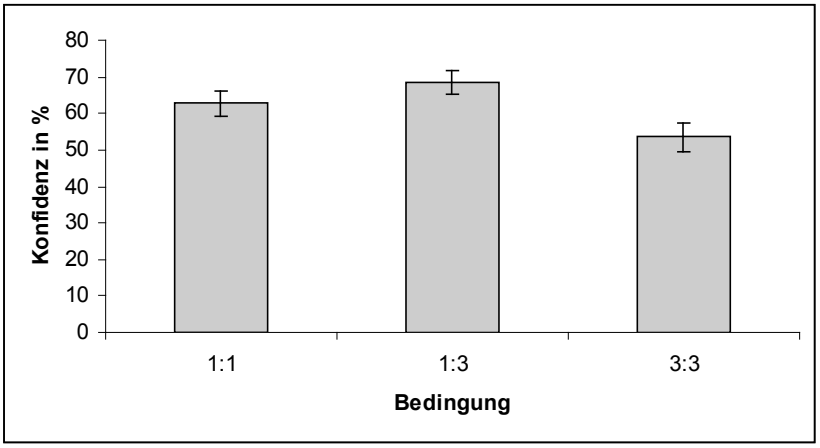


Abbildung 8: Konfidenz in die Entscheidung. Die Bedingungsbezeichnungen ergeben sich aus der Art des präsentierten Cue-Musters. Fehlerbalken geben Standardfehler an.

Eine post-hoc-Analyse mittels des Tukey HSD-Tests zeigt, dass sich nur die Bedingungen 2 und 3 signifikant voneinander unterscheiden. Die Konfidenz ist bei Versuchspersonen der zweiten Bedingung am höchsten (*MW* = 68.26, *SD* = 19.11), gefolgt von Versuchspersonen der ersten (*MW* = 62.71, *SD* = 19.65) und schließlich der dritten Bedingung (*MW* = 53.63, *SD* = 20.67). Folglich ist die Hypothese teilweise bestätigt.

DISKUSSION

Mit dem vorliegenden Experiment wurden zwei Ziele verfolgt: Erstens sollte überprüft werden, ob sich ein Effekt der Ausgeglichenheit der Informationslage auch mit anderem Material und methodischen Verbesserungen finden lässt. Zweitens sollte der

Einfluss von Persönlichkeitsvariablen auf die Umwertung subjektiver Cue-Validitäten untersucht werden.

Auch in diesem Experiment konnte ein Einfluss der Informationslage gefunden werden: So zeigte sich auf der Ebene der Optionen erneut ein signifikantes *spreading apart* der Alternativen, wenn die Versuchspersonen mit einer unausgeglichenen Informationslage konfrontiert worden waren. Auf der Ebene der Cues konnten allerdings keine signifikanten systematischen Umwertungen gefunden werden. Diese Umwertungen wären nach dem PCS-Modell aber zu erwarten gewesen. Zudem steht das Ausbleiben eines *spreading apart* bei einer ausgeglichenen Informationslage nicht in Einklang mit dem im PCS-Modell postulierten generellen Konsistenzmaximierungsprozess. Folglich konnten auf der einen Seite die Ergebnisse des vorherigen Experiments repliziert werden, wobei die Evidenz im vorliegenden Experiment allerdings schwächer ist. Auf der anderen Seite ist das Ergebnis, dass keine Umwertungen bei einer ausgeglichenen Informationslage gefunden werden konnten, nicht in Einklang mit den Annahmen des PCS-Modells. An dieser Stelle soll noch keine Diskussion dieses Aspekts vorgenommen werden, da diese unter Berücksichtigung auch der vorherigen Ergebnisse in Kapitel 3.5.4 erfolgen soll.

Bezüglich der Persönlichkeitsvariablen konnte ein Zusammenhang mit lediglich einem Faktor ermittelt werden, und zwar mit der Präferenz für Konsistenz: Eine hohe Ausprägung auf der entsprechenden Skala geht mit einer stärkeren mittleren Cue-Umwertung einher. Unter Berücksichtigung der Konsistenz-Definition von Cialdini et al. (1995), nach der Konsistenz als Tendenz, seine Antworten auf vorherige Variablen zu gründen, zu verstehen ist, lässt sich folgende vorläufige Vermutung aufstellen: Personen basieren ihre Validitätseinschätzungen im Posttest auf die mentale Repräsentation, die sie bei der Konfrontation mit der Entscheidungssituation konstruiert haben. Bei Personen mit einer hohen Präferenz für Konsistenz ist dies verstärkt der Fall. Selbstverständlich bedarf es einer Replikation dieses Effektes, bevor fundierte Interpretationen vorgenommen werden können. Bezüglich der Relevanz dieses Maßes gilt es zudem einen Hinweis von Cialdini et al. (1995) zu beachten, nach dem die Präferenz für Konsistenz nicht nur bei den Teilnehmern seiner Experimente (Studenten) relativ schwach ausgeprägt sein könnte, sondern auch in der Gesamtbevölkerung. Daher erachtet er das Finden von Konsistenzeffekten als schwierig. Zunächst bedarf es aber wie angesprochen einer Replikation.

Der von Glöckner und Betsch (2008c) formulierte positive Zusammenhang zwischen der Höhe der Differenz zwischen den gewichteten Summen der Optionen und der Höhe der

Konfidenz konnte in dem vorliegenden Experiment teilweise bestätigt werden. So war die Konfidenz in der zweiten Bedingung, in der auch die Differenz zwischen den gewichteten Summen mit Abstand am höchsten war, höher als in den anderen beiden Bedingungen. Allerdings zeigten post hoc-Analysen lediglich einen signifikanten Unterschied zwischen der zweiten und der dritten Bedingung. Hierbei muss jedoch bedacht werden, dass die Umwertungen im zweiten Muster im Vergleich zu den vorherigen Experimenten von Glöckner et al. (im Druck) nur schwach ausgeprägt waren. Es ist vor dem Hintergrund der bisherigen Forschungsbefunde (z.B. Glöckner und Betsch, 2008c; Russo et al., 2008) davon auszugehen, dass eine stärkere Umwertung – wie sie auch in vorherigen Experimenten gefunden wurde – zu einem deutlicheren Unterschied geführt hätte.

Zusammenfassend konnte erneut ein Einfluss der Ausgeglichenheit der Informationslage auf die Veränderung subjektiver Cue-Validitäten beobachtet werden. Des Weiteren erwies sich die Präferenz für Konsistenz als weiterer Einflussfaktor. Inwiefern dieses Ergebnis replizierbar ist, soll im folgenden Experiment überprüft werden.

3.5.3 Experiment 5: Charakteristika der Person und die Veränderung subjektiver Cue-Validitäten

ZIELSETZUNG

In dem vorzustellenden Experiment wurde der Einfluss personaler Faktoren auf die prädezisionale Umwertung subjektiver Cue-Validitäten untersucht. Nachdem im vorherigen Experiment ein signifikanter Einfluss der Präferenz für Konsistenz gefunden worden war, sollte in dem vorliegenden Experiment die Replizierbarkeit dieses Ergebnisses überprüft werden. Des Weiteren sollte der Einfluss weiterer Charakteristika der Person untersucht werden.

METHODE

Design. Wie in den vorherigen Experimenten wurde *within subjects* die Einschätzung der Cue-Validitäten sowohl im Pretest als auch im Posttest vorgenommen. Alle Versuchspersonen wurden mit einer Entscheidungsaufgabe konfrontiert, in der ein Cue gegen drei andere Cues sprach. Des Weiteren bearbeiteten alle Versuchspersonen drei Persönlichkeitsfragebögen, die in 3! = 6 möglichen Reihenfolgen und mit dem Ziel der Gleichverteilung ausgehändigt wurden.

Hypothesen. Gemäß dem automatischen Prozess der Konsistenzmaximierung und angesichts der Befunde aus den vorherigen Experimenten wurde erwartet, dass eine systematische Reevaluation der Cue-Validitäten stattfindet, und zwar sowohl auf der Ebene der Optionen als auch auf der Ebene der Cues. Des Weiteren wurde erwartet, dass die Stärke dieser Umwertung von der Ausprägung verschiedener Persönlichkeitsfaktoren beeinflusst wird. So sollte sich der Einfluss der PFC (Cialdini et al., 1995), der im vorherigen Experiment gefunden werden konnte, replizieren lassen. Zweitens wurden als weitere Maße, die später detailliert vorgestellt werden, der HAKEMP-90 (Kuhl, 1994) und eine Kurzform zur Erfassung der *Big Five* (Rammstedt und John, 2007) eingesetzt. Während für den HAKEMP-90 aufgrund der Befunde von Beckmann und Kuhl (Beckmann und Kuhl, 1984) die Hypothese aufgestellt werden konnte, dass handlungsorientierte Personen zu einer stärkeren Umwertung neigen als lageorientierte Personen, wurde der Einfluss der *Big Five* explorativ untersucht.

Material. Das verwendete Material entsprach im Wesentlichen demjenigen der ersten Bedingung des ersten Experiments. Die einzigen Abweichungen bestanden darin, dass das vorliegende Experiment papierbasiert erhoben wurde und dass die Versuchspersonen nach der Bearbeitung des Posttests tatsächlich eine Entscheidung trafen sowie ihre Konfidenz in diese Entscheidung angaben (das Material und die Instruktionen sind Anhang A-2 und A-6 zu entnehmen). Zur Erfassung der Persönlichkeitsmaße wurden erneut der PFC-B (Cialdini et al., 1995), der Hakemp-90 (Kuhl, 1994) sowie der BFI-10 (Rammstedt und John, 2007) eingesetzt. Der PFC-B wurde bereits in Kapitel 3.5.2 vorgestellt. Der HAKEMP-90 erfasst individuelle Unterschiede hinsichtlich der Handlungskontrolle, wobei einzelne Beurteilungen auf Basis der drei jeweils 12 Items umfassenden Skalen „Handlungsorientierung nach Mißerfolgserfahrungen“ (HOM), „Grad der Entscheidungs- und Handlungsplanung“ (HOP) und „Handlungsorientierung bei Tätigkeitsausführung“ (HOT) möglich sind (Kuhl, 1994).[40] Mit dem von Rammstedt und John (2007) entwickelten BFI-10 können die *Big Five* Extraversion, Verträglichkeit, Gewissenhaftigkeit, Neurotizismus und Offenheit für Erfahrungen in kurzer Zeit erfasst werden, was zur ökonomischen Zeitplanung in dem vorliegenden Experiment beitrug. Jede Skala umfasst 2 Items, wobei die Validität und Reliabilität der Verträglichkeits-Skala laut Autoren optional mit einem dritten Item erhöht werden können. Dieses Vorgehen wurde in dem vorliegenden Experiment gewählt. In

[40] Ich danke Julius Kuhl herzlich für die Zusendung der deutschen Übersetzung des Fragebogens.

Tabelle 6 sind einige Items der verschiedenen Fragebögen exemplarisch aufgeführt, um die zugrunde liegenden Konstrukte zu veranschaulichen.

Tabelle 6: Beispielitems der verwendeten Persönlichkeitsmaße

Fragebogen	Skala	Item	Antwortformat
PFC-B (Cialdini et al., 1995)		Es stört mich nicht sehr, wenn meine Handlungen inkonsistent/ widersprüchlich sind.[a,b]	9 Stufen „ich widerspreche stark“ bis „ich stimme stark zu“
BFI-10 (Rammstedt & John, 2007)	Extraversion	Ich bin eher zurückhaltend, reserviert.[a]	5 Stufen „trifft überhaupt nicht zu“ bis „trifft voll und ganz zu“
	Verträglichkeit	Ich bin rücksichtsvoll zu anderen, einfühlsam.	
	Gewissenhaftigkeit	Ich erledige Aufgaben gründlich.	
	Neurotizismus	Ich werde leicht nervös und unsicher.	
	Offenheit	Ich habe eine aktive Vorstellungskraft, bin phantasievoll.	
HAKEMP-90 (Kuhl, 1994)	HOM	Wenn ich nach einem Einkauf zu Hause merke, daß ich zu viel bezahlt habe, aber das Geld nicht mehr zurückbekomme, a) fällt es mir schwer, mich auf irgend etwas anderes zu konzentrieren. b) fällt es mir leicht, die Sache auszublenden.	Ankreuzen der Antwortalternative
	HOP	Wenn ich vorhabe, eine umfassende Arbeit zu erledigen, dann a) denke ich manchmal zu lange nach, womit ich anfangen soll. b) habe ich keine Probleme loszulegen.	
	HOT	Wenn ich etwas Interessantes lese, dann a) beschäftige ich mich zwischendurch zur Abwechslung auch mit anderen Dingen. b) bleibe ich oft sehr lange dabei.	

Anmerkungen: [a] Das Item ist zu invertieren. [b] Das Item wurde ins Deutsche übersetzt.

Stichprobe. An dem Experiment nahmen insgesamt 51 Studenten der Universität Erfurt (8 männlich, 43 weiblich; Durchschnittsalter: 21.53 Jahre, *SD* = 1.98) teil, die über das Rekrutierungssystem *ORSEE* (Greiner, 2004) angeworben worden waren. Das Experiment wurde papierbasiert durchgeführt und war Teil einer Experimentalbatterie. Deren Gesamtdauer betrug eine Stunde, wobei etwa 20 Minuten auf das vorliegende Experiment entfielen. Die Versuchspersonen wurden mit bis zu 14 € entlohnt.

Durchführung. Nach einer anschaulichen Erläuterung des Validitätskonzepts schätzten die Versuchspersonen die subjektive Validität der Cues, d.h. der Wetterberichte von „www.wetteronline.de“, „Sat.1“, „Bild“ und „ZDF“, ein. Im Anschluss an ein thematisch

unverbundenes Experiment, das etwa 15 Minuten dauerte, erfolgte das Treatment, wobei als Coverstory wieder die Planung einer Urlaubsreise diente. Die Versuchspersonen wurden mit einer Entscheidungsaufgabe konfrontiert, in der sie darüber nachdenken sollten, an welchem von zwei Ferienorten sie auf Basis der Aussagen der verschiedenen Wetterberichte eher sonniges Wetter erwarten. „Sat.1“ sagte sonniges Wetter lediglich für Ferienort A vorher, während die übrigen Cues „www.wetteronline.de“, „Bild“ und „ZDF“ sonniges Wetter nur für Ferienort B vorhersagten. Die endgültige Entscheidung wurde mit dem Hinweis, dass die Versuchspersonen noch eine weitere Quelle (den Wetterbericht der ARD am Ende der Tagesthemen) berücksichtigen wollen, hinausgezögert. Nach einer erneuten Einschätzung der subjektiven Validitäten wurde den Versuchspersonen mitgeteilt, dass sie doch ohne die Information der zusätzlichen Quelle eine Entscheidung treffen müssten. Zudem wurden sie um eine Einschätzung der Konfidenz in ihre Entscheidung sowie um demografische Angaben gebeten. Anschließend bearbeiteten die Versuchspersonen die drei Persönlichkeitsfragebögen.

Ergebnisse

Vorausgehende Analysen. Keine der Versuchspersonen hat durchgängig in einem der Tests oder in beiden Tests 50% oder 100% angekreuzt. Eine einzige Person hat sich dafür entschieden, dass das Wetter eher an Ferienort A sonnig sein wird. Diese Versuchsperson wurde aus den nachfolgenden Analysen ausgeschlossen, da sie auf der Ebene der Optionen einen extremen Ausreißer darstellt und auf der Ebene der Cues eine im Vergleich zu den übrigen Personen entgegengesetzte Auf- und Abwertung zu erwarten ist. Folglich verblieben 50 Versuchspersonen (8 männlich, 42 weiblich; Durchschnittsalter: 21.52 Jahre, $SD = 2.00$). Die mittleren Validitätseinschätzungen sind Tabelle B-8 in Anhang B zu entnehmen.

Veränderung der Cue-Validitäten. Die Umwertung wurde wieder auf der Ebene der Optionen und auf der Ebene der Cues geprüft. Für die Analyse auf der Ebene der Optionen wurden zunächst die korrigierten gewichteten Summen der Optionen zum Pretest und zum Posttest berechnet, woraufhin auf Basis der jeweils getroffenen Entscheidung die Differenz zwischen den korrigierten gewichteten Summen zu beiden Messzeitpunkten ermittelt wurde. Eine explorative Analyse der Differenzwerte anhand von Boxplots offenbarte keinen extremen Ausreißer. In einer einfaktoriellen Varianzanalyse mit Messwiederholung (Innersubjektfaktor „Messzeitpunkt“) wurde die Veränderung dieser Differenzen untersucht. Eine post-hoc durchgeführte Teststärkenanalyse mittels GPower (Faul et al., 2007) zeigte,

dass unter Anlegen eines α-Niveaus von .05 mit $N = 50$ ein mittlerer Effekt ($f = .25$, Cohen, 1988) mit einer Teststärke, die höher als .99 ist, gefunden werden kann. Der Haupteffekt „Messzeitpunkt“ wurde signifikant, $F(1, 49) = 10.49$, $p < .01$, $\eta_p^2 = .18$, wobei die Posttest-Differenzen wie erwartet signifikant höher als die Pretest-Differenzen waren. Es hat also ein *spreading apart* der Alternativen stattgefunden. Die Differenz zwischen den korrigierten gewichteten Summen im Posttest korrelierte mit $r = .25$ positiv und signifikant mit der Konfidenz in die Entscheidung, $p < .05$ (einseitig). Die mittlere Konfidenz betrug 86.28 ($SD = 7.56$).

Für die Prüfung auf der Ebene der Cues wurden die einzelnen Cue-Validitäten zunächst einer explorativen Analyse anhand von Boxplots unterzogen. Es konnten keine extremen Ausreißer ermittelt werden. Es wurde eine zweifaktorielle messwiederholte Varianzanalyse mit Messwiederholung auf den beiden Faktoren „Messzeitpunkt“ und „Cue“ berechnet. Der Haupteffekt „Messzeitpunkt“ wurde nicht signifikant, $F(1, 49) = 1.40$, $p = .24$, $\eta_p^2 = .03$, was mit der gleichzeitigen Auf- und Abwertung der unterschiedlichen Cues zu begründen ist und damit zu erwarten war. Ebenfalls erwartet war der signifikante Effekt des Faktors „Cue“, $F(2.40, 117.54) = 56.08$, $p < .001$, $\eta_p^2 = .53$, der aus der unterschiedlichen Einschätzung der einzelnen Cues resultiert.[41] Für die vorliegende Fragestellung war die Interaktion „Messzeitpunkt x Cue“ interessant, die ebenfalls signifikant wurde, $F(2.08, 101.68) = 6.62$, $p < .01$, $\eta_p^2 = .12$.[42] Wie eine post-hoc durchgeführte Teststärkenanalyse (Faul et al., 2007) ergab, kann bei $N = 50$ und unter Anlegen des konventionellen α-Niveaus von .05 ein mittlerer Effekt ($f = .25$, Cohen, 1988) mit einer Teststärke von .95 gefunden werden.[43] Um diesen Effekt näher zu untersuchen, wurden für jeden Cue gepaarte *t*-Tests (einseitig) berechnet. Für lediglich einen Cue wurde der gepaarte *t*-Test signifikant, und zwar für den Cue „Bild“, $t(49) = 2.72$, $p < .01$, $d_z = .38$, der signifikant aufgewertet wurde. Die Abwertung des Cues „Sat.1“ scheiterte knapp am konventionellen Signifikanzniveau, $t(49) = -1.70$, $p = .05$, $d_z = -.24$. Obgleich alle Cues mindestens tendenziell systematisch umgewertet wurden (siehe Abbildung 9), konnte lediglich für einen Cue eine signifikante Umwertung nachgewiesen werden.

[41] Die Freiheitsgrade wurden nach Huynh-Feldt korrigiert, da $\varepsilon > .75$.
[42] Die Freiheitsgrade wurden nach Greenhouse-Geisser korrigiert, da $\varepsilon < .75$.
[43] Die Verletzung der Sphärizitätsannahme wurde bei der Berechnung berücksichtigt.

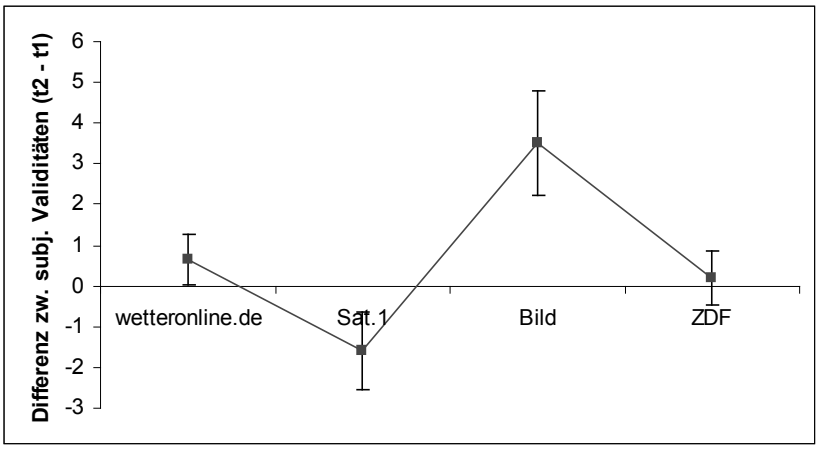


Abbildung 9: Differenz der subjektiven Cue-Validitäten zwischen den beiden Messzeitpunkten t1 (Pretest) und t2 (Posttest). Die Einschätzung des Pretests wurde von der des Posttests subtrahiert. Werte > 0 zeigen eine Auf-, Werte < 0 eine Abwertung an. Fehlerbalken geben Standardfehler an.

Zusammenhang mit Persönlichkeitsmaßen. Aufgrund bisheriger Ergebnisse und bestehender Befunde wurden Einflüsse der Präferenz für Konsistenz und der Handlungsorientierung, hierbei insbesondere der Handlungsorientierung bei der Handlungsplanung, erwartet. Explorativ wurde der Einfluss der *Big Five* untersucht. Tabelle 7 gibt einen Überblick über die internen Konsistenzen (Werte aus Originalarbeiten in Klammern), die Mediane und die Itemanzahl der einzelnen Skalen.

Tabelle 7: Überblick über interne Konsistenzen, Mediane und Anzahl der Items der verwendeten Persönlichkeitsmaße

Fragebogen	Skala	Cronbachs α	Median	Anzahl an Items
PFC-B		.76 (.84)	5	9
HAKEMP-90	HOM	.63 (.70)	3	12
	HOP	.59 (.78)	4	12
	HOT	.44 (.74)	10	12
BFI-10	Extraversion		4	2
	Verträglichkeit		3.5	3
	Gewissenhaftigkeit		3.5	2
	Neurotizismus		3	2
	Offenheit		4.5	2

Anmerkungen: Die in den Originalarbeiten angegebenen Cronbachs α-Werte stehen, sofern vorhanden, in Klammern. Sie wurden für den PFC-B Cialdini et al. (1995), für den HAKEMP-90 Kuhl (1994) entnommen.

Für die BFI-10 wurden aufgrund der Kürze der jeweiligen Skalen keine internen Konsistenzen berechnet. Die Mediane für die Subskalen HOM und HOP sind relativ niedrig, die übrigen liegen meist um den Skalenmittelpunkt. Auffällig sind die relativ niedrigen internen Konsistenzen der Subskalen HOM, HOP und HOT, die deutlich unter den von Kuhl (1994) ermittelten Werten liegen.

Um den Einfluss der einzelnen Faktoren zu ermitteln, wurden wieder bivariate Korrelationen berechnet. Zunächst wurde wie im vorherigen Experiment die mittlere Veränderung der Cues berechnet, mit der die Ausprägung des jeweiligen Persönlichkeitsmaßes korreliert wurde. Eine Berechnung der Teststärke mittels GPower (Faul et al., 2007) ergab, dass mit $N = 50$ unter Anlegen eines α-Niveaus von .05 und zweiseitiger Testung mittlere Effekte ($r = .3$, Cohen, 1988) mit einer Teststärke von .59 gefunden werden können. Durch eine Anhebung des α-Niveaus auf .10 konnte die Teststärke auf .71 erhöht werden. In Tabelle 8 sind sämtliche Ergebnisse zusammengefasst.

Tabelle 8: Korrelationen der Persönlichkeitsmaße mit der mittleren Cue-Umwertung

Skala	r	r^2	$t(48)$
PFC	.11	.01	0.75
HOM	-.05	.00	-0.36
HOP	-.01	.00	-0.10
HOT	-.24	.06	-1.75^{+}
Extraversion	.18	.03	1.29
Verträglichkeit	.29	.08	2.10*
Gewissenhaftigkeit	.18	.03	1.23
Neurotizismus	.04	.00	0.27
Offenheit	.12	.02	0.86

Anmerkungen: *: $p < .05$, $^{+}$: $p < .10$.

Wie der Tabelle zu entnehmen ist, bestehen lediglich zwischen den Persönlichkeitsmaßen „HOT“ und „Verträglichkeit“ signifikante Zusammenhänge mit der mittleren Cue-Umwertung. Je stärker die Handlungsorientierung bei der Tätigkeitsausführung ausgeprägt ist, desto schwächer ist die Umwertung. Außerdem geht ein hohes Maß an Verträglichkeit mit einer stärkeren Umwertung einher. Folglich konnte die Hypothese, dass die Präferenz für Konsistenz und die Handlungsorientierung bei der Handlungsplanung in *positivem*

Zusammenhang mit der Stärke der prädezisionalen Umwertung stehen, nicht bestätigt werden. Vielmehr zeigten sich ein *negativer* Zusammenhang mit einer anderen Subskala der Handlungsorientierung sowie ein positiver Zusammenhang mit dem Ausmaß an Verträglichkeit.

Diskussion

Mit dem vorliegenden Experiment wurde das Ziel verfolgt, den im vorherigen Experiment gefundenen Einfluss der Präferenz für Konsistenz zu replizieren und den Einfluss weiterer Charakteristika der Person zu untersuchen. Folglich wurden als Persönlichkeitsfaktoren, für die ein Einfluss auf die Stärke der prädezisionalen Umwertung angenommen wurde, die Präferenz für Konsistenz, die Handlungsorientierung sowie die *Big Five* erfasst. Voraussetzung für die Untersuchung eines Einflusses personaler Faktoren ist das Auftreten prädezisionaler Umwertungen. Sowohl auf der Ebene der Optionen als auch auf der Ebene der Cues konnte der entsprechende Nachweis erbracht werden. Allerdings ist zu berücksichtigen, dass auf der Ebene der Cues zwar alle Cues systematisch verändert wurden, diese Veränderung aber lediglich für einen Cue signifikant war. Folglich ist die Evidenz schwächer als nach dem PCS-Modell zu erwarten gewesen wäre. Die Entscheidung ging mit einer relativ hohen – insbesondere im Vergleich zum vorherigen Experiment – Konfidenz einher. Während diese in der vergleichbaren Bedingung 2 von Experiment 4 durchschnittlich lediglich 68.26 (SD = 19.11) betrug, lag sie im vorliegenden Experiment durchschnittlich bei 86.28 (SD = 7.56). Dies ist nicht durch eine a priori höhere Differenz zwischen den korrigierten gewichteten Summen im Pretest zu erklären, da diese in der Bedingung 2 von Experiment 4 mit 130.57 (SD = 47.27) im Vergleich zu dem im vorliegenden Experiment erzielten Wert von 121.84 (SD = 34.94) sogar höher war. Vielmehr scheint die stärkere Umwertung im vorliegenden Experiment zu einer vergleichsweise höheren Konfidenz beigetragen zu haben. Obgleich diese Interpretation mit Vorsicht zu betrachten ist, da sie auf einem deskriptiven Vergleich zwischen zwei Experimenten basiert und daher nicht vertieft wird, gibt sie doch einen wichtigen Hinweis auf den Zusammenhang zwischen der Umwertung und der Erhöhung der Konfidenz.

Ob Persönlichkeitsfaktoren einen Einfluss auf die Stärke der prädezisionalen Informationsumwertung haben, wurde anhand von Korrelationen analysiert. Entgegen der Erwartung konnte der Zusammenhang zwischen der mittleren Cue-Umwertung und der Präferenz für Konsistenz nicht repliziert werden. Lediglich eine Facette der Handlungsorien-

tierung, nämlich die Handlungsorientierung bei der Tätigkeitsausführung, sowie ein *Big Five*-Faktor, nämlich die Verträglichkeit, wiesen signifikante Korrelationen mit dieser Variable auf. Gemäß der Befunde von Beckmann und Kuhl (1984) wäre am ehesten ein Zusammenhang mit der Handlungsorientierung bei der Handlungsplanung zu erwarten gewesen, da sie diese Facette als entscheidungsrelevant erachten. Die Skala HOT hatten sie selbst in ihrem Experiment nicht berücksichtigt. Nach Kuhl (ohne Jahresangabe) erfasst diese Skala „das Ausmaß, in dem eine Person in einer Tätigkeit ‚aufgeht', ohne daß die Aufmerksamkeit von der Tätigkeitsausführung abgelenkt wird" (S. 2). Folglich besteht kein direkter Bezug zu der im vorliegenden Experiment verwendeten Entscheidungsaufgabe. Die Feststellung, dass hohe Kennwerte u.a. mit einer hohen Ambiguitätstoleranz einhergehen, könnte für die vorliegende Fragestellung zwar relevant sein. Übertragen auf das dargestellte Experiment besteht die Möglichkeit, dass eine niedrige Ambiguitätstoleranz mit einer höheren Informationsumwertung einhergeht. Doch ist diese Interpretation angesichts des angesprochenen fehlenden Bezugs nicht überzeugend. Ebenfalls mit Vorsicht zu interpretieren sind die Ergebnisse bezüglich des *Big Five*-Faktors Verträglichkeit. Zu bedenken ist nämlich, dass die entsprechende Skala aus lediglich drei Items bestand, die zudem nicht signifikant interkorrelierten. Folglich bestehen begründete Zweifel an der Aussagekraft des Skalenwertes.

Es stellt sich nun die Frage, weshalb keine beständigen Einflüsse von Charakteristika der Person nachgewiesen werden können. Nahe liegend ist, dass diese nicht (wesentlich) auf den Prozess der Konsistenzmaximierung einwirken. Diese Vermutung gilt es in der umfassenderen Diskussion in Kapitel 3.5.4, in der auch die Ergebnisse der vorherigen Experimente berücksichtigt werden, zu erörtern. Zwei Einschränkungen, die in dem vorliegenden Experiment zu bedenken sind, seien auch angesprochen. Erstens waren die internen Konsistenzen der meisten betrachteten Skalen kleiner als .70 und damit noch nicht einmal befriedigend. Selbstverständlich konnten aufgrund der Kürze der meisten Skalen keine hohen α-Koeffizienten erwartet werden, allerdings liegen die Konsistenzwerte der Skalen PFC-B, HOM, HOP und HOT auch unter den ursprünglich erhaltenen Werten. Zudem wird der Standpunkt Schmitts (1996) vertreten, nach dem die Kürze eines Tests zwar dessen geringe Reliabilität erklären kann, dass diese aber dennoch bei der Interpretation zu berücksichtigen ist. Zweitens besteht die Möglichkeit, dass extremere Ausprägungen der einzelnen Personcharakteristika durchaus einen Einfluss haben. Diese Hypothese bedarf natürlich der weiteren empirischen Testung.

3.5.4 Zusammenfassung und Diskussion

Im Folgenden sollen die bisherigen Ergebnisse zusammengefasst und diskutiert werden, wobei auf die zentralen Aspekte dieses Kapitels, nämlich den Einfluss des Informationsmusters und von Charakteristika der Person, fokussiert wird. Eine übergreifendere und umfassendere Diskussion findet sich im Abschlusskapitel 4.

Ausgehend von der Beobachtung, dass keine systematische Umwertung von Cue-Validitäten bei einer ausgeglichenen Informationslage erfolgte (Experiment 2), wurde der Einfluss der Ausgeglichenheit des Informationsmusters in zwei Experimenten (Experimente 3 und 4) untersucht. Gemäß dem im PCS-Modell für probabilistische Inferenzentscheidungen postulierten Prozess der automatischen Konsistenzmaximierung wurde angenommen, dass unabhängig von der Ausgeglichenheit der Informationslage eine systematische Umwertung von Cue-Validitäten stattfindet. Angesichts der Befunde des zweiten Experiments wurde des Weiteren erwartet, dass die Stärke dieser Umwertung von der Ausgeglichenheit der Informationslage moderiert wird: So wurde eine stärkere Umwertung bei einer unausgeglichenen Informationslage angenommen. Das Ausbleiben einer Umwertung bei einer ausgeglichenen Informationslage, wie sie in Experiment 2 gefunden wurde, wurde auch in den Experimenten 3 und 4 beobachtet, und zwar unabhängig von der Anzahl der Cues. Zwar konnte in Experiment 3 auf Cue-Ebene für das dritte Muster eine signifikante Interaktion „Messzeitpunkt x Cue“ nachgewiesen werden, jedoch wurde lediglich ein einziger Cue signifikant abgewertet. Auch für das erste Muster wurde die signifikante Umwertung nur eines Cues beobachtet. Im Gegensatz dazu wurden bei einer unausgeglichenen Informationslage überzeugende signifikante Umwertungen sowohl auf der Optionen- als auch auf der Cue-Ebene beobachtet. Zudem konnte die erwähnte Interaktion für das dritte Muster in Experiment 4 nicht repliziert werden und es fanden keine signifikanten und systematischen Umwertungen bei ausgeglichenen Informationslagen statt. Allerdings waren in diesem Experiment auch die Umwertungen bei einer unausgeglichenen Informationslage weniger überzeugend, da selbige auf der Cue-Ebene nicht signifikant waren. Die Ergebnisse auf der Optionen-Ebene sprechen allerdings dafür, dass nur bei einer unausgeglichenen Informationslage deutliche systematische Umwertungen auftreten. In den Experimenten 4 und 5 wurden Charakteristika der Person hinsichtlich ihres potenziellen Einflusses auf die Stärke der Umwertung subjektiver Cue-Validitäten untersucht. Während sich in Experiment 4 ein Einfluss der Präferenz für Konsistenz beobachten ließ, konnte dieser in Experiment 5 nicht repliziert werden. In Experiment 5 fanden sich signifikante

Zusammenhänge mit der Handlungsorientierung bei der Tätigkeitsausführung sowie mit der Verträglichkeit. Da diese Ergebnisse nicht vorherzusagen waren und zum Teil nicht intuitiv sind, sind sie eher mit Vorsicht zu interpretieren. Personcharakteristika wie der *Need for Cognitive Closure*, die Tendenz zu maximieren sowie die weiteren Facetten der Handlungsorientierung und die übrigen Faktoren der *Big Five* wiesen ebenfalls keinen nennenswerten Zusammenhang mit der Stärke der prädezisionalen Informationsumwertung auf. Weitere Ergebnisse zeigten, dass weder die Reliabilität der Einschätzung noch die Häufigkeit der Erfahrung einen Einfluss auf die Stärke der Verzerrungen haben. Zudem war die Konfidenz nach einer Entscheidung mit einem unausgeglichenen Informationsmuster höher als nach einer Entscheidung mit einem ausgeglichenen Informationsmuster. Ein deskriptiver Vergleich zwischen zwei Experimenten zeigte darüber hinaus, dass eine stärkere Umwertung mit einer höheren Konfidenz einhergeht. Die zusammenfassende Betrachtung der Ergebnisse weist auf vier Aspekte hin, die es zu diskutieren gilt: (1) Weder die zum Teil schwache Evidenz für systematische Umwertungen bei unausgeglichenen Informationslagen noch das Ausbleiben selbiger bei ausgeglichenen Informationslagen waren nach dem PCS-Modell für probabilistische Inferenzen zu erwarten. (2) Systematische und signifikante Umwertungen sind bei einer unausgeglichenen im Gegensatz zu einer ausgeglichenen Informationslage deutlich zu beobachten. (3) Es konnten keine personalen Faktoren ermittelt werden, die einen konstanten Einfluss auf die Veränderung subjektiver Cue-Validitäten haben. (4) Die Reliabilität der Validitätseinschätzungen und die Häufigkeit der Erfahrung mit Cues stehen nicht in Zusammenhang mit der Veränderung von Cue-Validitäten. Eine umfassendere Diskussion der Ergebnisse vor dem Hintergrund der theoretischen Ausführungen des Kapitels 2 erfolgt im abschließenden Kapitel 4.

Gemäß dem PCS-Modell für probabilistische Inferenzen von Glöckner und Betsch (2008b) wären Umwertungen bei jedem Cue zu erwarten gewesen. Es besteht aber die Möglichkeit, dass zwar Umwertungen erfolgt sind, dass aber die Effekte zu klein sind, um mit den verwendeten Methoden ermittelt zu werden. Damit wird bereits die Frage der Bedeutung von Konsistenzmaximierungsprozessen im Allgemeinen und in probabilistischen Inferenzentscheidungen im Speziellen angesprochen, die in Kapitel 4 thematisiert werden soll. Zudem ist zu bedenken, dass ein stabiler Zustand auch in einem lokalen Optimum münden kann (Read et al., 1997), was nicht die Umwertung jedes einzelnen Cues voraussetzen würde. Dass bei ausgeglichenen Informationslagen eine (deutliche) Umwertung ausbleibt, ist jedoch schwerer zu erklären. Nicht angezweifelt werden soll und kann die

Allgegenwärtigkeit des Phänomens der automatischen Konsistenzmaximierung. Es besteht auch in diesem Fall die Möglichkeit, dass die Effekte schlichtweg zu klein sind. Es kann aber nicht ausgeschlossen werden, dass eine erfolgreiche Konsistenzmaximierung gewisser Voraussetzungen, die in den entsprechenden Bedingungen eventuell nicht gegeben waren, bedarf. Bevor weitere Vermutungen angestellt werden erscheint es ratsam zu überprüfen, welche Vorhersagen das PCS-Modell für die gewonnenen Daten machen würde. Dies erfolgt in Kapitel 3.6 anhand von Simulationsstudien. Abschließend ist aber festzuhalten, dass die Befunde deutlich der Annahme unidirektionaler Schlussfolgerungen von Cues auf Optionen, wie sie etwa von Gigerenzer et al. (1999) vertreten wird, widersprechen.

Dass systematische Umwertungen bei einer unausgeglichenen Informationslage im Gegensatz zu einer ausgeglichenen Informationslage deutlich zu beobachten sind, ist vor dem Hintergrund der bisherigen Forschungsergebnisse überraschend, wenn man ausgeglichen als ambigue interpretiert. Wie in Kapitel 2.4 dargestellt, wird Ambiguität als Moderatorvariable diskutiert: Bei höherer Ambiguität werden stärkere Umwertungen erwartet. Obgleich Ambiguität meist hinsichtlich der Widersprüchlichkeit der Informationen erörtert wird, wurde sie auch hinsichtlich der Informationslage untersucht. Russo et al. (1998) fanden, dass eine balancierte im Gegensatz zu einer unbalancierten Informationslage zu einer stärkeren Informationsumwertung führt. Obgleich sie in diesem Experiment nicht die Verteilung der Cues, sondern deren Inhalt variiert haben, erscheint es doch nicht plausibel, weshalb sich dieser Effekt bei einer Variation der Verteilung der Cues umkehren sollte. Folglich scheint eine andere Erklärung heranzuziehen zu sein, wobei zwei plausible Ansätze, die sich nicht ausschließen, erörtert werden sollen. Denkbar ist etwa, dass der kognitive Aufwand, der mit einer Umwertung von Cues in einer ausgeglichenen Informationslage verbunden ist, höher ist als in einer unausgeglichenen Informationslage. Ähnlich argumentiert Brownstein (2003), wenn er festhält “that decision makers may be less likely to conduct biased processing when they would have to distort information to an unreasonably high degree” (S. 561). In einer weiteren Arbeit (Brownstein et al., 2004) diskutiert er direkt den kognitiven Aufwand als Erklärungsmöglichkeit für sein Ergebnis, dass eher die gewählte Alternative aufgewertet als die nicht gewählte Alternative abgewertet wird. Bezogen auf die vorliegenden Ergebnisse könnte der kognitive Aufwand insofern eine Rolle spielen, als dass in einer unausgeglichenen Informationslage die Richtung quasi vorgegeben ist, was den Prozess erleichtern könnte. Da aber bereits Umwertungen bei ausgeglichenen Informationslagen in anderen Entscheidungsaufgaben gefunden wurden (z.B. Holyoak und Simon, 1999),

könnte auch der Aufgabentyp relevant sein. An dieser Stelle könnte die Ambiguität doch relevant sein, aber in anderer Hinsicht als oben angesprochen: Die Informationen in den verwendeten probabilistischen Inferenzaufgaben sind im Vergleich zu anderen Aufgaben wenig ambigue, weshalb der kognitive Aufwand, der mit der Umwertung von Informationen verbunden ist, hier stärker zum Tragen kommen könnte. Somit scheinen sowohl die Widersprüchlichkeit der Informationslage als auch die Widersprüchlichkeit der Informationen von Relevanz zu sein. Eine weitere mögliche interessante Erklärung betrifft die Rolle der Konfidenz. Den positiven Zusammenhang, der zwischen der Differenz zwischen den korrigierten gewichteten Summen und der Höhe der Konfidenz besteht, konnten Glöckner und Betsch (2008c) nachweisen und er wurde auch in einem der dargestellten Experimente (Experiment 4) bestätigt. Einen kausalen Zusammenhang zwischen der Höhe der Konfidenz und der Höhe der Informationsumwertung formulierten Russo et al. (2008): "The phenomenon of ID [Information Distortion] is systematic, increasing linearly with the confidence in the leading alternative (Russo et al., 1998, 2000)" (S. 457). Obgleich diese Interpretation nicht ohne weiteres auf die vorliegenden Ergebnisse übertragbar ist, da die Konfidenz erst nach der Entscheidung erfasst wurde, stellt sie doch einen potenziellen und plausiblen Erklärungsansatz dar, der in weiteren Untersuchungen zu prüfen ist. Konfidenz in die *leading option* könnte eine Voraussetzung für das Auftreten erfolgreicher Konsistenzmaximierung sein, besonders wenn der Interpretationsspielraum der Informationen aufgrund geringer Ambiguität beschränkt ist. Diese verschiedenen Erklärungsansätze stellen potenzielle Begründungen für die gefundenen Ergebnisse dar und schließen sich gegenseitig nicht aus. Im Rahmen der abschließenden Diskussion in Kapitel 4 werden diese unter Berücksichtigung weiterer Ergebnisse wieder aufgegriffen und Möglichkeiten vorgeschlagen, die Erklärungskraft dieser Ansätze zu überprüfen.

Bezüglich der Charakteristika der Person konnte kein beständiger Einfluss eines Faktors nachgewiesen werden. Zwei Erklärungen sind hierfür denkbar: Erstens könnte es sein, dass Persönlichkeitsfaktoren tatsächlich zu vernachlässigen sind. Eventuell haben sie erst dann einen Einfluss, wenn sie extrem ausgeprägt sind, was zu überprüfen ist. Zweitens besteht durchaus die Möglichkeit, dass die in der bisherigen Forschung verwendeten Maße unzureichend sind. Spezifischere Maße könnten hier Abhilfe schaffen. Allerdings ist es dann überraschend, dass bisher keinerlei systematische Zusammenhänge mit verschiedensten Maßen gefunden werden konnten. Somit sollte der mögliche Einfluss individueller

Unterschiede bedacht werden, doch scheint er nicht wesentlich für den basalen Prozess der Konsistenzmaximierung zu sein.

Dass weder eine niedrige Reliabilität einer Validitätseinschätzung noch seltene Erfahrungen mit einem Cue das Auftreten prädezisionaler Umwertungen in probabilistischen Inferenzentscheidungen erklären können, deutet ebenfalls darauf hin, dass es sich bei der Konsistenzmaximierung um einen grundlegenden Prozess handelt. Der Inhalt der Cues selbst ist natürlich von Bedeutung, da er die Einschätzung der Validität determiniert. In diesem Fall ist eher die Art der Erfahrung von Bedeutung. Dass besonders glaubwürdige Cues wie „Stiftung Warentest“ kaum verändert werden (vgl. Glöckner, 2006) und besonders unglaubwürdige Cues in einer bestimmten Konstellation deutlich umgewertet werden (vgl. Experiment 3 und 4), führt zu der Frage, wann eine bestimmte Konstellation eine besonders positive oder negative Einschätzung eines Cues überlagern und zu Umwertungen führen kann. Dieser Aspekt war kein Bestandteil der Zielsetzung dieser Experimente, doch bietet er interessante Ansatzpunkte für künftige Forschungsarbeiten.

Eine offene Frage ist das Auftreten sogenannter Switcher. Dies betraf, berücksichtigt man auch die Experimente 1 und 2, die Experimente 2, 3 und 4. In den Experimenten 2 und 3 wurden als Switcher diejenigen Personen klassifiziert, bei denen gemäß den korrigierten gewichteten Summen der Optionen im Pretest eine andere Entscheidung zu erwarten war als im Posttest. Bei diesen Personen konnte folglich keine überzeugende Erwartung hinsichtlich der Entscheidung gebildet werden, weshalb sie aus den Analysen ausgeschlossen wurden. Dadurch wurden auch Verzerrungen durch vergleichsweise extreme und nicht eindeutig zu erklärende Reevaluationen vermieden. In Experiment 4 hatten die Versuchspersonen eine Entscheidung getroffen, auf deren Basis die Differenz zwischen den korrigierten gewichteten Summen für den Pre- und für den Posttest gebildet werden konnte. Auch in diesem Fall traten Switcher auf, da sich die Relation der Optionen vom Pre- zum Posttest bei einigen Personen umkehrte. Dies weist erneut auf vergleichsweise extreme Umwertungen hin, deren Ursache unklar ist, so dass diese Versuchspersonen nicht berücksichtigt wurden. Interessanterweise waren in den meisten Fällen (10 von 12) systematische Umwertungen in Richtung der späteren Entscheidung zu beobachten. Es mag nun eingewendet werden, dass hiervon Personen, die in einer Bedingung mit einem ausgeglichenen Informationsmuster waren, mit höherer Wahrscheinlichkeit betroffen sind. Hierzu ist zu entgegnen, dass selbst die Berücksichtigung dieser Switcher das Ergebnismuster nicht veränderte. Der Begriff Switcher wurde in Anlehnung an Simon und Kollegen (Holyoak und Simon, 1999; Simon et al.,

2001; Simon, Snow et al., 2004) gewählt, wobei sie hiermit einen Wechsel von der tendenziellen zur finalen Entscheidung bezeichnen. In beiden Fällen ist der Grund für die Umkehrung unklar. Letztlich können keine begründeten Erklärungen gegeben werden, weshalb es zu diesen Umkehrungen gekommen ist. Dies liegt natürlich auch daran, dass eine gezielte Untersuchung dieser Beobachtung nicht Ziel der vorliegenden Arbeit war. Damit bleibt es eine Aufgabe für künftige Arbeiten, zu ergründen, weshalb teilweise Umkehrungen auftreten.

Zusammenfassend ist festzuhalten, dass Umwertungen in probabilistischen Inferenzentscheidungen dann deutlich auftreten, wenn eine unausgeglichene Informationslage vorliegt. Eine mögliche Begründung ist, dass in diesem Fall die Umwertung weniger kognitiv aufwändig ist. Charakteristika der Person scheinen keinen wesentlichen Einfluss zu haben, was die Interpretation des Konsistenzmaximierungsprozesses als basalen Mechanismus unterstreicht. Dass allerdings keine signifikanten Umwertungen bei ausgeglichenen Informationslagen zu beobachten waren, ist unter dieser Perspektive überraschend. Folglich soll in Simulationsstudien geklärt werden, wie gut die Ergebnisse mittels PCS vorhergesagt werden können.

3.6 SIMULATIONSSTUDIEN

3.6.1 Methodische Vorüberlegungen

In der PCS-Forschung wurde und wird eine Vielzahl an Simulationsstudien durchgeführt. Diese können verschiedenen Zwecken dienen, etwa der Simulation bekannter Phänomene (z.B. Read und Miller, 1994), der Generierung von Hypothesen (z.B. Glöckner, 2006) und der Simulation eigener Daten (z.B. Spellman und Holyoak, 1992). In den vorzustellenden Simulationsstudien wurde die dritte Zielsetzung verfolgt: Unter Verwendung der Daten aus den Experimenten 3, 4 und 5 wurden Simulationsstudien durchgeführt, um zu überprüfen, wie gut das PCS-Modell die empirischen Ergebnisse vorhersagen kann. In den vorzustellenden Simulationsstudien wurde für jeden Datensatz ein analoges Vorgehen gewählt, das im vorliegenden Kapitel einführend erläutert werden soll. Hierbei wird zunächst die Zielsetzung präzisiert, bevor das Berechnen der Simulationen und die Auswertungsmethode detailliert geschildert werden.

Zielsetzung

Das wesentliche Ziel der im Folgenden darzustellenden Simulationsstudien bestand darin zu überprüfen, wie gut die empirischen Ergebnisse mit dem PCS-Modell vorhergesagt werden können. Hierbei wurden individuelle Simulationen pro Versuchsperson berechnet. Darüber hinaus wurde jede Simulation noch einmal unter Modifikation von einem oder zwei kritischen Parameter(n) berechnet, was einen umfassenden Einblick verspricht.

Berechnung der Simulationen

Das Modell. Die Simulationen wurden in einem von Glöckner in *Visual Basic* 6.0 erstellten Programm berechnet (vgl. Glöckner, 2006).[44] Sie basierten auf dem PCS-Netzwerk, das von Glöckner und Betsch (2008b) für probabilistische Inferenzaufgaben vorgestellt wurde (siehe Kapitel 2.3.1). Die Kernaspekte des Modells sollen noch einmal kurz erläutert werden. Die wesentlichen Komponenten stellen die Knoten und die Verbindungen zwischen den Knoten dar. Die Optionen bzw. Cues werden durch die Knoten repräsentiert, während Informationen über die Knoten in den Verbindungen repräsentiert sind. Eine besondere Stellung kommt dem generellen Validitätsknoten zu, der das System mit Energie versorgt. Die von ihm ausgehenden Verbindungen zu den Cues haben ein Gewicht w_v, welches die im Pretest angegebene subjektive Cue-Validität repräsentiert. Dementsprechend nimmt dieses Gewicht stets einen positiven Wert an. Die Verbindungsgewichte zwischen Cues und Optionen können hingegen positiv oder negativ sein: Ein positives Gewicht kennzeichnet eine exzitatorische Verbindung, d.h. dass der Cue eine positive Aussage über eine Option macht. Ein negatives Gewicht kennzeichnet eine inhibitorische Verbindung, d.h. dass der Cue eine negative Aussage über eine Option macht. Dass bei einer probabilistischen Inferenzaufgabe nur eine Option gewählt werden kann, wird durch eine starke inhibitorische Verbindung zwischen den Optionen realisiert. Zwischen den Cues bestehen keine Verbindungen.

Es wurden probabilistische Inferenzentscheidungen zwischen zwei Optionen simuliert, wobei je nach Experimentalmuster zwei, vier oder sechs Cues vorlagen. Somit wurde das Netzwerk für jedes Muster leicht modifiziert. Aus Verständlichkeitsgründen wird auch im Folgenden von „Mustern“ gesprochen. Bei den Simulationsstudien für die Experimente 3 und 4 wurden drei verschiedene Muster verwendet, die denjenigen aus den jeweiligen

[44] Ich danke Andreas Glöckner herzlich dafür, dass er mir das Programm zur Verfügung gestellt hat.

Experimenten entsprachen. Bei der Simulationsstudie für Experiment 5 wurde lediglich das zweite Muster verwendet. Tabelle 9 gibt einen Überblick über die Art der Verbindungen zwischen Cues und Optionen pro Muster. Dass die Simulationsmuster die in den Experimenten verwendeten Muster widerspiegeln, ist offensichtlich.

Tabelle 9: Überblick über die Art der Verbindungen zwischen Cues und Optionen

	Muster 1		Muster 2		Muster 3	
	Option 1	Option 2	Option 1	Option 2	Option 1	Option 2
Cue 1	inhib.	exzit.	inhib.	exzit.	exzit.	inhib.
Cue 2	exzit.	inhib.	exzit.	inhib.	exzit.	inhib.
Cue 3			inhib.	exzit.	inhib.	exzit.
Cue 4			inhib.	exzit.	inhib.	exzit.
Cue 5					inhib.	exzit.
Cue 6					exzit.	inhib.

Anmerkungen: Option 1 und 2 sowie Cue 1 bis 6 entsprechen Knoten im Netzwerk. Inhib.: Inhibitorische Verbindung zwischen Cue und Option. Exzit.: Exzitatorische Verbindung zwischen Cue und Option.

In jedem Muster bestand zudem eine stark inhibitorische Verbindung zwischen den beiden Optionen. Wurde zusätzlich ein Persönlichkeitsfaktor berücksichtigt, wurde je nach Ausprägung ein höheres oder ein niedrigeres Gewicht gewählt.

Die Parameter. Die Festlegung der Parameter orientierte sich an den Arbeiten von Glöckner et al. (im Druck). Der allgemeine Validitätsknoten hat eine konstante Aktivation von 1. Exzitatorische Verbindungen zwischen Cues und Optionen wurden mit einem Gewicht von $w_{C\text{-}O} = .01$, inhibitorische Verbindungen mit einem Gewicht von $w_{C\text{-}O} = -.01$ verwirklicht. Die stark inhibitorische Verbindung zwischen den Optionen wurde mit einem Gewicht von $w_{O1\text{-}O2} = -.50$ realisiert, wenn kein Persönlichkeitsfaktor berücksichtigt wurde. Ansonsten betrug das Gewicht entweder $w_{O1\text{-}O2} = -.40$ oder $w_{O1\text{-}O2} = -.60$. Entgegen der Entscheidung von Glöckner et al. (im Druck) wurde damit stets ein höherer Parameter festgelegt, was aber angesichts der Simulation von probabilistischen Inferenzentscheidungen nachvollziehbar ist: Wie bereits dargestellt, gibt es bei diesem Entscheidungstyp nur eine richtige Lösung. Eine stark inhibitorische Verbindung zwischen den Optionen repräsentiert

genau diese Eigenschaft.[45] Der Zerfallsparameter wurde auf .05 festgesetzt. Die minimale Aktivation der Cues betrug -1, die maximale Aktivation der Cues betrug 1.

Da das zentrale Interesse der vorliegenden Simulationsstudien darin bestand, zu bestimmen, wie gut das PCS-Modell die empirischen Ergebnisse vorhersagen kann, stellte die Übersetzung empirisch gewonnener Pretest-Validitäten in initiale Gewichte[46] einen besonders wichtigen Schritt dar. Dies erfolgte zum einen mittels einer linearen und zum anderen mittels einer quadratischen Funktionsvorschrift. Letztere trägt der Möglichkeit Rechnung, dass Unterschiede zwischen Cues betont werden, und würdigt den Einfluss sehr valider Cues (vgl. Glöckner und Bröder, 2010). Bei der linearen Funktionsvorschrift wurde von den empirisch ermittelten Validitäten (v) 50 subtrahiert, da ein Cue mit einer Validität von 50% keine Vorhersagekraft hat und dementsprechend auch in den empirischen Analysen stets 50 subtrahiert wurde. Die anschließende Multiplikation mit $\lambda = .005$ diente einer Begrenzung des maximal möglichen Gewichtes auf .25. Damit wurde der Empfehlung von Glöckner et al., dass Gewichte in einem Range von 0 bis .25 liegen sollten (im Druck), gefolgt:

$$w_v = (v - 50) * \lambda$$

Die initialen Aktivationen wurden nach der quadratischen Funktionsvorschrift folgendermaßen berechnet:

$$w_v = \left[\frac{(v - 50)}{100} \right]^2$$

Die Division durch 100 diente wieder der Begrenzung des maximal möglichen Gewichts. Auf diese Weise wird die maximale initiale Aktivation wie zuvor auf .25 beschränkt. Für die Berechnung der initialen Gewichte wurden die Pretest-Validitätswerte, die in den Experimenten ermittelt wurden und den entsprechend der Optionshierarchie oder der Entscheidung als aufzuwertend oder als abzuwertend klassifizierten Cues zuzuordnen sind, verwendet. Versuchspersonen, die sowohl im Pre- als auch im Posttest durchgehend 50% oder 100% angekreuzt haben, wurden ebenso wie Switcher ausgeschlossen.

Der Berechnungsalgorithmus. Die Ausbreitung der Aktivation wird in dem angesprochenen Programm mit einem iterativen Updating-Algorithmus simuliert. Hierbei wird

[45] Auch Glöckner (2006) hatte in früheren Studien den Inhibitionsparameter auf -.50 gesetzt.
[46] Der Begriff „initiales Gewicht“ drückt lediglich aus, dass das Gewicht zwischen dem allgemeinen Validitätsknoten und dem jeweiligen Cue dessen Pretest-Validität repräsentiert. Es bedeutet nicht, dass das Gewicht verändert wird (siehe Kapitel 2.2.3).

auf eine sigmoide Aktivationsfunktion zurückgegriffen, wie sie von McClelland und Rumelhart (1981) vorgeschlagen und von verschiedenen Forschergruppen (z.B. Glöckner und Betsch, 2008b; Thagard, 1989) in ähnlicher Weise verwendet wurde. Sie lautet nach Glöckner und Betsch (2008b, S. 218):

$$a_i(t+1) = a_i(t) * (1 - decay) + \begin{cases} if \quad input_i(t) < 0 & input_i(t)(a_i(t) - floor) \\ if \quad input_i(t) \geq 0 & input_i(t)(ceiling - a_i(t)) \end{cases}$$

$$input_i(t) = \sum_{j=1 \rightarrow n} w_{ij} a_j(t)$$

Der Iterationsprozess wird dann beendet, wenn in zehn aufeinander folgenden Iterationen keine Energieveränderung, die größer als 10^{-6} war, eintritt. Als abhängige Variable waren die finalen Aktivationen der einzelnen Cues, die die Validitätseinschätzungen im Posttest repräsentieren, von Interesse.

AUSWERTUNG

Mithilfe der Simulationen sollte überprüft werden, wie gut die empirischen Ergebnisse mittels PCS vorhergesagt werden können. Hierzu ist zunächst ein Vergleich der initialen Gewichte mit den finalen Aktivationen der Cues erforderlich, um systematische Umwertungen aufzuzeigen. Da die Gewichte und Aktivationen aufgrund unterschiedlicher Skalen nicht ohne weiteres vergleichbar sind, wurden in einem ersten Schritt jeweils die relativen Gewichte berechnet. In einem zweiten Schritt wurden dann die Summe der relativen Gewichte der aufzuwertenden Cues und die Summe der relativen Gewichte der abzuwertenden Cues gebildet. Dieser Schritt war notwendig, da lediglich eine begrenzte Höhe an Aktivation zur Verfügung steht, so dass die Erhöhung des relativen Gewichtes eines aufzuwertenden Cues die Reduzierung des relativen Gewichts eines anderen aufzuwertenden Cues mit sich bringt. Da diese Beschränkung eher dem Programm geschuldet ist und nicht dem Modell inhärent ist, erscheint die Bildung von Summen gerechtfertigt. In einem dritten Schritt wurden dann pro Muster messwiederholte ANOVAs berechnet, wobei „Messzeitpunkt“ und „Cue“ die *within subjects*-Faktoren waren. Die Ergebnisse wurden daraufhin mit den auf Cue-Ebene ermittelten Experimentalergebnissen verglichen.

Es wurden des Weiteren die finalen Aktivationen mit den empirisch ermittelten Posttestwerten korreliert, was eine Überprüfung der Übereinstimmung der Simulationsergebnisse mit den empirisch ermittelten Ergebnissen ermöglicht (vgl. Glöckner et al., im Druck). Der

zentrale Auswertungsschritt bestand dann darin zu untersuchen, ob die finalen Aktivationen auch dann einen signifikanten Beitrag zur Prädiktion der Posttest-Werte liefern, wenn der Einfluss der Pretest-Werte kontrolliert wird. Hierfür wurden Regressionsanalysen mit den empirisch ermittelten Posttest-Werten als Kriterium und den Pretest-Werten sowie den finalen Aktivationen als Prädiktoren berechnet (vgl. Glöckner et al., im Druck). Bezüglich der quadratischen Transformationsfunktion kann der Einwand erhoben werden, dass zwar auf der einen Seite postuliert wird, dass Pretest-Validitäten gemäß einer derartigen Funktion in eine mentale Repräsentation übersetzt werden, dass aber andererseits mit den resultierenden Aktivationen die untransformierten Posttest-Validitäten vorhergesagt werden sollen. Um diesem Einwand zu begegnen, wurden in einem weiteren Schritt alle entsprechenden Korrelationen auch mit transformierten Posttest-Werten und alle Regressionen mit transformierten Posttest- und Pretest-Werten berechnet.

3.6.2 Simulationsstudie 1: Die Vorhersage der Posttest-Valditäts-einschätzungen aus Experiment 3

ZIELSETZUNG

Mit den folgenden Simulationsstudien sollte überprüft werden, inwieweit sich die in Experiment 3 gefundenen Ergebnisse mit dem PCS-Modell vorhersagen lassen. Von besonderer Relevanz war es, den Beitrag der Simulationsergebnisse an der Vorhersage der Posttest-Werte zu ermitteln.

METHODE

Die methodische Vorgehensweise entsprach der in Kapitel 3.7.1 geschilderten, wobei das Gewicht der Verbindung zwischen den Optionen -.50 betrug. Es wurde also pro Versuchsperson jeweils eine Simulation unter Verwendung der linearen und der quadratischen Transformationsfunktion berechnet. Die Klassifikation in aufzuwertende und abzuwertende Cues basierte auf der gemäß einer korrigierten WADD-Strategie zu erwartenden Entscheidung. Für die Versuchspersonen, die die Cue-Validitäten durchgängig mit 50% oder 100% eingeschätzt haben und deren Optionenhierarchie wechselte (Switcher), wurden keine Simulationen berechnet.

In Anlehnung an die erste Bedingung des dritten Experiments wurde eine Situation simuliert, in der basierend auf zwei Cues, die entgegengesetzte Vorhersagen machten, eine

Entscheidung zwischen zwei Optionen getroffen werden musste. Die Gewichte des aufzuwertenden und des abzuwertenden Cues wurden so eingesetzt, dass Cue 1 dem aufzuwertenden Cue entsprach und damit exzitatorisch mit Option 2 verbunden war, während Cue 2 dem abzuwertenden Cue entsprach und damit exzitatorisch mit Option 1 verbunden war. Es wurden Simulationen für 33 Versuchspersonen berechnet.

In Anlehnung an die zweite Bedingung des dritten Experiments wurde eine Situation simuliert, in der basierend auf vier Cues, von denen einer (Cue 2) exzitatorisch mit Option 1, aber inhibitorisch mit Option 2 verbunden war, und drei (Cues 1, 3, 4) exzitatorisch mit Option 2, aber inhibitorisch mit Option 1 verbunden waren, eine Entscheidung zwischen zwei Optionen getroffen werden musste. Die Gewichte wurden so eingesetzt, dass die Cues 1, 3 und 4 die aufzuwertenden Cues darstellten, während Cue 2 den abzuwertenden Cue repräsentierte. Es wurden Simulationen für 38 Versuchspersonen berechnet.

In Anlehnung an die dritte Bedingung des dritten Experiments wurde eine Situation simuliert, in der basierend auf sechs Cues, von denen drei (Cues 1, 2, 6) exzitatorische Verbindungen zu Option 1, aber inhibitorische Verbindungen zu Option 2, und drei (Cues 3, 4, 5) exzitatorische Verbindungen zu Option 2, aber inhibitorische Verbindungen zu Option 1 aufwiesen, eine Entscheidung zwischen zwei Optionen getroffen werden musste. Die Gewichte wurden so eingesetzt, dass die Cues 1, 2 und 6 den drei aufzuwertenden Cues entsprachen, während die Cues 3, 4 und 5 den drei abzuwertenden Cues entsprachen. Diese Simulationen wurden für insgesamt 37 Versuchspersonen berechnet.

ERGEBNISSE

Nachdem die initialen Gewichte und die finalen Aktivationen, wie in Kapitel 3.7.1 beschrieben, in relative Gewichte übersetzt worden waren und die Summe der relativen Gewichte aufzuwertender bzw. abzuwertender Cues berechnet worden war, wurden messwiederholte ANOVAs mit „Messzeitpunkt“ und „Cue“ als *within subjects*-Faktoren berechnet. Wurden die Simulationen mit initialen Gewichten, die über die lineare Transformationsfunktion ermittelt worden waren, berechnet, so wurde die interessierende Interaktion „Messzeitpunkt x Cue“ im ersten Muster signifikant, $F(1, 32) = 6.13$, $p < .05$, $\eta_p^2 = .16$. Eine Inspektion der Mittelwerte zeigt allerdings, dass es entgegen der Erwartung zu einer Aufwertung des abzuwertenden Cues und zu einer Abwertung des aufzuwertenden Cues gekommen ist. Statt eines *spreading apart* fand vielmehr ein *approaching* statt. Für das zweite Muster wurde hingegen eine signifikante systematische Umwertung gefunden, $F(1,$

37) = 8.80, $p < .01$, $\eta_p^2 = .19$. Die Interaktion von „Messzeitpunkt x Cue“ wurde beim dritten Muster wiederum nicht signifikant, $F(1, 36) < 1$, $p = .81$, $\eta_p^2 = 00$. Für die Ergebnisse, die auf den über die quadratische Transformationsfunktion ermittelten initialen Gewichten beruhen, zeigte sich ein ähnliches Ergebnismuster: Für das erste Muster wurde die Interaktion „Messzeitpunkt x Cue“ nicht signifikant, $F(1, 31) = 2.95$, $p = .10$, $\eta_p^2 = .09$, wobei aber eine Inspektion der Mittelwerte auf eine zumindest tendenziell systematische Umwertung schließen lässt.[47] Eine signifikante systematische Umwertung zeigte sich erneut für das zweite Muster, $F(1, 37) = 10.04$, $p < .01$, $\eta_p^2 = .21$, sowie nun auch für das dritte Muster, $F(1, 36) = 18.63$, $p < .001$, $\eta_p^2 = .34$. Im zugehörigen Experiment 3 wurden lediglich für die Muster 2 und 3 systematische Umwertungen auf der Cue-Ebene gefunden.

Im Rahmen der Analyse, wie gut die empirischen Ergebnisse mit dem PCS-Modell vorhergesagt werden können, wurden in einem ersten Schritt die finalen Aktivationen der Cues mit den Posttest-Validitäten (bzw. mit den transformierten Posttest-Validitäten) korreliert. Die Ergebnisse sind in Tabelle 10 wiedergegeben.

Tabelle 10: Korrelationen zwischen den empirischen Posttest-Validitäten und den finalen Aktivationen

	Muster 1			Muster 2			Muster 3		
	linear	quadratisch		linear	quadratisch		linear	quadratisch	
		a	b		a	b		a	b
Cue 1	.77**	.82**	.78**	.82**	.80**	.76**	.42*	.48**	.46**
Cue 2	.73**	.78**	.73**	.64**	.64**	.64**	.78**	.86**	.79**
Cue 3				.62**	.66**	.63**	.68**	.80**	.79**
Cue 4				.87**	.85**	.83**	.49**	.62**	.60**
Cue 5							.66**	.75**	.72**
Cue 6							.66**	.79**	.76**
n	33	32		38	38		37	37	

Anmerkungen: Die Bezeichnungen der Cues gelten stets spezifisch für die einzelnen Muster (Erläuterung siehe Text). Es werden die Korrelationen für die Studien, in denen die lineare und in denen die quadratische Transformationsfunktion angewendet wurde, getrennt berichtet. Bei letzterer wird zudem unterschieden, ob die empirischen Posttest-Werte ebenfalls transformiert wurden (b) oder nicht (a). *n*: Anzahl an Versuchspersonen, *: $p < .05$, **: $p < .01$.

[47] Eine Versuchsperson wurde aus der Analyse ausgeschlossen, da die finalen Aktivationen der Knoten 0 betrugen.

Die Korrelationen sind gemäß den Konventionen von Cohen (1988) mehrheitlich groß, was darauf hindeutet, dass die Posttest-Werte sehr gut durch das PCS-Modell vorhergesagt werden können. Es treten zwar Abweichungen bezüglich der Korrelationshöhe zwischen den Transformationsfunktionen bzw. innerhalb der quadratischen Transformationsfunktion auf, jedoch sind diese eher moderat. Lediglich beim dritten Muster sind diese deskriptiv etwas größer, wobei die auf Grundlage der quadratischen Transformationsfunktion und mit untransformierten Posttest-Werten berechneten Korrelationen am höchsten sind. Bei der Interpretation dieser Korrelationen ist natürlich zu bedenken, dass der Einfluss des Pretests nicht herauspartialisiert wurde. Um nun die Bedeutung der PCS-Prozesse unter Berücksichtigung der Pretest-Werte näher zu beleuchten, wurden in einem nächsten Schritt Regressionen berechnet, in denen die Posttest-Werte das Kriterium und die Pretest-Einschätzungen („Pretest“) sowie die finalen Aktivationen („PCS“) die Prädiktoren darstellten. Wurden auch die Posttest-Werte transformiert, wurden diese als Kriterium und die initialen Gewichte (das heißt die transformierten Pretest-Werte) als der zweite Prädiktor berücksichtigt. Durch die Berechnung von Regressionen sollte der Beitrag des Modells ermittelt werden, wenn für die Pretest-Werte kontrolliert wird. Da es sich um messwiederholte Regressionen handelte und damit n (= Anzahl der Versuchspersonen pro Bedingung) Cluster an Beobachtungen vorliegen, werden entsprechend cluster-robuste Standardfehler berechnet (Rogers, 1993). Die Regressionen wurden für jede Bedingung und jede Transformationsfunktion berechnet. Eine Teststärkenanalyse mittels GPower (Faul, Erdfelder, Buchner & Lang, 2009) ergab, dass bei $n_{\text{Muster 1}} = 33$, $n_{\text{Muster 2}} = 38$ und $n_{\text{Muster 3}} = 37$ und jeweils einem Signifikanzniveau von $\alpha = .05$ ein mittlerer Effekt ($f^2 = .15$, Cohen, 1988) der Prädiktoren mit Teststärken von .70, .76 und .75 gefunden werden kann. Durch das Anheben von α auf .10 konnten die Teststärken auf .82, .86 und .85 erhöht werden. Die Ergebnisse der Regressionsanalysen können Tabelle 11 entnommen werden. Ein signifikanter Beitrag der finalen Aktivationen ist deutlich in Muster 2 zu beobachten. In Muster 3 leistet der PCS-Prädiktor nur dann einen signifikanten Beitrag, wenn die Pretest-Werte quadratisch transformiert wurden und keine Rücktransformation stattfand. In Muster 1 leistet der PCS-Prädiktor in keiner der Analysen einen signifikanten Beitrag. Diese Ergebnisse entsprechen weitestgehend den empirischen Befunden: So waren für das zweite Muster deutliche Umwertungen zu beobachten. Beim dritten Muster waren zwar auf der Cue-Ebene Umwertungen festgestellt worden, aber lediglich bei einem Cue. Beim ersten Muster waren keine Umwertungen zu beobachten gewesen. Entsprechend dem PCS-Modell wurde ein

signifikanter Beitrag in jedem Muster erwartet, doch zeigen die Simulationen in Übereinstimmung mit den empirischen Daten, dass nicht in jedem Muster systematische und signifikante Umwertungen auftreten. Das Ergebnis des dritten Musters ist mit Vorsicht zu interpretieren, da lediglich in einer Berechnungsvariante ein signifikanter Beitrag des PCS-Prädiktors zu verzeichnen war.

Tabelle 11: Ergebnisse der Regressionsanalysen

	Posttest								
	Muster 1			Muster 2			Muster 3		
	linear	quadratisch		linear	quadratisch		linear	quadratisch	
		a	b		a	b		a	b
Pretest	.83*** (8.88)	.79*** (5.67)	.76*** (6.50)	.62*** (4.58)	.62*** (5.53)	.66*** (7.69)	.81*** (10.95)	.63*** (5.59)	.75*** (9.16)
PCS	.07 (0.81)	.10 (0.75)	.13 (1.21)	.21^{+} (1.74)	.21* (2.14)	.19* (2.63)	.02 (0.34)	.21^{+} (2.01)	.10 (1.32)
N_B	66	64	64	152	152	152	222	222	222
R^2	.79	.77	.76	.67	.67	.68	.69	.70	.69

Anmerkungen: Es werden die β-Koeffizienten berichtet, und zwar getrennt für Analysen, in denen die lineare oder quadratische Transformationsfunktion verwendet wurde. Bei letzterer wird zudem unterschieden, ob die empirischen Posttest-Werte ebenfalls transformiert wurden (b) oder nicht (a). *t*-Werte stehen in Klammern. Aufgrund der Messwiederholung wurden cluster-robuste Standardfehler verwendet. N_B: Anzahl an Beobachtungen, $^{+}$: $p < .10$, *: $p < .05$, ***: $p < .001$.

DISKUSSION

Das Ziel der vorgestellten Simulationsstudien bestand darin zu überprüfen, wie gut die empirischen Ergebnisse mittels des PCS-Modells vorhergesagt werden können. Es zeigte sich, dass die in Experiment 3 gefundenen Unterschiede zwischen den einzelnen Mustern unter Verwendung einer linearen Transformationsfunktion teilweise und unter Verwendung einer quadratischen Transformationsfunktion vollständig vorhergesagt werden können. Des Weiteren ist festzuhalten, dass die Ergebnisse der Simulationen (die finalen Aktivationen) lediglich für das zweite (1:3-) Muster unabhängig von der Berechnungsvariante einen signifikanten Beitrag zur Vorhersage der Posttest-Werte leisten. Die Ergebnisse für das dritte (3:3-) Muster sind inkonsistent und demnach eher mit Vorsicht zu interpretieren. Obgleich Umwertungen aus theoretischer Sicht zu erwarten gewesen wären, liegt unter Berücksichtigung der eher schwachen empirischen Evidenz bezüglich der Umwertung in diesem Muster

und der gescheiterten Replikation in einem späteren Experiment (Experiment 4) die Vermutung nahe, dass die Konsistenzmaximierungsprozesse in diesem Muster wie in dem ersten (1:1-) Muster ebenfalls kaum zum Tragen kommen. Zusammenfassend ist Folgendes festzuhalten: Es hätten entsprechend den theoretischen Annahmen des PCS-Modells sowohl Umwertungen als auch ein signifikanter Beitrag des PCS-Prädiktors in jedem Muster erwartet werden können. Die Simulationsergebnisse sind damit in Übereinstimmung mit den empirischen Ergebnissen, in denen lediglich für das zweite Muster signifikante Umwertungen gefunden werden konnten. Es wurden Simulationen nach einer linearen und nach einer quadratischen Transformationsfunktion berechnet. Um der etwaigen Kritik, dass unter Verwendung der quadratischen Funktion Analysen mit untransformierten Werten durchgeführt werden, zuvorzukommen, wurden sämtliche diesbezüglichen Analysen zusätzlich auch mit transformierten Pretest- und Posttest-Validitäten berechnet. Es zeigten sich lediglich geringe Abweichungen.

3.6.3 Simulationsstudie 2: Die Vorhersage der Posttest-Validitäts-einschätzungen aus Experiment 4

ZIELSETZUNG

Das wesentliche Ziel der vorzustellenden Simulationen bestand in der Überprüfung der Frage, inwieweit sich mit dem PCS-Modell die Resultate des vierten Experimentes vorhersagen lassen. Da im zugrunde liegenden Experiment ein Einfluss eines Persönlichkeitsfaktors, nämlich der Präferenz für Konsistenz, auf die Stärke der Umwertungen beobachtet werden konnte, sollte des Weiteren überprüft werden, ob die zusätzliche Berücksichtigung dieses Persönlichkeitsfaktors zu einer verbesserten Vorhersage führt.

METHODE

Die Muster, die in ein PCS-Modell übersetzt wurden, entsprachen denen der vorherigen Simulationsstudien. Die Klassifikation in aufzuwertende und abzuwertende Cues basierte aber – wie im Experiment selbst – auf der tatsächlich getroffenen Entscheidung. Die Simulationen wurden entsprechend dem in Kapitel 3.7.1 berichteten Vorgehen berechnet. In einem ersten Set an Studien wurde das Gewicht der Verbindung zwischen den Optionen wie bisher auf -.50 festgesetzt. In einem zweiten Set wurde das Gewicht der Verbindung zwischen den Optionen hingegen variiert: Um die Unterschiedlichkeit in der Präferenz für

Konsistenz zu berücksichtigen, betrug dieser Parameter entweder -.40 oder -.60. Ein höheres Gewicht geht mit einem stärkeren Konsistenzmaximierungsprozess einher, was auf empirischer Ebene einer stärkeren Umwertung von Cue-Validitäten entspricht. Folglich wurde das Gewicht bei Personen mit einer hohen Präferenz für Konsistenz (oberhalb des Medians) auf -.60 und bei Personen mit einer niedrigen Präferenz für Konsistenz (unterhalb des Medians) auf -.40 festgesetzt. Die Mediane wurden pro Bedingung berechnet und betrugen 4.50 in Bedingung 1, 5.00 in Bedingung 2 und 4.89 in Bedingung 3. Auf diesem Wege soll entschieden werden, ob die zusätzliche Berücksichtigung der Präferenz für Konsistenz zu einer höheren Varianzaufklärung führt. Erneut wurde in jedem Set pro Versuchsperson und pro Transformationsfunktion eine Simulation berechnet. Es wurden Simulationen für 32 (Muster 1), 32 (Muster 2) und für 25 Versuchspersonen (Muster 3) berechnet.

Ergebnisse

Um die Frage zu beantworten, ob sich die empirischen Ergebnisse mit dem PCS-Modell vorhersagen lassen, wurde wieder in einem ersten Schritt die Summe der relativen Gewichte aufzuwertender bzw. abzuwertender Cues berechnet.[48] Anschließende messwiederholte ANOVAs wurden erneut pro Muster und pro Transformationsfunktion berechnet. Unter Zugrundelegung der linearen Transformationsfunktion für die Berechnung der Simulationen wurde die interessierende Interaktion „Messzeitpunkt x Cue“ für das erste Muster wieder nicht signifikant, $F(1, 31) = 1.03$, $p = .32$, $\eta_p^2 = .03$, wie auch für das zweite Muster, $F(1, 31) < 1$, $p = .70$, $\eta_p^2 = .01$, und das dritte Muster, $F(1, 24) = 2.66$, $p = .12$, $\eta_p^2 = .10$. Wurden die initialen Gewichte mit der quadratischen Transformationsfunktion berechnet, so wurde die Interaktion „Messzeitpunkt x Cue“ sowohl für das erste Muster signifikant, $F(1, 31) = 4.86$, $p < .05$, $\eta_p^2 = .14$, als auch für das zweite Muster, $F(1, 31) = 13.07$, $p < .01$, $\eta_p^2 = .30$. Eine Inspektion der Mittelwerte zeigte, dass die Umwertungen systematisch verliefen. Für das dritte Muster wurde die Interaktion hingegen nicht signifikant, $F(1, 24) < 1$, $p = .38$, $\eta_p^2 = .03$. Im Experiment war keine der Interaktionen auf Cue-Ebene signifikant geworden, auf Optionen-Ebene hatte sich hingegen ein signifikantes *spreading apart* beim zweiten Muster gezeigt.

[48] Die folgenden Ergebnisse basieren auf den Simulationsstudien, in denen das Gewicht zwischen den Optionen -.50 betrug. Berechnet man die Analysen mit den Ergebnissen aus dem anderen Set, so zeigt sich das gleiche Ergebnismuster.

In einem weiteren Schritt wurden wieder die finalen Aktivationen der Knoten mit den empirisch ermittelten Posttestwerten (bzw. mit den transformierten Werten) korreliert. Tabelle 12 gibt einen Überblick über die Korrelationen.

Tabelle 12: Korrelationen zwischen den empirischen Posttest-Validitäten und den finalen Aktivationen

	Muster 1			Muster 2			Muster 3		
	linear	quadratisch		linear	quadratisch		linear	quadratisch	
		a	b		a	b		a	b
C1	.44* (44*)	.44* (.44*)	.57** (.57*)	.87** (.87**)	.94** (.94**)	.87** (.87**)	.49* (.49*)	.64** (.64*)	.59** (.59**)
C2	.75** (75**)	.80** (.80**)	.77** (.77**)	.76** (.76**)	.81** (.81**)	.73** (.73**)	.77** (.77**)	.82** (.82**)	.73** (.73**)
C3				.78** (.78**)	.80** (.80**)	.73** (.73**)	.76** (.76**)	.78** (.79**)	.77** (.77**)
C4				.88** (.88**)	.98** (.98**)	.91** (.90**)	.79** (.79*)	.82** (.82**)	.74** (.74**)
C5							.85** (.85**)	.89** (.89**)	.82** (.82**)
C6							.84** (.84**)	.82** (.82**)	.69** (.69**)
n	32	32		32	32		25	25	

Anmerkungen: Die Bezeichnungen der Cues (C1 bis C6) gelten stets spezifisch für die einzelnen Muster (Erläuterung siehe Text). Es werden die Korrelationen für die Simulationen, in denen die lineare und in denen die quadratische Transformationsfunktion verwendet wurde, getrennt berichtet. Bei letzterer wird zudem unterschieden, ob die empirischen Posttest-Werte ebenfalls transformiert wurden (b) oder nicht (a). In Klammern stehen die Ergebnisse für die Simulationen, in denen zusätzlich die Höhe des PFC berücksichtigt wurde. *n*: Anzahl an Versuchspersonen, *: $p < .05$, **: $p < .01$.

Die Korrelationen sind wieder überwiegend hoch (vgl. Cohen, 1988), und zwar unabhängig von der Berechnungsvariante. Zudem unterscheiden sich die Korrelationen je nach dem, ob die Ausprägung der Präferenz für Konsistenz berücksichtigt wurde oder nicht, kaum. Einen überzeugenderen Test dafür, wie gut die Ergebnisse mit dem PCS-Modell vorhergesagt werden können, stellen Regressionsanalysen dar: So wurde auch in diesen Simulationsstudien anhand von messwiederholten Regressionen analysiert, ob die finalen Aktivationen einen signifikanten Beitrag zur Prädiktion der empirischen Posttest-Werte (bzw. der transformierten Werte) liefern, wenn für die Pretest-Werte (bzw. die transformier-

ten Werte) kontrolliert wird. Hierbei war es auch von Interesse, ob die zusätzliche Berücksichtigung von Persönlichkeitsfaktoren zu einer höheren Varianzaufklärung führt oder nicht. Wie bei den vorherigen Studien wurden die Regressionen für jede Bedingung und jede Transformationsfunktion berechnet. Die Standardfehler wurden aufgrund der Messwiederholung erneut für *n* (= Anzahl der Versuchspersonen pro Bedingung) Cluster an Beobachtungen angepasst. Teststärkenanalysen mittels GPower (Faul et al., 2009) ergaben, dass ein Signifikanzniveau von $\alpha = .05$ die Entdeckung von Effekten der Größe $f^2 = .15$ (vgl. Cohen, 1988) mit Teststärken von .69 für Prädiktoren in Muster 1 ($n = 32$) und in Muster 2 ($n = 32$) und von .59 für Prädiktoren in Muster 3 ($n = 25$) erlaubt. Eine Anhebung des α-Niveaus von .10 führte zu einer Erhöhung der Teststärken auf .81 (Muster 1 und 2) und .73 (Muster 3). Die Ergebnisse der Regressionsanalysen werden separat für beide Sets an Simulationsstudien (ohne Berücksichtigung des PFC und mit Berücksichtigung) dargestellt (Tabellen 13 und 14).

Tabelle 13: Ergebnisse der Regressionsanalysen für das erste Set an Simulationsstudien

	Posttest								
	Muster 1			Muster 2			Muster 3		
	linear	quadratisch		linear	quadratisch		linear	quadratisch	
		a	b		a	b		a	b
Pretest	.57** (3.33)	.42 (1.64)	.59** (3.57)	.84*** (6.94)	.77*** (6.40)	.84*** (10.71)	.95*** (11.96)	.96*** (12.24)	.87*** (13.71)
PCS	.21* (2.19)	.35^{+} (1.82)	.24* (2.43)	.05 (0.73)	.13^{+} (1.79)	.07 (1.67)	-.07 (-0.79)	-.08 (-0.96)	.02 (0.35)
N_B	64	64	64	128	128	128	150	150	150
R^2	.57	.57	.64	.78	.79	.80	.79	.79	.78

Anmerkungen: In diesen Simulationsstudien wurde die Höhe der Präferenz für Konsistenz nicht berücksichtigt. Es werden die β-Koeffizienten berichtet, und zwar getrennt für Analysen, in denen die lineare oder quadratische Transformationsfunktion verwendet wurde. Bei letzterer wird zudem unterschieden, ob die empirischen Posttest-Werte ebenfalls transformiert wurden (b) oder nicht (a). *t*-Werte stehen in Klammern. Aufgrund der Messwiederholung wurden cluster-robuste Standardfehler verwendet. N_B: Anzahl an Beobachtungen, $^{+}$: $p < .10$, *: $p < .05$, **: $p < .01$, ***: $p < .001$.

Tabelle 14: Ergebnisse der Regressionsanalysen für das zweite Set an Simulationsstudien

	Posttest								
	Muster 1			Muster 2			Muster 3		
	linear	quadratisch		linear	quadratisch		linear	quadratisch	
		a	b		a	b		a	b
Pretest	.57** (3.32)	.41 (1.64)	.59** (3.56)	.84*** (6.89)	.77*** (6.32)	.84*** (10.64)	.95*** (11.82)	.96*** (12.18)	.87*** (13.69)
PCS	.21* (2.20)	.35^{+} (1.84)	.24* (2.44)	.06 (0.75)	.13^{+} (1.81)	.07 (1.7)	-.07 (-0.77)	-.08 (-0.94)	.02 (0.36)
N_B	64	64	64	128	128	128	150	150	150
R^2	.57	.57	.65	.78	.79	.80	.79	.79	.78

Anmerkungen: Bei diesen Simulationsstudien wurde die Höhe der Präferenz für Konsistenz durch eine Modifikation des inhibitorischen Gewichts zwischen den Optionen berücksichtigt. Es werden die β-Koeffizienten berichtet, und zwar für Analysen auf Basis der linearen und der quadratischen Transformationsfunktion. Bei letzterer wurden die empirischen Posttest-Werte entweder auch transformiert (b) oder nicht (a). *t*-Werte stehen in Klammern. Aufgrund der Messwiederholung wurden cluster-robuste Standardfehler verwendet. N_B: Anzahl an Beobachtungen, $^{+}$: $p < .10$, *: $p < .05$, **: $p < .01$, ***: $p < .001$.

Wie den Tabellen zu entnehmen ist, ist die Prädiktionskraft der auf Basis des PCS-Modells berechneten finalen Aktivationen je nach Muster und je nach Transformationsfunktion unterschiedlich. Im ersten Muster leisten die finalen Aktivationen einen signifikanten Beitrag zur Vorhersage der Posttest-Werte. Obgleich dieses Ergebnis mit den theoretischen Vorhersagen des PCS-Modells in Einklang ist, ist es vor dem Hintergrund der empirischen und der vorherigen Simulationsergebnisse überraschend. Im zweiten Muster leistet der PCS-Prädiktor dann einen signifikanten Beitrag, wenn die quadratische Transformationsfunktion verwendet wurde und keine Transformation der Posttest-Werte stattfand. In letzterem Fall ist der Test mit $p = .11$ (ohne PFC) und $p = .10$ allerdings nur knapp an der Signifikanzgrenze gescheitert. Ein über die Berechnungsvarianten konsistentes Ergebnis wäre zwar theoretisch zu erwarten und auch eindeutiger gewesen. Allerdings können die Ergebnisse auch dahingehend interpretiert werden, dass die quadratische Transformationsfunktion die empirischen Ergebnisse besser abzubilden scheint. Diese Vermutung bedarf natürlich der weiteren Testung. Im dritten Muster weisen die finalen Aktivationen in keiner Analyse signifikante Prädiktionskraft auf. Erneut entsprechen die Ergebnisse nicht den generellen Vorhersagen des PCS-Modells und sind zudem nur zum Teil in Übereinstimmung mit den empirischen Daten.

In keinem der Muster machte es bezüglich der Vorhersagekraft der Prädiktoren oder bezüglich der Varianzaufklärung einen Unterschied, ob die Präferenz für Konsistenz berücksichtigt wurde oder nicht.

Diskussion

Das Ziel der vorliegenden Simulationsstudie bestand in der Überprüfung, inwieweit das PCS-Modell die in Experiment 4 ermittelten Ergebnisse vorhersagen kann. Die Ergebnisse entsprechen weder vollkommen den theoretischen Vorhersagen des PCS-Modells noch den empirischen Ergebnissen des zugrunde liegenden Experimentes. So konnte unter Verwendung der linearen Transformationsfunktion in keinem Muster eine signifikante Interaktion zwischen dem Messzeitpunkt und dem Cue gefunden werden. Obgleich im Experiment ein *spreading apart* der Optionen für das zweite Muster gefunden werden konnte, war dies auf der Ebene der Cues nicht der Fall, auch nicht bei den anderen Mustern, so dass dieses Ergebnis zwar im Sinne von PCS nicht zu erwarten war, aber vor dem Hintergrund der Experimentalresultate verständlich ist. Schwer zu interpretieren ist allerdings, dass unter Verwendung der quadratischen Transformationsfunktion nicht nur im zweiten Muster, sondern auch im ersten Muster eine signifikante Interaktion gefunden werden konnte. Es kann hier weniger ein systematischer Effekt als vielmehr ein Fehler 1. Art vermutet werden, zumal diese Interaktion für das dritte Muster nicht signifikant war und folglich das Ergebnismuster nicht in Einklang mit dem PCS-Modell ist. Auch hinsichtlich der Prädiktionskraft der mit PCS generierten Aktivationen ist die Befundlage geteilt. So ist zunächst festzuhalten, dass die Korrelationen zwischen Posttest-Validitäten und finalen Aktivationen wieder sehr hoch sind. Problematisch ist allerdings, dass der PCS-Prädiktor unter Verwendung der linearen und der quadratischen Funktion einen signifikanten Beitrag zur Vorhersage der Posttest-Validitäten in Muster 1 leistet. Da dies nur das erste, nicht aber das dritte Muster betrifft und angesichts vorheriger Simulationsergebnisse nicht zu erwarten war, ist dieses Ergebnis sowohl vor dem Hintergrund des PCS-Modells als auch der Experimentalergebnisse überraschend und nicht plausibel zu erklären. Für das zweite Muster wurde entgegen der Erwartung ein signifikanter Beitrag des PCS-Prädiktors nur unter der quadratischen Berechnungsvariante, wenn die Posttest-Werte nicht transformiert waren, gefunden. Eventuell ist dies damit zu begründen, dass auch auf empirischer Ebene keine signifikanten Cue-Umwertungen aufgetreten waren und die Evidenz für systematische Informationsverzerrungen dementsprechend schwach war. Dass aber der PCS-Prädiktor nur

knapp an der Signifikanzgrenze scheiterte, wenn die Posttest-Werte transformiert waren, könnte als erster Hinweis darauf gedeutet werden, dass eine quadratische Transformationsfunktion besser geeignet ist, um die empirischen Ergebnisse abzubilden. Dies wird gestützt durch die Befunde der ANOVAs, obgleich auch hier die Resultate für das erste Muster schwer zu vereinbaren sind. Für das dritte Muster konnte ebenfalls kein signifikanter Beitrag des PCS-Prädiktors verzeichnet werden. Die Berücksichtigung der Ausprägung des PFC hatte in keiner Analyse einen nennenswerten Einfluss. Einschränkend ist anzuführen, dass die Übersetzung der Ausprägung des PFC in einen Simulationsparameter ungenau war. Da aber auch in einem nachfolgenden Experiment der Einfluss des PFC nicht repliziert werden konnte, scheint die Bedeutung dieses Faktors zu vernachlässigen zu sein. Zusammenfassend ist Folgendes festzuhalten: Während zwar aus theoretischer Sicht eine Umwertung und damit ein signifikanter Beitrag des PCS-Prädiktors in allen Mustern hätten erwartet werden können, hätte auf Basis der bisherigen empirischen und Simulationsergebnisse zumindest ein Effekt für das zweite Muster erwartet werden können. Für dieses Muster konnten zum Teil signifikante Umwertungen bzw. ein signifikanter Beitrag des Prädiktors ermittelt werden, allerdings sind die Effekte nicht eindeutig. Eventuell ist dies der schwachen empirischen Evidenz und damit der Datenlage geschuldet. Dies erklärt allerdings nicht die Ergebnisse für das erste Muster. Hinsichtlich der eventuellen Überlegenheit der quadratischen Transformationsfunktion (bezogen auf die Abbildung der empirischen Ergebnisse) könnte eine weitere Simulationsstudie ein deutlicheres Bild vermitteln.

3.6.4 Simulationsstudie 3: Die Vorhersage der Posttest-Validitäts-einschätzungen aus Experiment 5

ZIELSETZUNG

Die vorzustellenden Simulationsstudien wurden auf Basis der Daten des fünften Experimentes durchgeführt. Erneut bestand die Zielsetzung im Wesentlichen darin zu überprüfen, ob der PCS-Prädiktor einen signifikanten Beitrag zur Vorhersage der Posttest-Validitäten macht. Hierbei wurden wieder die lineare und die quadratische Transformationsfunktion angewandt, wobei bei letzterer zu unterscheiden ist, ob die nachfolgenden Analysen mit untransformierten oder mit transformierten Pretest- und Posttest-Werten erfolgten. Es wurde zudem überprüft, ob sich die im Experiment gefundene Umwertung von Cues durch das PCS-Modell vorhersagen lässt.

METHODE

Methodisch wurde entsprechend der Darstellung in Kapitel 3.7.1 vorgegangen. Das Gewicht zwischen den Optionen $w_{O1\text{-}O2}$ betrug -.50. Wie in den vorherigen Studien wurde pro Versuchsperson eine Simulation unter Verwendung der linearen Transformationsfunktion und eine Simulation unter Verwendung der quadratischen Transformationsfunktion berechnet. Die Person, die sich als einzige in Experiment 6 für Option A entschieden hatte, wurde ausgeschlossen. Folglich wurden alle Cues, die für Option B sprachen, als aufzuwertend klassifiziert, während der Cue, der für Option A sprach, als abzuwertend klassifiziert wurde. Entsprechend dem Material in Experiment 5 wurde ein PCS-Muster aus zwei Optionen und vier Cues, von denen einer (Cue 2) exzitatorisch (inhibitorisch) mit Option 1 (2) verbunden war und drei (Cues 1, 3, 4) exzitatorisch (inhibitorisch) mit Option 2 (1) verbunden waren. Es wurden Simulationen auf Basis der Daten von 50 Versuchspersonen berechnet.

ERGEBNISSE

In einem ersten Schritt wurde überprüft, ob die Umwertung von Cues durch das PCS-Modell vorhergesagt wird. Dazu wurden zunächst wie in den vorherigen Simulationsstudien die initialen Gewichte sowie die finalen Aktivationen in relative Gewichte übersetzt. Es wurde eine messwiederholte ANOVA mit „Messzeitpunkt“ und „Cue“ als Innersubjektfaktoren berechnet. Die ANOVA wurde jeweils für die Daten, die unter Verwendung der linearen und die unter Verwendung der quadratischen Transformationsfunktion gewonnen wurden, berechnet. Wurde für die Berechnung der initialen Gewichte die lineare Transformationsfunktion verwendet, so wurde die Interaktion „Messzeitpunkt x Cue“ nicht signifikant, $F(1, 49) < 1$, $p = .84$, $\eta_p^2 = .00$. Ein anderes Bild zeigte sich, wenn die Berechnung der initialen Gewichte nach der quadratischen Transformationsfunktion erfolgte. Die Interaktion „Messzeitpunkt x Cue“ wurde signifikant, $F(1, 49) = 8.41$, $p < .01$, $\eta_p^2 = .15$. Das heißt, dass systematische Umwertungen – wie sie auch in den experimentellen Daten gefunden worden waren – aufgetreten sind. Somit werden die experimentellen Ergebnisse unter Verwendung der quadratischen Transformationsfunktion, nicht aber unter Verwendung der linearen Transformationsfunktion, vom PCS-Modell vorhergesagt. Um gezielter die Frage zu beantworten, ob das PCS-Modell die empirisch ermittelten Posttest-Werte vorhersagen kann, wurden wieder Korrelations- und Regressionsanalysen durchgeführt. Zunächst wurden die finalen Aktivationen der Cues mit den Posttest-Validitäten bzw. – im Falle der quadratischen

Transformationsfunktion – zusätzlich mit den transformierten Posttest-Validitäten korreliert. Die Ergebnisse für jede Transformationsfunktion können Tabelle 15 entnommen werden.

Tabelle 15: Korrelation zwischen den empirischen Posttestwerten und den finalen Aktivationen

	Linear	Quadratisch	
		a	b
Cue 1	.88**	.89**	.84**
Cue 2	.73**	.75**	.72**
Cue 3	.58**	.73**	.66**
Cue 4	.85**	.86**	.81**
N	50	50	50

Anmerkungen: Cues 1, 3 und 4 entsprechen den aufzuwertenden Cues, Cue 2 entspricht dem abzuwertenden Cue. Es werden die Korrelationen für die Simulationen, in denen die lineare und in denen die quadratische Transformationsfunktion angewendet wurde, getrennt berichtet. Bei letzterer wird zudem unterschieden, ob die empirischen Posttest-Werte ebenfalls transformiert wurden (b) oder nicht (a). **: $p < .01$.

Sowohl unter Verwendung der linearen als auch unter Verwendung der quadratischen Funktion sind die Korrelationen zwischen den finalen Aktivationen eines Cues und der entsprechenden Posttest-Validität gemäß der Konventionen von Cohen (1988) als groß einzustufen. Besonders bei Cue 3 ist auffällig, dass die Korrelationen etwas höher sind, wenn die initialen Gewichte mit der quadratischen Funktion bestimmt wurden. Auch an dieser Stelle ist allerdings zu berücksichtigen, dass diese Korrelationen noch nicht überzeugend für die Vorhersagekraft des PCS-Modells sprechen, da nicht für die Pretest-Werte kontrolliert wurde. Dies erfolgte jedoch in einem zweiten Schritt, nämlich bei der Berechnung von messwiederholten Regressionsanalysen (siehe Tabelle 16). Hierbei stellten wieder die Posttest-Validitäten das Kriterium und die Pretest-Validitäten sowie die finalen Aktivationen die Prädiktoren dar. Wurden auch die Posttest-Werte transformiert, wurden diese als Kriterium und die transformierten Pretest-Werte als weiterer Prädiktor (neben den finalen Aktivationen) verwendet. Erneut wurden aufgrund der Messwiederholung die Standardfehler für *n* (= Anzahl der Versuchspersonen) Cluster an Beobachtungen angepasst. Eine mittels GPower (Faul et al., 2009) durchgeführte Teststärkenanalyse für die Prädiktoren ergab eine akzeptable Teststärke von .85, um bei $N = 50$ und $\alpha = 0.05$ einen mittleren Effekt ($f^2 = .15$, Cohen, 1988) zu finden.

Tabelle 16: Ergebnisse der Regressionsanalysen

	Posttest		
	linear	quadratisch	
		a	b
Pretest	.84** (10.36)	.64** (6.51)	.76** (13.45)
PCS	.01 (0.07)	.24** (3.22)	.12* (2.32)
N_B	200	200	200
R^2	.72	.74	.76

Anmerkungen: Es werden die β-Koeffizienten berichtet, und zwar getrennt für Analysen auf Basis der linearen und der quadratischen Transformationsfunktion. Bei letzterer wurden die empirischen Posttest-Werte entweder auch transformiert (b) oder nicht (a). *t*-Werte stehen in Klammern. Aufgrund der Messwiederholung wurden cluster-robuste Standardfehler verwendet. N_B: Anzahl an Beobachtungen, *: $p < .05$, **: $p < .01$

Der PCS-Prädiktor leistet dann einen signifikanten Beitrag zur Vorhersage der Posttest-Validitäten, wenn die initialen Gewichte entsprechend der quadratischen Transformationsfunktion berechnet wurden. Hierbei macht es keinen gewichtigen Unterschied, ob die Analysen mit untransformierten oder mit transformierten Pretest- und Posttest-Validitäten berechnet wurden.

Diskussion

Mit der vorliegenden Simulationsstudie sollte überprüft werden, ob die im Experiment gefundenen Resultate durch das PCS-Modell vorhergesagt werden. Dies war teilweise der Fall. Erneut konnten Unterschiede zwischen den einzelnen Transformationsfunktionen, das heißt der linearen und der quadratischen Funktion, beobachtet werden. Es zeigte sich nämlich in messwiederholten ANOVAs, dass die auf Basis des PCS-Modells gewonnenen Simulationsergebnisse die im Experiment gefundenen Umwertungen dann vorhersagen können, wenn die initialen Gewichte unter Verwendung der quadratischen Transformationsfunktion berechnet wurden. Die höhere Adäquatheit der quadratischen Transformationsfunktion spiegelt sich auch bei der Vorhersage der Posttest-Validitäten durch den PCS-Prädiktor wider: Dieser leistet nur dann einen signifikanten Beitrag, wenn die initialen Gewichte gemäß der quadratischen Transformationsfunktion ermittelt wurden. Hierbei ist es irrelevant, ob die Analysen mit ebenfalls quadratisch transformierten Pretest- und Posttest-Validitäten

oder mit untransformierten Werten berechnet wurden. Die bereits in der Diskussion der Simulationsergebnisse der vorherigen Studie (Kapitel 3.6.3) geäußerte Vermutung, dass diese Funktion besser geeignet ist, um die empirischen Daten abzubilden, hat sich in diesem Experiment bestätigt.

3.6.5 Zusammenfassung und Diskussion

In diesem Kapitel sollen die Ergebnisse der Simulationsstudien zusammenfassend diskutiert werden. Die aus dem PCS-Modell abzuleitende Hypothese, dass in jedem Muster systematische Umwertungen auftreten, die mit dem PCS-Modell vorhergesagt werden können, konnte nicht bestätigt werden. Vielmehr spiegelt sich in den Simulationsergebnissen eher das empirische Ergebnismuster wider, nach dem deutliche Reevaluationen lediglich bei einer unausgeglichenen Informationslage zu beobachten sind. Einschränkend ist jedoch festzuhalten, dass die empirischen Ergebnisse tendenziell dann besser abgebildet werden, wenn die Übersetzung von subjektiven Cue-Validitäten in initiale Gewichte nach einer quadratischen Transformationsfunktion erfolgte. Des Weiteren stimmen die Simulationsergebnisse vereinzelt auch bei dieser Transformationsfunktion nicht mit den empirischen Ergebnissen überein. So lässt sich die signifikante Interaktion „Messzeitpunkt x Cue“, die bei der Simulation der Daten des vierten Experiments für das erste Muster gefunden wurde, nur schwer erklären. Insgesamt ist die Evidenz damit zwar schwach, weist aber in die Richtung der empirischen Ergebnisse. Die Berücksichtigung der Präferenz für Konsistenz führte hierbei zu keiner Veränderung des Ergebnismusters. Somit kann festgehalten werden, dass wenn empirisch systematische Umwertungen zu beobachten sind, diese durch PCS vorhergesagt werden können und damit die Annahme stabiler subjektiver Cue-Validitäten widerlegt wird. Dass sich aber der Einfluss der Informationslage auch in den Simulationsergebnissen zeigt, stützt die in Kapitel 3.5.4 geäüßerte Vermutung, dass diese in probabilistischen Inferenzentscheidungen relevant ist. Dies hätte aus dem PCS-Modell nicht vorhergesagt werden können. Erklärungsmöglichkeiten wurden bereits in der Diskussion in Kapitel 3.5.4 angesprochen. Die Simulationsergebnisse können diesbezüglich zur weiteren Klärung beitragen: So interpretieren Spellman und Holyoak (1992) niedrige Aktivationen als Zeichen von Spannung. In den vorliegenden Simulationsstudien wären nach dem Vorgehen von Spellman und Holyoak hierfür die Aktivationen der Optionen zu betrachten, da sie die Stützung der jeweiligen Option repräsentieren. Zwar ist nach Smith (1996) zu bedenken, dass Aktivationen einzelner Knoten stets im Zusammenhang mit dem gesamten Aktivati-

onsmuster zu betrachten sind, doch wenn das Ziel nicht in einer Beurteilung einzelner Aktivationen, sondern in einem relativen Vergleich derselben besteht, erscheint eine Betrachtung der Aktivationen gerechtfertigt. Deskriptiv zeigt sich in der Tat, dass die Aktivation der gewählten Option im zweiten Muster stets höher ist als die Aktivation der gewählten Option im ersten und im dritten Muster. Dies gilt analog, nur mit anderem Vorzeichen, für die nicht gewählte Option. Damit scheint nicht die Anzahl der Cues ausschlaggebend zu sein, sondern deren Verteilung. Einschränkend ist darauf hinzuweisen, dass die Unterschiede gering sind. Bei einer minimal bzw. maximal möglichen Aktivation von -1 bzw. +1 zeigen sie sich überwiegend ab der 3. Nachkommastelle. Da diese Unterschiede aber unabhängig von der Transformationsfunktion und bei beiden Datensätzen (von Experiment 3 und 4) zu beobachten sind, sollen sie – obige Einschränkung beachtend und daher mit Vorsicht – interpretiert werden. Dass die Aktivationen bei den unausgeglichenen Informationslagen im Vergleich zum ausgeglichenen Informationsmuster niedriger sind und gemäß Spellman und Holyoak (1992) von vergleichsweise mehr Spannung zeugen, spricht nämlich dafür, dass vergleichsweise mehr Inkonsistenz besteht. Damit kann das mögliche Argument entkräftet werden, dass die Versuchspersonen die Informationslage aus irgendeinem Grund als konsistent empfinden. Dass dennoch keine deutlichen Umwertungen erfolgt sind, um Konsistenz zu maximieren, könnte auf eine der in Kapitel 3.5.4 geäußerten Erklärungen zurückzuführen sein. Diese Vermutungen sollten natürlich einer empirischen Testung unterzogen werden. Entsprechende Vorschläge werden in Kapitel 4 vorgestellt. Die tendenziell höhere Vorhersagekraft der Simulationsergebnisse unter Verwendung einer quadratischen Transformationsfunktion gibt einen wichtigen Anhaltspunkt für die Übersetzung subjektiv eingeschätzter Cue-Validitäten in initiale Simulationsgewichte. Weitere Studien, die sich der Untersuchung der Eignung verschiedener Funktionen widmen, sind wünschenswert. Auch individuelle Transformationsfunktionen sind denkbar. Dass die Berücksichtigung des PFC keinen Mehrwert bezüglich der Vorhersagekraft der Simulationsergebnisse erwirkte, spricht gegen einen Einfluss desselben auf die Umwertung von Cue-Validitäten. Zwar ist das gewählte Vorgehen dahingehend zu kritisieren, dass die Ausprägung dieses Faktors nur vergröbert, nämlich in zwei Stufen, als Modellparameter umgesetzt wurde, doch erscheint es vor dem Hintergrund des gescheiterten Replikationsversuchs fraglich, ob eine feinere Abstufung zu einem anderen Ergebnis geführt hätte. Dem möglichen prinzipiellen Einwand, dass die selbstständige Festlegung der Parameter die Gefahr einer gewissen Willkür bei der Durchführung von Simulationen birgt, sei mit

Thagard (1989) entgegnet: “However, if a fixed set of default parameters applies to a large range of cases, then the arbitrariness is much diminished” (S. 443). Dieses Vorgehen liegt den dargestellten Simulationsstudien zugrunde, wobei sich die Wahl der Parameter – wie in Kapitel 3.6.1 dargestellt – an früheren Studien von Glöckner et al. (im Druck) orientierte. Dies schließt aber keine begründete Variation der Parameter, wie sie auch bezüglich des PFC vorgenommen wurde, aus.

3.7 DER EINFLUSS DER WIEDERHOLUNG DER ENTSCHEIDUNGSSITUATION AUF DIE VERÄNDERUNG SUBJEKTIVER CUE-VALIDITÄTEN

Nachdem in den vorherigen Experimenten und Simulationsstudien gezeigt werden konnte, dass die Informationslage einen Einfluss auf die Stärke prädezisionaler Umwertungen ausübt, wurde in den vorzustellenden Experimenten 6 und 7 ein weiterer möglicher Einflussfaktor untersucht. So wurde erneut die Frage aufgegriffen, welchen Einfluss die Wiederholung der Entscheidungssituation auf diese Umwertungen ausübt. Wie bereits bei der Darstellung von Experiment 2 erläutert wurde, stellt die Erweiterung des Designs von singulären zu wiederholten probabilistischen Inferenzentscheidungen gerade im Hinblick auf die Befunde aus dem Bereich der Präferenzentscheidungen eine interessante Neuerung dar. So konnten Hoeffler und Ariely (1999) zeigen, dass das wiederholte Treffen von Präferenzentscheidungen zu einer Stabilisierung der Präferenz führt: Personen gewinnen zunehmend größeres Vertrauen in die Wichtigkeit der von ihnen verwendeten Attribute (erkennbar an einem kleiner werdenden Konfidenzintervall). Zu untersuchen ist, welchen Einfluss die Wiederholung der Entscheidungssituation bei probabilistischen Inferenzentscheidungen hat. Anhand der Daten des zweiten Experimentes konnte diese Frage nicht zufriedenstellend beantwortet werden, da weder bei der einmaligen noch bei der wiederholten Darbietung einer Entscheidungssituation Umwertungen beobachtet werden konnten. Dies ließ sich – wie die folgenden Experimente zeigten – darauf zurückführen, dass ein ausgeglichenes Informationsmuster verwendet worden war. Folglich sollte in den Experimenten 6 und 7 der Einfluss der Wiederholung unter Verwendung eines anderen Informationsmusters, das starke Umwertungen zulässt, untersucht werden. Aufgrund der bisherigen Ergebnisse fiel die Wahl auf ein 1:3-Muster.

3.7.1 Experiment 6: Die Veränderung von Cue-Validitäten bei wiederholter im Vergleich zu einmaliger Darbietung der Entscheidungssituation

ZIELSETZUNG

In dem vorliegenden Experiment wurde – wie einleitend bereits erwähnt – noch einmal die Frage aufgegriffen, welchen Einfluss die wiederholte im Vergleich zur einmaligen Darbietung von Entscheidungssituationen auf die Stärke der Veränderung subjektiver Cue-Validitäten hat.

METHODE

Design. Die Veränderung der subjektiven Cue-Validitäten wurde entsprechend dem bisherigen Vorgehen *within subjects* in einem Pre-/Post-Design gemessen. Als *between subjects*-Faktor wurde die Anzahl an Entscheidungssituationen variiert. Während den Versuchspersonen der ersten Bedingung lediglich eine Entscheidungssituation präsentiert wurde (einmalige Darbietung), wurden denjenigen der zweiten Bedingung in Anlehnung an Hoeffler und Ariely (1999) zwölf Entscheidungssituationen präsentiert (wiederholte Darbietung). Es wurden tendenzielle Entscheidungen, die entweder vor oder nach der Bearbeitung des Posttests abgegeben wurden, und finale Entscheidungen erfasst.

Hypothesen. Entsprechend dem Prozess der automatischen Konsistenzmaximierung wurde für die erste Bedingung erwartet, dass die Cues, die für die spätere Entscheidung sprechen, aufgewertet werden, während die anderen Cues abgewertet werden. Obgleich im PCS-Modell die Konsistenzmaximierung nur für einmalige Entscheidungssituationen beschrieben wird, erschien es vor diesem theoretischen Hintergrund plausibel anzunehmen, dass die wiederholte Darbietung einer identisch konstruierten Entscheidungssituation ebenfalls zu systematischen Umwertungen führt (siehe Kapitel 3.4). Der Einfluss der Wiederholung des Cue-Musters auf die Stärke der Veränderung der Cue-Validitäten wurde explorativ untersucht.

Material. Im vorliegenden Experiment wurden die Versuchspersonen mit einer Entscheidungssituation konfrontiert, in der basierend auf verschiedenen Fakten bzw. Cues über die Schuld bzw. Unschuld eines Verdächtigen entschieden werden sollte. Diese Entscheidungssituation sowie die entsprechenden Fakten sind an den *Jason Wells Case* von Simon, Snow et al. (2004) angelehnt, in dem ein Mitarbeiter verdächtigt wird, Geld aus dem

Firmensafe gestohlen zu haben.[49] Die entsprechenden Cues bzw. Fakten wurden in eine neue Cover-Story eingebettet und gemäß der Zielsetzung des vorliegenden Experimentes in ihrem Inhalt und der Formulierung deutlich modifiziert. Während die Versuchspersonen bei Simon und Kollegen einzuschätzen hatten, wie sehr sie bestimmten Aussagen zustimmen, wurde im vorliegenden Experiment eine Einschätzung der Vorhersagekraft einzelner Fakten für die Schuld oder Unschuld eines Verdächtigen erbeten. Die Versuchspersonen wurden sowohl im Pre- als auch im Posttest instruiert, die Validität von acht Fakten (Cues) für die Schuld bzw. Unschuld eines Verdächtigen möglichst gut und differenziert einzuschätzen. Das Validitätskonzept wurde wie folgt erklärt: „Die Validität eines Faktes ist definiert als die Wahrscheinlichkeit, dass bei Vorliegen des Faktes ein Verdächtiger eher schuldig (unschuldig) ist als unschuldig (schuldig).“ Die Versuchspersonen schätzten die Validitäten auf einer Skala, die von 0% schuldig bis 100% unschuldig bzw. von 100% unschuldig bis 0% unschuldig reichte, ein (siehe Abbildung 10).

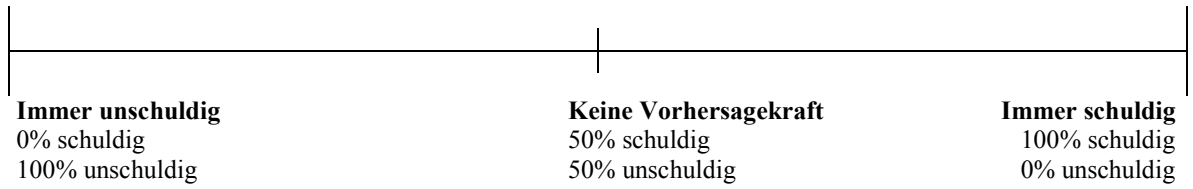


Abbildung 10: Skala zur Einschätzung der subjektiven Cue-Validitäten

Mit der Positionierung des Kreuzes auf der linken oder rechten Seite gaben die Versuchspersonen damit an, ob der jeweilige Fakt für sie eher für die Schuld oder für die Unschuld des Verdächtigen sprach. Die Schuld- und die Unschuldskala reichten folglich wie in den vorherigen Untersuchungen von 50% bis 100%, da ein Fakt bzw. Cue mit einer Validität von 50% keine Vorhersagekraft hat. Hierbei waren die Fakten so formuliert worden, dass ein Cue („Augenzeuge“) für die Schuld des Verdächtigen sprach (*Schuld-Cue*), während die sieben anderen Cues für die Unschuld des Verdächtigen sprachen (*Unschuld-Cues*). In Tabelle 17 sind die jeweiligen Fakten, deren Validität einzuschätzen war, sowie die im Folgenden verwendeten Abkürzungen aufgelistet.

[49] Herzlicher Dank an Dan Simon und Douglas Stenstrom für die Zusendung des Materials.

Tabelle 17: Die einzelnen Fakten sowie ihre Abkürzung

Fakt	Abkürzung
Ein Augenzeuge hat den Verdächtigen als die Person identifiziert, die er vom Tatort hat wegeilen sehen.	*Augenzeuge*
Der Verdächtige gilt als fleißiger und guter Mitarbeiter.	*Arbeitsweise*
Der Verdächtige hat finanzielle Ausgaben getätigt, die in etwa dem Betrag des gestohlenen Geldes entsprechen. Er kann aber belegen, woher er das Geld für diese Ausgaben hat.	*Ausgaben*
Der Verdächtige ist noch nie mit dem Gesetz in Konflikt geraten.	*Gesetz*
Eine Analyse ergibt, dass der einzige Fingerabdruck, der am Tatort gefunden wurde, nicht mit dem Fingerabdruck des Verdächtigen übereinstimmt.	*Fingerabdruck*
Der Verdächtige hat etwa fünf Minuten nach der Tat einen Anruf von einem Ort aus gemacht, der ca. 30 Minuten weit entfernt vom Tatort ist.	*Anruf*
Wenige Minuten nach dem Diebstahl wurde von der Überwachungskamera gefilmt, wie ein Auto rasant den Firmenparkplatz verlässt. Der Verdächtige fährt nicht das gleiche Auto, das von der Überwachungskamera gefilmt wurde.	*Auto*
Der Verdächtige wurde etwa 20 Minuten nach der Tat an einem Ort gesehen, der etwa 40 Minuten vom Tatort entfernt liegt.	*Ort*

Für das vorliegende Experiment wurden dann verschiedene Fälle konstruiert, in denen stets ein Mitarbeiter des Diebstahls verdächtigt wurde und seine Schuld bzw. Unschuld auf Basis verschiedener Fakten einzuschätzen war. Als Treatment wurden die Versuchspersonen mit einem Fall (Bedingung 1) bzw. mit zwölf voneinander unabhängigen Fällen (Bedingung 2) von Diebstahl in verschiedenen Unternehmen konfrontiert. Es wurden sowohl der Fall geschildert als auch die Fakten präsentiert. Hierbei sprach nun jeweils ein Fakt bzw. Cue für die Schuld des Verdächtigen, während drei andere Fakten bzw. Cues für die Unschuld des Verdächtigen sprachen. In der ersten Bedingung sprach der Cue „Augenzeuge“ für die Schuld, während die Cues „Anruf“, „Auto“ und „Ort“ für die Unschuld sprachen. In der zweiten Bedingung trat der Cue „Augenzeuge“ in jedem der zwölf Fälle auf. Des Weiteren waren die Fakten so zusammengestellt, dass der Fakt „Ort“ ebenfalls in allen Fällen auftrat, die Fakten „Auto“ und „Arbeitsweise“ in sechs Fällen und die Fakten „Anruf“, „Ausgaben“, „Fingerabdruck“ und „Gesetz“ in drei Fällen. Somit blieb das

Informationsmuster stets konstant, da immer ein Cue („Augenzeuge“) für die Schuld des Verdächtigen sprach und drei Cues für die Unschuld des Verdächtigen sprachen, während die Cue-Inhalte von Fall zu Fall wechselten. Der Fakt Augenzeuge war jeweils dreimal an erster, zweiter, dritter und vierter Position (zufällige Zuteilung). Die Position der anderen Fakten war ebenso wie die Reihenfolge der Fälle randomisiert. Der Fall, der in Bedingung 1 dargeboten wurde, war ebenfalls unter den in Bedingung 2 präsentierten Fällen. Das Material sowie die Instruktionen können Anhang A-7 und A-8 entnommen werden.

Stichprobe. An dem Experiment nahmen 75 Versuchspersonen (35 männlich, 39 weiblich, 1 ohne Angabe; Durchschnittsalter: 25.68 Jahre, *SD* = 7.42), hauptsächlich Studenten aus Bonn, teil. Sie wurden über das Online-Rekrutierungssystem *ORSEE* (Greiner, 2004) sowie mündliche Mitteilung angeworben. Das Experiment, das papierbasiert durchgeführt wurde, war Teil einer etwa einstündigen Experimentalbatterie. Auf das vorliegende Experiment entfielen je nach Bedingung 15 (Bedingung 1) oder 35 Minuten (Bedingung 2). Als Aufwandsentschädigung erhielten die Versuchspersonen 12 € pro Stunde. Sie wurden den beiden Bedingungen zufällig, aber mit dem Ziel einer Gleichverteilung zugeteilt.

Durchführung. Das Experiment gliederte sich in vier Teile: Nach einer Erläuterung des Validitätskonzepts schätzten die Versuchspersonen im ersten Teil, dem Pretest, die Validität der Fakten (Cues) „Augenzeuge“, „Arbeitsweise“, „Ausgaben“, „Gesetz“, „Fingerabdruck“, „Anruf“, „Auto“ und „Ort“ auf der vorgestellten Skala ein. Es folgte ein thematisch unverbundenes Experiment, das etwa 15-20 Minuten dauerte. Im anschließenden zweiten Teil des Experimentes erfolgte das Treatment: Die Versuchspersonen der ersten Bedingung wurden instruiert sich vorzustellen, sie arbeiteten als Detektiv bei einer Wirtschaftsdetektei. Diese Detektei decke Fälle von Wirtschaftskriminalität auf, wobei sich die Versuchsperson auf Fälle von Diebstahl spezialisiert habe. Der Versuchsperson läge nun ein Fall eines Unternehmens vor, in dem es zu einem Diebstahl gekommen sei. Die Versuchsperson sei damit beauftragt worden, den Täter zu finden, und habe daher in den letzten Wochen und Monaten recherchiert. Die Ermittlungen hätten sich auf einen Mitarbeiter konzentriert, wobei einige Fakten für und einige gegen die Schuld des Mitarbeiters sprächen. Die Versuchspersonen wurden mit dem Hinweis, dass sie noch eine DNA-Analyse abwarten wollten, gebeten, noch keine endgültige Entscheidung bzgl. der Schuld oder Unschuld des Mitarbeiters zu treffen, sondern lediglich zu überlegen, ob sie ihn aufgrund der vorliegenden Informationen tendenziell eher für schuldig oder tendenziell eher

für unschuldig halten. Im Anschluss an die Cover-Story lasen die Versuchspersonen die Schilderung des Diebstahls und die Fakten. Etwa die Hälfte der Versuchspersonen gab ihre tendenzielle Entscheidung an. An das Treatment schloss als dritter Experimentalteil der Posttest an: Alle Versuchspersonen schätzten erneut die Validitäten der acht Fakten ein. Im letzten Teil des Experimentes wurde eine endgültige Entscheidung getroffen. Alle Versuchspersonen wurden unter dem Vorwand, der Auftraggeber wolle eine sofortige Entscheidung, gebeten, ihre endgültige Entscheidung zu treffen. Außerdem gaben zuvor die Versuchspersonen, die noch keine Tendenz abgegeben hatten, diese an.

Der Experimentalablauf war für Versuchspersonen der zweiten Bedingung nahezu identisch. Der wichtige Unterschied bestand im Treatment, in dem sie nicht nur mit einem Fall, sondern mit insgesamt zwölf Fällen konfrontiert wurden. Hierbei wurden sie darauf aufmerksam gemacht, dass die einzelnen Fälle unabhängig voneinander sind. Etwa die Hälfte gab ihre Tendenz jeweils nach dem Lesen der Informationslage an, während die andere Hälfte dies erst gemeinsam mit der Angabe der endgültigen Entscheidung in der Nachbefragung machte.

ERGEBNISSE

Vorausgehende Analysen. Von den 75 Versuchspersonen, die an dem Experiment teilgenommen hatten, mussten fünf Personen aufgrund von Fehlern im Versuchsablauf ausgeschlossen werden. Weitere zwei Teilnehmer wurden in den folgenden Analysen nicht berücksichtigt, da sie durchgängig Validitätseinschätzungen von 50% oder 100% machten. Von den verbleibenden 68 Versuchspersonen (31 männlich, 36 weiblich, 1 ohne Angabe; Durchschnittsalter: 25.57 Jahre, SD = 7.60) waren 34 zufällig der ersten und 34 der zweiten Bedingung zugeteilt worden.

Interpretation der Informationen. Wie bereits angesprochen, waren die Cues so formuliert worden, dass ein Cue, nämlich der Cue „Augenzeuge“, für die Schuld des Verdächtigen sprach, während die übrigen sieben Cues für die Unschuld des Verdächtigen sprachen. Anhand der von den Versuchspersonen vorgenommenen Einschätzung auf der dafür vorgesehenen Skala konnte nicht nur die subjektive Validität der einzelnen Cues bestimmt werden, sondern es konnte auch überprüft werden, inwiefern die Teilnehmer die Cues tatsächlich als Schuld-Cue bzw. als Unschuld-Cues wahrnahmen. Es wurde eine Klassifikation der Cues vorgenommen, die auf der absoluten Mehrheit der Richtungsangabe basierte. So konnten die Cues danach unterteilt werden, ob sie überwiegend als Schuld-Cue

oder als Unschuld-Cue interpretiert wurden. Zwei weitere Möglichkeiten bestanden darin, dass die Cues am ehesten keine Vorhersagekraft hatten oder aber nicht klassifizierbar waren. Die jeweiligen Häufigkeitsangaben sowie die Klassifikation sind den Tabellen C-1 und C-2 des Anhangs C zu entnehmen. Während die Cues „Augenzeuge“, „Fingerabdruck“, „Anruf“, „Auto“ und „Ort“ stets von der überwiegenden Zahl der Versuchspersonen wie intendiert gedeutet wurden, waren die übrigen Cues „Arbeitsweise“, „Ausgaben“ und „Gesetz“ entweder nicht klassifizierbar oder sie hatten für die Versuchspersonen am ehesten keine Vorhersagekraft. Berechnet man für den Pretest über beide Bedingungen hinweg die prozentuale Häufigkeit korrekter (im Sinne von intendierter) Richtungsangaben, so zeigt sich eine Aufteilung der Cues in drei Gruppen: Mit 88.20%, 75.00%, 73.50% und 83.80% wurden die Cues „Augenzeuge“, „Fingerabdruck“, „Anruf“ und „Ort“ besonders häufig wie intendiert interpretiert. Die Cues „Auto“ und „Ausgaben“ wurden mit 55.90% und 47.1% etwa zur Hälfte korrekt gedeutet. Die Cues „Arbeitsweise“ (42.60%) und „Gesetz“ (41.20%) wurden hingegen eher selten wie erwartet interpretiert. Betrachtet man die Inhalte der Cues genauer, so lassen sich post hoc diese drei Gruppen folgendermaßen beschreiben: Die erste Kategorie umfasst tatnahe Cues, die den Verdächtigen eindeutig be- oder entlasten. Es handelt sich um die Cues „Augenzeuge“, „Fingerabdruck“, „Anruf“ und „Ort“, die besonders häufig wie intendiert interpretiert wurden. In die zweite Gruppe sind die Cues „Auto“ und „Ausgaben“ einzuordnen, die zwar ebenfalls tatnah sind, aber den Verdächtigen verglichen mit den vorherigen Cues weniger eindeutig entlasten. Dies hat sich in weniger korrekten Zuordnungen gezeigt. In die dritte Kategorie fallen schließlich die Cues „Arbeitsweise“ und „Gesetz“, die den Verdächtigen zwar entlasten, aber als eher tatfern zu bezeichnen sind. Dies könnte erklären, weshalb diese Cues relativ selten wie intendiert gedeutet wurden. Obgleich diese Klassifikation lediglich post hoc durchgeführt wurde, gibt sie doch einen Anhaltspunkt dafür, weshalb die Cues zum Teil nicht wie intendiert interpretiert wurden.

Die Cues, deren Richtungsangabe nicht mit der intendierten übereinstimmte, wurden umkodiert, indem die angegebene Validität von 100% subtrahiert wurde. Ein Überblick über die mittleren Validitätseinschätzungen – separat für beide Bedingungen und Messzeitpunkte – findet sich in Tabelle B-9 in Anhang B. Bevor die Veränderung der Cue-Validitäten untersucht wird, soll geklärt werden, ob die Angabe der tendenziellen Entscheidung vor oder nach der zweiten Einschätzung der Cue-Validitäten einen Einfluss auf die Veränderung derselben hat.

Zeitpunkt der Entscheidungstendenz. In Anlehnung an Simon et al. (2001) wurde variiert, ob die Versuchspersonen ihre tendenzielle Entscheidung vor oder nach der zweiten Einschätzung der Cue-Validitäten trafen. So war pro Bedingung zu überprüfen, ob die Veränderung der Cue-Validitäten, gemessen an der Differenz zwischen dem Pretest- und dem Posttest-Wert, je nach Zeitpunkt der Angabe der tendenziellen Entscheidung unterschiedlich ist. Da dieser zusätzliche Gruppierungsfaktor dazu führt, dass pro Bedingung zwei Stichproben, deren Größe kleiner als 30 war, verglichen wurden, wurden entsprechend der Empfehlung von Rasch et al. (2006b) Mann-Whitney-U-Tests durchgeführt. Bei beiden Bedingungen wurde bei keinem Cue ein signifikantes Ergebnis ($\alpha = 5\%$) erzielt, so dass gefolgert werden kann, dass der Zeitpunkt der Angabe der tendenziellen Entscheidung keinen Einfluss auf die Veränderung der subjektiven Cue-Validitäten hat (vgl. Simon et al., 2001). In den folgenden Analysen wurden daher beide Gruppen zusammengefasst.

Tendenzielle und finale Entscheidungen. Die Entscheidungssituationen waren so konstruiert worden, dass die Informationslage für die Unschuld des Verdächtigen sprach. Da aber einige Cues nicht wie intendiert interpretiert wurden, sind vermehrt Entscheidungen für die Schuld des Verdächtigen denkbar. In Bedingung 1 entschieden sich eine Person tendenziell für schuldig und 33 Personen tendenziell für unschuldig. Final entschieden sich zwei Personen für schuldig und 32 Personen für unschuldig. Somit wechselte in Bedingung 1 eine Person ihre Entscheidung. Tabelle 18 gibt einen Überblick über die in Bedingung 2 getroffenen tendenziellen und finalen Entscheidungen sowie über die Anzahl der Personen, deren tendenzielle und finale Entscheidung unterschiedlich war (Switcher).

Tabelle 18: Tendenzielle und finale Entscheidungen bei wiederholter Darbietung der Entscheidungsaufgabe

		Fall											
		1	2	3	4	5	6	7	8	9	10	11	12
T	schuldig	6	1	2	1	10	7	0	3	2	4	6	6
	unschuldig	27	33	32	32	23	25	34	30	30	28	26	26
E	schuldig	4	5	3	4	10	8	0	5	4	7	10	4
	unschuldig	30	29	31	30	24	26	34	29	30	26	24	30
Switcher		4	4	1	6	1	5	0	5	7	8	5	10

Anmerkungen: T = tendenzielle Entscheidung. E = finale Entscheidung. Aufgrund von Fehlern in der Bearbeitung konnten die tendenziellen Entscheidungen nicht von allen Teilnehmern berücksichtigt werden.

Obgleich bei jedem Fall mehr Entscheidungen für die Unschuld als für die Schuld des Verdächtigen getroffen wurden, ist die Zahl der schuldig-Urteile doch höher als erwartet. Da die Hypothesen bezüglich der Umwertung des Schuld-Cues und der Unschuld-Cues auf der Annahme eines unschuldig-Urteils basieren, gilt es die getroffene Entscheidung in den späteren Analysen zu berücksichtigen. Bei einigen Versuchspersonen unterschied sich die tendenzielle von der finalen Entscheidung. Während insgesamt 346 tendenzielle Entscheidungen für die Unschuld des Verdächtigen getroffen wurden, wurden 343 finale Entscheidungen für die Unschuld getroffen. Das bedeutet, dass das Urteil annähernd gleich häufig von Schuld zu Unschuld und von Unschuld zu Schuld gewechselt wurde.

Veränderung der subjektiven Cue-Validitäten. Ob subjektive Cue-Validitäten systematisch umgewertet wurden, wurde erneut auf der Ebene der Optionen und auf der Ebene der Cues analysiert. Für die Analyse auf der Ebene der Optionen wurden zunächst die korrigierten gewichteten Summen für den präsentierten Fall (Bedingung 1) bzw. für die zwölf präsentierten Fälle (Bedingung 2) berechnet. Dies erfolgte jeweils für den Pretest sowie für den Posttest. Daraufhin wurden für beide Messzeitpunkte die Differenzen zwischen den korrigierten gewichteten Summen berechnet, wobei für die Berechnung die final getroffene Entscheidung zugrunde gelegt wurde. Zunächst wurde der Frage nachgegangen, ob in beiden Bedingungen ein *spreading apart* der Alternativen stattgefunden hat. Hierzu wurde für Bedingung 1 ein gepaarter *t*-Test berechnet, anhand dessen geprüft werden konnte, ob sich die Differenz zwischen den gewichteten Summen vom Pre- zum Posttest vergrößerte. Eine post-hoc durchgeführte Teststärkenanalyse mittels GPower (Faul et al., 2007) ergab, dass bei $n = 34$ und $\alpha = .05$ ein Effekt d_z der Größe .55 mit einer Teststärke von .93 gefunden werden kann.[50] Es zeigte sich ein signifikantes Ergebnis, $t(33) = 1.73$, $p < .05$ (einseitig), $d_z = .30$. Für die Analyse für Bedingung 2 wurde eine messwiederholte ANOVA mit Messwiederholung auf den Faktoren „Messzeitpunkt“ und „Fall“ berechnet. Eine post-hoc mit GPower (Faul et al., 2007) durchgeführte Teststärkenanalyse für den interessierenden Haupteffekt „Messzeitpunkt“ zeigte, dass bei $\alpha = .05$ und $n = 33$ ein mittlerer Effekt ($f = .25$, Cohen, 1988) mit einer Teststärke von .77 gefunden werden kann.[51] Der Haupteffekt „Messzeitpunkt“ wurde wider Erwarten nicht signifikant, $F(1, 32) < 1$, $p = .78$, $\eta_p^2 = .00$. Somit konnte über alle Fälle hinweg kein signifikantes *spreading apart* verzeichnet werden.

[50] Bei der Berechnung der Teststärke wurde der Tatsache Rechnung getragen, dass die Konventionen von Cohen (1988) für unabhängige Stichproben gelten und bei der Festlegung entsprechender Effektstärken für abhängige Stichproben stets die Korrelation zu beachten ist.

[51] Eine Versuchsperson gab nicht für jeden Fall eine Entscheidung ab und wurde ausgeschlossen.

Deskriptiv zeigte sich sogar über alle Fälle eine Abnahme der Differenz zwischen den korrigierten gewichteten Summen. Der Haupteffekt „Fall“ wurde hingegen signifikant, $F(2.92, 93.41) = 4.02$, $p < .05$, $\eta_p^2 = .11$.[52] Folglich unterschieden sich die Fälle hinsichtlich der Höhe der über beide Messzeitpunkte gemittelten Differenz zwischen den korrigierten gewichteten Summen. Dies lässt sich plausibel mit der variierenden Zusammensetzung der Cues, die ihrerseits zum Teil sehr unterschiedlich in ihrer Validität eingeschätzt wurden, erklären. Es zeigte sich keine signifikante Interaktion „Messzeitpunkt x Fall“, $F(3.72, 118.96) = 1.45$, $p = .23$, $\eta_p^2 = .04$.[53]

Für die Analyse auf der Ebene der Cues wurden in einem ersten Schritt anhand von Boxplots wiederholte extreme Ausreißer pro Bedingung ermittelt. In Bedingung 1 gab es einen wiederholten extremen Ausreißer, in Bedingung 2 konnten fünf wiederholte extreme Ausreißer identifiziert werden. Es wurden dann getrennt für jede Bedingung messwiederholte ANOVAs berechnet. Für die erste Bedingung waren „Messzeitpunkt“ und „Cue“ die Innersubjektfaktoren, wobei als Cues die tatsächlich im Treatment präsentierten Cues berücksichtigt wurden, nämlich „Auto“, „Augenzeuge“, „Anruf“ und „Ort“. Der Haupteffekt „Messzeitpunkt“ wurde nicht signifikant, $F(1, 33) < 1$, $p = .57$, $\eta_p^2 = .01$, was mit der gleichzeitig auftretenden Auf- und Abwertung von Cues zu erklären ist. Signifikant wurden hingegen der auf die unterschiedliche Einschätzung der Cues zurückzuführende Haupteffekt „Cue“, $F(3, 99) = 8.64$, $p < .001$, $\eta_p^2 = .21$, sowie die interessierende Interaktion „Messzeitpunkt x Cue“, $F(2.65, 87.37) = 3.36$, $p < .05$, $\eta_p^2 = .09$[54]. Die Teststärke – berechnet mittels GPower (Faul et al., 2007) –, bei $n = 34$ und $\alpha = .05$ sowie unter Berücksichtigung der Sphärizitätskorrektur einen mittleren Effekt ($f = .25$, Cohen, 1988) zu finden, liegt nur bei .77. Um diese Interaktion näher zu untersuchen, wurde anhand von gepaarten t-Tests für jeden Cue berechnet, ob die Umwertung der Validitäten signifikant war. Dies war lediglich bei dem Cue „Augenzeuge“ der Fall, $t(33) = -2.98$, $p < .01$, $d_z = -.51$, der wie erwartet signifikant abgewertet wurde. Zwei der anderen Cues, nämlich „Anruf“ und „Ort“, wurden zumindest tendenziell wie erwartet aufgewertet, während sich die Einschätzung der Validität des Cues „Auto“ nahezu nicht veränderte.[55] Für die zweite Bedingung wurde ebenfalls eine messwiederholte ANOVA mit „Messzeitpunkt“ und „Cue“ als Innersubjektfaktoren berechnet. Der Haupteffekt „Messzeitpunkt“ wurde ebenso signifikant, $F(1, 33) = 12.02$, $p <$

[52] Die Korrektur der Freiheitsgrade erfolgte nach Greenhouse-Geisser.
[53] Die Korrektur der Freiheitsgrade wurde erneut nach dem Verfahren von Greenhouse-Geisser vorgenommen.
[54] Aufgrund einer Verletzung der Sphärizitätsannahme wurde eine Korrektur der Freiheitsgrade vorgenommen. Diese erfolgte nach Huynh-Feldt.
[55] Unter Ausschluss des einen wiederholten extremen Ausreißers zeigte sich das gleiche Ergebnismuster.

.01, η_p^2 = .27, wie der Haupteffekt „Cue“, $F(5.44, 179.61) = 24.51$, $p < .001$, $\eta_p^2 = .43$[56]. Von besonderem Interesse war die Interaktion „Messzeitpunkt x Cue“. Eine Teststärkenanalyse mittels GPower (Faul et al., 2007) ergab, dass ein Anlegen von $\alpha = .05$ bei $n = 34$ die Entdeckung eines mittleren Effektes ($f = .25$, Cohen, 1988) mit einer Teststärke von .68 erlaubt, wobei die Sphärizitätskorrektur berücksichtigt wurde. Diese Interaktion wurde nicht signifikant, $F(5.73, 188.93) = 1.84$, $p = .10$, $\eta_p^2 = .05$.[57] Dass der Haupteffekt „Messzeitpunkt“ signifikant geworden ist, deutet darauf hin, dass die Umwertungen der Cues vorwiegend in eine Richtung gehen. Diese Interpretation wird gestützt durch die fehlende Interaktion von Messzeitpunkt und Cue. Um dieses Ergebnis zu untersuchen, wurden für jeden Cue gepaarte *t*-Tests ($df = 33$) berechnet. Die Resultate sind Tabelle 19 zu entnehmen.

Tabelle 19: Ergebnisse der *t*-Tests für jeden Cue bei wiederholter Darbietung der Entscheidungsaufgabe

Cue	Mittlere Differenz	*SD*	*t*-Wert	*p*	d_z
Augenzeuge	-6.08	11.87	-2.99	< .01	-.51
Arbeitsweise	-1.58	14.07	-0.66	.52	-.11
Ausgaben	-11.43	20.32	-3.28	< .01	-.56
Gesetz	0.55	11.41	0.28	.78	.05
Fingerabdruck	-5.59	19.34	-1.69	.10	-.29
Anruf	-0.92	25.13	-0.21	.83	-.04
Auto	-5.00	22.23	-1.31	.20	-.22
Ort	-10.15	23.70	-2.50	< .05	-.43

Anmerkungen: Die Differenz wurde berechnet, indem der Pretest-Wert vom Posttest-Wert subtrahiert wurde. Negative Werte zeigen folglich eine Abwertung, positive eine Aufwertung an.

Von acht Cues wurden drei Cues signifikant abgewertet, wobei diese Abwertung lediglich für den Cue „Augenzeuge“ erwartet worden war. Vier weitere Cues wurden zumindest tendenziell abgewertet und damit nicht in die erwartete Richtung umgewertet.[58] Es zeigten sich damit weder auf der Optionen- noch auf der Cue-Ebene *systematische* Umwertungen.

[56] Die Korrektur der Freiheitsgrade wurde nach Huynh-Feldt vorgenommen.
[57] Da $\varepsilon > .75$, wurden die Freiheitsgrade nach Huynh-Feldt korrigiert.
[58] Unter Ausschluss der fünf wiederholten extremen Ausreißer blieb das Ergebnismuster im Wesentlichen bestehen. Es wurde nun aber auch der Cue „Auto“ signifikant abgewertet.

Unterstützt werden die Ergebnisse, wenn lediglich der Fall, der in beiden Bedingungen dargeboten wurde, betrachtet wird. Für die erste Bedingung hatte sich ein signifikantes *spreading apart* der Optionen gezeigt. Führt man die entsprechende Analyse für die zweite Bedingung und nur unter Berücksichtigung des „gemeinsamen" Falls durch, zeigt sich kein signifikanter Effekt, $t(33) = 0.91$, $p = .18$ (einseitig), $d_z = .16$. Es liegt somit ein Unterschied zwischen beiden Bedingungen vor. Die Unterschiede auf der Cue-Ebene für diesen Fall ergeben sich aus der obigen Darstellung: Während in der ersten Bedingung eine signifikante und systematische Abwertung des Cues „Augenzeuge" stattfand, wurde in der zweiten Bedingung erwartungskonträr zudem der Cue „Ort" signifikant abgewertet.

Zusammengefasst traten in der ersten Bedingung systematische Umwertungen auf, obgleich diese auf der Cue-Ebene nicht stark ausgeprägt waren, während die Umwertungen in der zweiten Bedingung weitestgehend von einer mindestens tendenziellen Abwertung einiger Cues gekennzeichnet waren.

Berücksichtigung der „schuldig"-Entscheidungen. Wie bereits angesprochen, gilt es die relativ häufige Zahl der tendenziellen und finalen Entscheidungen für schuldig zu bedenken. Folglich wurden die Analysen noch einmal unter Berücksichtigung dieses Entscheidungsverhaltens durchgeführt. Bei der ersten Bedingung wurden stets die Personen ausgeschlossen, die sich tendenziell bzw. final für schuldig entschlossen hatten. Bei der zweiten Bedingung wurde zunächst pro Versuchsperson die Anzahl tendenzieller und finaler Entscheidungen für schuldig aufsummiert. Um auf der einen Seite eine Verzerrung der Ergebnisse durch Entscheidungen für die Schuld des Verdächtigen zu verhindern und auf der anderen Seite nicht zu viele Versuchspersonen ausschließen und damit eine Verringerung der Teststärke akzeptieren zu müssen, wurde ein *Cutoff* gesetzt: Überschritt die Anzahl der tendenziellen bzw. finalen Entscheidungen den Wert 2, so wurde die entsprechende Versuchsperson ausgeschlossen. Im Folgenden soll lediglich auf die Ergebnisse eingegangen werden, die sich von den ursprünglichen Resultaten unterscheiden. Eine tabellarische Gegenüberstellung der Ergebnisse findet sich in den Tabellen C-3, C-4 und C-5 in Anhang C. Hierbei wurde auf die interessierenden Effekte und die signifikanten *t*-Tests fokussiert. In Bedingung 1 wird das *spreading apart* der Optionen knapp nicht mehr signifikant ($p = .06$, einseitig), wenn man nur die Personen berücksichtigt, die als tendenzielle Entscheidung unschuldig angegeben haben. Zu bedenken ist, dass hier die Person, die ihr Urteil von unschuldig zu schuldig wechselte, eingeschlossen ist. Berücksichtigt man nämlich nur die Personen, die als finale Entscheidung unschuldig angegeben haben, wird das *spreading*

apart wieder signifikant. In Bedingung 2 wird die Interaktion „Messzeitpunkt x Fall“ signifikant, wenn nur die Versuchspersonen in die Analyse eingeschlossen werden, die sich weniger als dreimal tendenziell für schuldig entschieden hatten. Die signifikante Interaktion ist darauf zurückzuführen, dass bei den Fällen 4, 7 und 8 ein tendenzielles (aber nicht signifikantes) *spreading apart* zu verzeichnen ist, während bei den übrigen Fällen eher eine Abnahme der Differenz zwischen den korrigierten gewichteten Summen auftrat. Auf der Cue-Ebene wird der Haupteffekt „Messzeitpunkt“ in dieser Bedingung nicht mehr signifikant (p = .07), wobei aber tendenziell die Abwertung bestehen bleibt. Wenn nur Personen berücksichtigt wurden, die sich höchstens zweimal tendenziell bzw. final für schuldig entschieden haben, wurden die Cues „Augenzeuge“ und „Ausgaben“ wie zuvor signifikant abgewertet, während der Cue „Ort“ nur noch tendenziell abgewertet wurde. Insgesamt ist festzuhalten, dass sich das Ergebnismuster nur geringfügig verändert, wenn die Anzahl an tendenziellen oder finalen Entscheidungen für die Schuld des Verdächtigen in den Analysen berücksichtigt wird.

DISKUSSION

Die zentrale Fragestellung dieses Experimentes betraf den Einfluss der Wiederholung einer Entscheidungssituation auf die Veränderung subjektiver Cue-Validitäten. Die gemäß dem Prinzip der automatischen Konsistenzmaximierung formulierte Hypothese, dass sich systematische Umwertungen von Cues zeigen, konnte für die erste Bedingung, in der nur eine Entscheidungssituation dargeboten wurde, bestätigt werden, wenn auch nicht vollends überzeugend. In dieser Bedingung zeigten sich sowohl auf der Optionen-Ebene als auch – aber in eher schwachem Maße – auf der Cue-Ebene systematische Umwertungen. In der zweiten Bedingung trat auf der Optionen-Ebene kein *spreading apart* der Optionen auf – vielmehr zeigte sich eine tendenzielle Abnahme der Differenz zwischen den korrigierten gewichteten Summen der Optionen. Diese Tendenz zur Abwertung von Cues zeigte sich noch deutlicher auf der Cue-Ebene: Sieben von acht Cues wurden signifikant oder tendenziell abgewertet. Der Unterschied zwischen den Bedingungen und der besondere Einfluss der Wiederholung wurden auch bei einem Vergleich der Umwertungen zwischen den Bedingungen deutlich. Da der in Bedingung 1 dargebotene Fall auch in Bedingung 2 präsentiert wurde, war ein derartiger Vergleich möglich. Im Gegensatz zur ersten Bedingung war in der zweiten Bedingung kein *spreading apart* der Optionen zu beobachten. Auf der Cue-Ebene zeigte sich erwartungskonträr eine Abwertung von Cues. Ein Unterschied

zwischen beiden Bedingungen ist also deutlich zu erkennen. Folglich scheint die wiederholte Präsentation einer Entscheidungssituation einen Einfluss auf die Veränderung subjektiver Cue-Validitäten auszuüben. Dass die Umwertungen aber nicht systematisch sind, entspricht nicht der Erwartung. Da die entsprechende Hypothese nicht strikt aus dem PCS-Modell abzuleiten, aus Plausibilitätsgründen aber zu rechtfertigen war, stellen diese Befunde weniger eine Herausforderung für das zugrunde liegende Modell als vielmehr einen Anhaltspunkt für einen weiteren möglichen Einflussfaktor dar. Eine umfassende Diskussion dieses Aspektes erfolgt in Kapitel 3.7.3. Die Berücksichtigung der Anzahl an tendenziellen und finalen Entscheidungen für die Schuld des Verdächtigen beeinflusste das Ergebnismuster nicht maßgeblich. Dies ist insofern überraschend, als dass diese eigentlich für die Richtung der Umwertungen ausschlaggebend sind.

Betrachtet man lediglich die Ergebnisse für die erste Bedingung, so gilt es zu diskutieren, weshalb die Evidenz eher moderat ist. So traten in der ersten Bedingung zwar sowohl auf der Optionen- als auch auf der Cue-Ebene systematische Umwertungen auf, doch wurde mit dem Cue „Augenzeuge“ nur ein Cue signifikant umgewertet. Gemäß dem Konsistenzmaximierungsprozess hätten systematische und signifikante Umwertungen aller Cues erwartet werden können. Zu beachten ist aber, dass eine Umwertung aller Cues nicht zwingend nötig ist, um eine konsistente Interpretation zu erzielen, da auch lokale Optima stabile und damit konsistente Zustände repräsentieren (vgl. Kapitel 3.5.4). Dennoch stellt sich die Frage, warum zum Teil Umwertungen aller Cues und zum Teil Umwertungen weniger Cues beobachtet werden. Dies gilt es übergreifend in Kapitel 4 zu diskutieren. Welchen Einfluss die Wiederholung einer identisch konzipierten Entscheidungssituation hat, ist auf Basis der vorliegenden Daten nicht eindeutig zu klären. Zum einen ist auch die Evidenz in der zweiten Bedingung nicht überzeugend, da sich auf der Optionen-Ebene die Abnahme der Differenz zwischen den korrigierten gewichteten Summen der Optionen nur tendenziell zeigte und auf der Cue-Ebene lediglich drei Cues signifikant abgewertet wurden. Dennoch sprechen die Ergebnisse insgesamt dafür, dass eher eine Abwertung statt einer systematischen Umwertung stattfindet. Zum anderen kann auch diese Schlussfolgerung nicht eindeutig gezogen werden, da methodische Schwächen vorliegen. So ist zu überlegen, ob die hohe Anzahl an Fehlinterpretationen die Ergebnisse in nicht intendierter Weise beeinflusst hat. Es ist denkbar, dass die Versuchspersonen die Cues im Kontext des Entscheidungsproblems anders verstanden haben, was zu einer generellen Verunsicherung und einer damit einhergehenden Abwertung von Cues geführt haben kann. Angesichts der hypothesenkon-

formen Ergebnisse in Bedingung 1 würde diese Interpretation aber bedeuten, dass lediglich die Versuchspersonen der zweiten Bedingung dieser Verunsicherung unterlagen. Dies erscheint aus zwei Gründen als sehr unwahrscheinlich: Erstens erfolgte die Zuteilung der Versuchspersonen zu den einzelnen Bedingungen randomisiert. Zweitens finden sich hinsichtlich der im Pretest vorgenommen Cue-Einschätzungen deskriptiv keine Anzeichen für derartige Unterschiede. Im Pretest wurden, betrachtet man lediglich die vier Cues des gemeinsamen Falls, in der ersten Bedingung 101 und in der zweiten Bedingung 104 Angaben in intendierter Richtung gemacht. Im Posttest betrug dieses Verhältnis 104 (Bedingung 1) zu 102 (Bedingung 2). Dieses Ergebnis spricht gegen die Interpretation, dass in der zweiten Bedingung das Verständnis der Cues stärker beeinträchtigt war als in der ersten Bedingung. Allerdings bleibt die Möglichkeit bestehen, dass die Fehlinterpretationen die Ergebnisse in nicht erwünschter Weise beeinflusst haben. Auch wenn dieser Einfluss nicht unbedingt in einer Verunsicherung bestehen muss, könnte diese eine zusätzliche oder die entscheidende Rolle gespielt haben. So weisen Harvey und Fischer (2005) auf die Bedeutung des Zusammenspiels von Feedback und Erwartungen hin. Versuchspersonen haben initiale Erwartungen in einem Experiment, die im vorliegenden Experiment im Finden der richtigen Antwort bzw. im korrekten Urteilen bestanden haben können. Die Versuchspersonen erhielten aber nie Feedback über ihre Leistungen, so dass ihre Erwartungen nicht bestätigt oder korrigiert werden konnten. Dies kann zu einer Verunsicherung geführt haben, die sich in einem reduzierten Vertrauen in die Cues ausgedrückt hat. Eventuell erschwerend kam hinzu, dass zwar der Cue „Augenzeuge“ immer für die Schuld des Verdächtigen sprach, während die anderen Cues immer für die Unschuld sprachen, die Zusammensetzung der Cues pro Fall aber wechselte. Vor dem Hintergrund der nicht intendierten Einschätzung der Cues verschärft sich dieses Problem und könnte zur Verunsicherung beigetragen haben. Dass die Berücksichtigung der getroffenen Entscheidungen keinen Einfluss hatte, kann ebenfalls als Hinweis auf eine Verunsicherung, die sich in einem unsystematischen Entscheidungsverhalten geäußert hat, verstanden werden. Aufgrund dieser Unwägbarkeiten ist der tatsächliche Einfluss der Wiederholung schwer zu bestimmen. Daher wurde im folgenden Experiment der Frage nachgegangen, ob die Fehlinterpretationen entscheidend für die vorliegenden Ergebnisse waren oder ob das mangelnde Feedback zu einer Verunsicherung und damit zu einer Abwertung von Cues geführt hat. Auf diesem Wege sollten auch Rückschlüsse über den Einfluss der Wiederholung gezogen werden können.

3.7.2 Experiment 7: Der Einfluss entscheidender Informationen für die Veränderung von Cue-Validitäten nach wiederholten Entscheidungssituationen

ZIELSETZUNG

Im sechsten Experiment wurden die Cue-Validitäten bei einer Wiederholung der Entscheidungssituation hypothesenkonträr nicht systematisch umgewertet, sondern fast ausschließlich signifikant oder tendenziell abgewertet. Vor dem Hintergrund der bisherigen Ergebnisse und der Aussagen von Harvey und Fisher (2005) wurde die Hypothese aufgestellt, dass das fehlende Feedback die Versuchspersonen verunsichert und sie daher moderatere Validitätseinschätzungen abgeben. Um dieser Frage nachzugehen, wurde in dem vorliegenden Experiment manipuliert, ob die Versuchspersonen eine zusätzliche, entscheidende Information erhielten, die mit höchster Deutlichkeit auf die richtige Antwort verwies und damit Unsicherheit reduzieren sollte. Des Weiteren wurden die Skalen zur Einschätzung der subjektiven Validitäten optimiert und die Konfidenz in die jeweilige Entscheidung erfasst.

METHODE

Design. Es wurde *between subjects* variiert, ob die Versuchspersonen nach jedem Fall eine weitere, entscheidende Information erhielten oder nicht (Information vs. reine Wiederholung). *Within subjects* wurde die subjektive Validität von acht Cues vor und nach der Präsentation von zwölf Entscheidungssituationen gemessen. Die tendenzielle Entscheidung wurde vor und die finale Entscheidung nach dem Posttest angegeben. Nach jeder tendenziellen und finalen Entscheidung wurden Konfidenzurteile erhoben.

Hypothesen. Unter Berücksichtigung der Ergebnisse des vorherigen Experimentes und der Aussagen von Harvey und Fischer (2005) wurde erwartet, dass die Veränderung der Cue-Validitäten davon abhängt, ob die Versuchspersonen die entscheidende Information erhalten oder nicht. Wird keine weitere Information gegeben, so sollte die Konstruktion einer adäquaten (im Sinne des Konsistenzmaximierungsprozesses) mentalen Repräsentation aufgrund der Verunsicherung behindert sein. Dies sollte sich sowohl in einer generellen Abwertung von Cues als auch in niedrigen Konfidenzurteilen ausdrücken. Wird die entscheidende Information gegeben, sollte der Aufbau einer passenden mentalen Repräsentation möglich sein und die Validitätseinschätzungen sollten entsprechend dem Prozess der

automatischen Konsistenzmaximierung systematisch verändert werden: Cues, die für die offensichtlich richtige Antwort sprechen, sollten in ihrer Validität aufgewertet werden, während Cues, die gegen diese Antwort sprechen, abgewertet werden sollten.

Material. Das Material, das im vorliegenden Experiment verwendet wurde, basierte weitestgehend auf dem Material, das bereits im sechsten Experiment verwendet wurde. Im Gegensatz zum vorherigen Experiment wurde den Versuchspersonen die anfängliche Erklärung des Validitätskonzeptes separat für einen Fakt, der für die Schuld eines Verdächtigen spricht, und für einen Fakt, der für die Unschuld des Verdächtigen spricht, gegeben. Dies ist damit zu begründen, dass die Versuchspersonen im vorherigen Experiment zum Teil nicht wie intendiert zwischen dem Schuld-Fakt (*Schuld-Cue*) und den Unschuld-Fakten (*Unschuld-Cues*) unterscheiden konnten. In dem vorliegenden Experiment wurde daher expliziert, ob der jeweilige Fakt bzw. Cue für die Schuld oder für die Unschuld des Verdächtigen spricht. Die separaten Validitätsdefinitionen sollten das Verständnis weiter stützen. Die Validitätsdefinition für den Schuld-Cue lautete folglich: „Die Validität eines Faktes für die Schuld eines Verdächtigen ist definiert als die Wahrscheinlichkeit, dass bei Vorliegen des Faktes ein Verdächtiger eher schuldig ist als unschuldig.“ Analog lautete die Validitätsdefinition für einen Unschuld-Cue: „Die Validität eines Faktes für die Unschuld eines Verdächtigen ist definiert als die Wahrscheinlichkeit, dass bei Vorliegen des Faktes ein Verdächtiger eher unschuldig ist als schuldig.“ Die Versuchspersonen schätzten die Validität für den Schuld-Cue auf einer Skala, die von „keine Vorhersagekraft/50% schuldig“ bis „Immer schuldig/100% schuldig“ reichte, ein, und sie schätzten analog die Validitäten für die Unschuld-Cues auf einer Skala, die von „keine Vorhersagekraft/50% unschuldig“ bis „Immer unschuldig/100% unschuldig“ reichte, ein. Die Fakten bzw. Cues entsprachen denjenigen des vorherigen Experiments.

Als Treatment wurden den Versuchspersonen wieder die zwölf Fälle von Diebstahl in verschiedenen Unternehmen, die bereits im vorherigen Experiment eingesetzt wurden, vorgelegt. Zusätzlich wurde den Versuchspersonen der Informations-Bedingung noch eine weitere entscheidende Information gegeben, die die Schuldfrage klärte. In jeweils drei Fällen wurde die Information gegeben, dass ein überzeugendes Alibi vorliegt, dass ein Überwachungsvideo aufgetaucht ist und dass jemand ein Geständnis abgegeben hat. Die jeweilige Information war so formuliert, dass sie keinen Zweifel daran ließ, dass der Verdächtige unschuldig ist. So lautete etwa die zusätzliche Information im Fall des Blumengroßhandels „Blütenzauber“, in dem Maria F. des Diebstahls verdächtigt ist: „Maria F. hat ein Alibi.

Dieses Alibi wurde von Ihren Mitarbeitern überprüft und ist vollkommen wasserdicht. Demnach ist Maria F. unschuldig.“ Welche Information nach welcher Falldarstellung gegeben wurde, wurde randomisiert und zwischen den Versuchspersonen konstant gehalten. Das Material und die Instruktionen können Anhang A-7 und A-9 entnommen werden.

Stichprobe. An dem computerbasierten Experiment nahmen 128 Studenten (96 weiblich, 32 männlich; Durchschnittsalter: 25.68 Jahre, *SD* = 2.32) der Universität Erfurt teil. Sie wurden über das Online-Rekrutierungssystem *ORSEE* (Greiner, 2004) angeworben und bearbeiteten das Experiment im Rahmen einer Experimentalbatterie. Diese Experimentalbatterie dauerte insgesamt ca. 60 Minuten, wobei zwischen 20 und 35 Minuten auf das vorliegende Experiment entfielen. Die Versuchspersonen erhielten durchschnittlich 8 € für die Teilnahme an der Experimentalbatterie. Die Versuchspersonen wurden den beiden Bedingungen zufällig, aber mit dem Ziel einer Gleichverteilung zugeteilt. So waren 64 Versuchspersonen in der ersten („reine Wiederholung“) und 64 Personen in der zweiten Bedingung („Information“).

Durchführung. Nachdem das Validitätskonzept für Fakten, die für die Schuld des Verdächtigen sprechen, und für Fakten, die für die Unschuld des Verdächtigen sprechen, erläutert worden war, schätzten die Versuchspersonen die Validität des Cues „Augenzeuge“ (pro Schuld) sowie der Cues „Arbeitsweise“, „Ausgaben“, „Gesetz“, „Fingerabdruck“, „Anruf“, „Auto“ und „Ort“ (pro Unschuld) auf der entsprechenden Skala ein. Wie in den Untersuchungen zuvor reichte die jeweilige Skala von 50% bis 100%, da ein Cue mit einer Validität von 50% keine Vorhersagekraft hat. Nach der Erfassung demografischer Angaben und einem thematisch unverbundenen Experiment, das 10 bis 15 Minuten dauerte, erfolgte das Treatment: Die Versuchspersonen sollten sich vorstellen, dass sie als Detektiv bei einer Wirtschaftsdetektei arbeiteten. Sie selbst hätten sich auf Fälle von Diebstahl spezialisiert und bearbeiteten nun Aufträge von verschiedenen Unternehmen, in denen es zu Diebstählen gekommen ist. Nach vorherigen Recherchen sei in jedem Unternehmen ein Mitarbeiter verdächtigt. Es gäbe Hinweise für die Schuld, aber auch für die Unschuld des Verdächtigen. Vor der Präsentation der Fälle wurde den Versuchspersonen der ersten Bedingung daraufhin mitgeteilt, dass sie zwar noch die Ergebnisse von DNA-Analysen abwarten möchten, dass sie aber dennoch überlegen sollten, ob sie den Mitarbeiter aufgrund der vorliegenden Informationen für schuldig oder unschuldig halten. Diese Tendenz sowie ihre Sicherheit in diese Einschätzung gaben sie nach jeder Fallschilderung ab, worauf die nächste Fallbeschreibung folgte. Nachdem alle zwölf Fälle präsentiert worden waren und die Versuchsper-

sonen für jeden Fall ihre tendenzielle Entscheidung und ihre Konfidenz in diese angegeben hatten, bearbeiteten die Versuchspersonen den Posttest, der analog zum Pretest aufgebaut war. Den Versuchsteilnehmern wurde im Anschluss mitgeteilt, dass die DNA-Analysen im Labor verloren wurden und sie deshalb allein auf Basis der vorliegenden Informationen eine Entscheidung über die Schuld bzw. Unschuld des Verdächtigen treffen müssten. Ihnen wurden die Fälle noch einmal vorgelegt und die Versuchspersonen entschieden endgültig, ob sie den jeweiligen Verdächtigen für schuldig oder unschuldig halten, und gaben zusätzlich ihre Sicherheit in diese Entscheidung an. Den Versuchspersonen der zweiten Bedingung wurde vor der Präsentation der Fälle ebenfalls mitgeteilt, dass sie sich nach jeder Fallschilderung überlegen sollten, ob sie den jeweiligen Mitarbeiter für schuldig oder unschuldig halten, und ihre Sicherheit in diese Tendenz angeben sollten. Es wurde außerdem angekündigt, dass sie nach Angabe ihrer Entscheidungstendenz und des Ausmaßes an Konfidenz eine weitere Information erhielten, die ihnen Aufschluss über die Schuld bzw. Unschuld des Verdächtigen gäbe. Die zusätzliche Information, die den Versuchspersonen nach jedem Fall gegeben wurde, sprach für die Unschuld des Verdächtigen, wobei die Art der Information variiert wurde (Alibi, Überwachungsvideo oder Geständnis). Nach dem Erhalt dieser weiteren Information wurde jeweils der nächste Fall präsentiert. Nachdem die Versuchspersonen den letzten Fall bearbeitet hatten, beantworteten sie die Fragen des Posttests. Da sie bereits über die Unschuld des Verdächtigen informiert worden waren, gaben sie im letzten Teil lediglich an, wie sicher sie sich sind, dass der jeweilige Verdächtige tatsächlich unschuldig ist. Für diese Einschätzung wurden den Versuchspersonen erneut sämtliche Informationen zur Verfügung gestellt.

ERGEBNISSE

Vorausgehende Analysen. Aufgrund erheblicher Sprachprobleme (Nicht-Muttersprachler) mussten zwei Versuchspersonen von den weiteren Analysen ausgeschlossen werden.[59] Von den verbleibenden 126 Teilnehmern (95 weiblich, 31 männlich; Durchschnittsalter: 22.00 Jahre, $SD = 2.33$) waren 62 zufällig der ersten und 64 zufällig der zweiten Bedingung zugeteilt worden.

Interpretation der Informationen. Im vorherigen Experiment konnten die Cues post hoc danach klassifiziert werden, ob sie (a) tatnah sind und den Verdächtigen eindeutig be-

[59] Eine weitere Versuchsperson gab im Posttest stets Validitäten von 50% oder 100% an und wich im Pretest nur bei einem Cue von diesem Antwortmuster ab. Die Ergebnisse unter Ausschluss dieser Person werden gesondert berichtet.

bzw. entlasten, (b) tatnah sind und den Verdächtigen weniger eindeutig entlasten oder (c) tatfern sind und den Verdächtigen entlasten. Aufgrund der modifizierten Skala war die Richtung (Schuld vs. Unschuld), in der die Cues zu interpretieren waren, im vorliegenden Experiment eindeutig. Die (unbeeinflussten) Pretest-Validitäten lassen aber eine Überprüfung der angesprochenen Klassifikation zu. So sollten die Einschätzungen von (a) zu (c) zunehmend geringer ausfallen, da die jeweiligen Cues inhaltlich zunehmend weniger Vorhersagekraft für die Entscheidung zwischen Schuld und Unschuld des Verdächtigen haben. Tabelle 20 gibt einen Überblick über die mittleren Pretest-Validitätseinschätzungen und die jeweiligen Standardabweichungen, und zwar separat für beide Bedingungen. Außerdem sind die Cues den auf Basis der Ergebnisse aus Experiment 6 ermittelten Kategorien zugeordnet:

Tabelle 20: Mittlere Pretest-Validitätswerte (*M*) und Standardabweichungen (*SD*) pro Bedingung

			Reine Wiederholung		Information	
			M	*SD*	*M*	*SD*
Pretest	Augenzeuge	a	74.28	12.91	72.78	12.19
	Fingerabdruck	a	77.54	16.06	75.01	15.61
	Anruf	a	79.12	18.20	81.35	17.96
	Ort	a	80.71	18.11	81.33	16.79
	Auto	b	65.59	16.37	64.92	14.92
	Ausgaben	b	70.55	17.51	69.55	16.35
	Arbeitsweise	c	60.62	11.84	59.56	12.75
	Gesetz	c	65.14	13.61	64.84	14.87
n			62		64	

Anmerkungen: Die Cues wurden den in Experiment 6 ermittelten Kategorien zugeordnet. a: tatnahe Cues, die eindeutig be- oder entlasten; b: tatnahe Cues, die weniger eindeutig entlasten; c: tatferne Cues.

Wie den Tabellen zu entnehmen ist, stimmt die Zuordnung zu den Kategorien mit der Rangfolge der Cues im Pretest überein. Allerdings ist die Abgrenzung zwischen „Gesetz“ und „Auto“ in beiden Bedingungen eher gering und damit nicht überzeugend. Dennoch erscheint insgesamt die im vorherigen Experiment post hoc durchgeführte Klassifikation gerechtfertigt. Einen deskriptiven Überblick über die Validitätseinschätzungen im Pretest und im Posttest pro Bedingung gibt Tabelle B-10 in Anhang B.

Tendenzielle und finale Entscheidungen. Die Entscheidungssituationen wurden aus dem vorherigen Experiment übernommen, so dass die Informationslage nach wie vor für die Unschuld des Verdächtigen sprach. Folglich sollte die Mehrzahl der tendenziellen und finalen Entscheidungen „unschuldig" sein. In der Informations-Bedingung wurden keine finalen Entscheidungen erfasst, da die zusätzliche Information und die weiteren Instruktionen eindeutig auf die Unschuld des Verdächtigen verwiesen. Folglich konnten für diese Bedingung auch keine Switcher ermittelt werden. Die Tabellen 21 und 22 geben einen Überblick über die tendenziellen und finalen Entscheidungen in der jeweiligen Bedingung.

Tabelle 21: Tendenzielle und finale Entscheidungen in der Bedingung „reine Wiederholung"

		Fall											
		1	2	3	4	5	6	7	8	9	10	11	12
T	schuldig	10	6	2	3	10	9	1	11	6	7	23	11
	unschuldig	52	56	60	59	52	53	61	51	56	55	39	51
E	schuldig	6	8	3	5	14	9	3	14	5	10	24	6
	unschuldig	56	54	59	57	48	53	59	48	57	52	38	56
Switcher		8	8	3	4	8	12	2	5	5	7	7	7

Anmerkungen: T = tendenzielle Entscheidung. E = finale Entscheidung.

Tabelle 22: Tendenzielle Entscheidungen in der Bedingung „Information"

		Fall											
		1	2	3	4	5	6	7	8	9	10	11	12
T	schuldig	9	8	1	0	6	5	3	1	3	2	5	4
	unschuldig	55	56	63	64	58	59	61	63	61	62	59	60

Anmerkungen: T = tendenzielle Entscheidung. Finale Entscheidungen wurden in dieser Bedingung nicht getroffen.

Wie erwartet lautete die Mehrzahl der tendenziellen und finalen Entscheidungen „unschuldig". Da es für die erwartete Umwertung von Cue-Validitäten allerdings maßgeblich ist, dass die Versuchspersonen den Verdächtigen in allen oder zumindest den meisten Fällen für unschuldig halten, gilt es die Entscheidungen auf individueller Ebene in den Analysen zu berücksichtigen.

Die Veränderung subjektiver Cue-Validitäten. Zunächst wurde auf der Optionen-Ebene überprüft, ob ein systematisches *spreading apart* der Alternativen stattgefunden hat. Hierfür wurden wie in den vorherigen Analysen die korrigierten gewichteten Summen für die zwölf präsentierten Fälle berechnet, und zwar basierend auf den Pretest- und auf den Posttest-Werten. Anschließend wurden die Differenzen zwischen den gewichteten Summen pro Messzeitpunkt ermittelt, wobei für diese Berechnung die finale Entscheidung bzw. die Entscheidung „unschuldig" (Bed. „Information") zugrunde gelegt wurde. Eine pro Bedingung vorgenommene explorative Analyse dieser Differenzwerte zeigte keine extremen Ausreißer. Um zu überprüfen, ob der Erhalt einer entscheidenden Information einen Einfluss auf die Veränderung subjektiver Cue-Validitäten ausübt, wurde eine messwiederholte Varianzanalyse mit „Messzeitpunkt" und „Fall" als Innersubjektfaktoren und „Bedingung" als Zwischensubjektfaktor berechnet (siehe Tabelle 23). Von besonderem Interesse war die Interaktion „Messzeitpunkt x Bedingung". Eine post-hoc durchgeführte Teststärkenanalyse mittels GPower (Faul et al., 2007) ergab, dass ein Anlegen von $\alpha = .05$ bei $N = 126$ das Finden eines mittleren Effekts ($f = .25$, Cohen, 1988) mit einer Teststärke > .99 erlaubt.

Tabelle 23: Ergebnisse der messwiederholten Varianzanalyse

Faktor	*F*	*df*	*p*	η_p^2
Messzeitpunkt	18.15	1, 124	< .001	.13
Messzeitpunkt x Bedingung	17.53	1, 124	< .001	.12
Fall	17.43	5.36, 665.14[a]	< .001	.12
Fall x Bedingung	4.78	11, 1364	< .001	.04
Messzeitpunkt x Fall	2.97	3.89, 482.75[a]	< .01	.02
Messzeitpunkt x Fall x Bedingung	< 1	11, 1364	.99	.00
Bedingung	27.50	1, 124	< .001	.18

Anmerkungen: [a]Korrektur der Freiheitsgrade nach Greenhouse-Geisser.

Unter Hinzuziehung der deskriptiven Werte lassen sich die Ergebnisse wie folgt interpretieren: Über beide Bedingungen und Fälle hinweg zeigt sich ein Haupteffekt „Messzeitpunkt" derart, dass die Differenz zwischen den korrigierten gewichteten Summen der Optionen im Posttest höher ist als im Pretest. Es hat also ein *spreading apart* der Alternativen stattgefunden. Allerdings gibt der Haupteffekt „Bedingung" einen Hinweis auf

einen Unterschied zwischen den beiden Bedingungen, der durch die Interaktion „Messzeitpunkt x Bedingung“ und eine Inspektion der Mittelwerte expliziert wird: Das *spreading apart* ist nur in der Informations-Bedingung zu beobachten, während sich in der reinen Wiederholungs-Bedingung keine wesentlichen Veränderungen zeigen. Folglich wird auch die Interaktion „Fall x Bedingung“ signifikant, da aufgrund des *spreading aparts* in der Informations-Bedingung höhere mittlere Differenzwerte vorliegen als in der reinen Wiederholungsbedingung. Der Haupteffekt „Fall“ ist darauf zurückzuführen, dass pro Fall unterschiedliche Cues und damit unterschiedliche Validitäten vorliegen, die in verschiedenen Differenzwerten resultieren. Zudem ist, wie die signifikante Interaktion „Messzeitpunkt x Fall“ zeigt, das *spreading apart* für die einzelnen Fälle unterschiedlich. Pro Bedingung durchgeführte ANOVAs bestätigen, dass das *spreading apart* lediglich in der Informations-Bedingung signifikant ist, $F(1, 63) = 25.69$, $p < .001$, $\eta_p^2 = .29$ und $F(1, 61) < 1$, $p = .95$, $\eta_p^2 = .00$ (siehe Abbildung 11).[60]

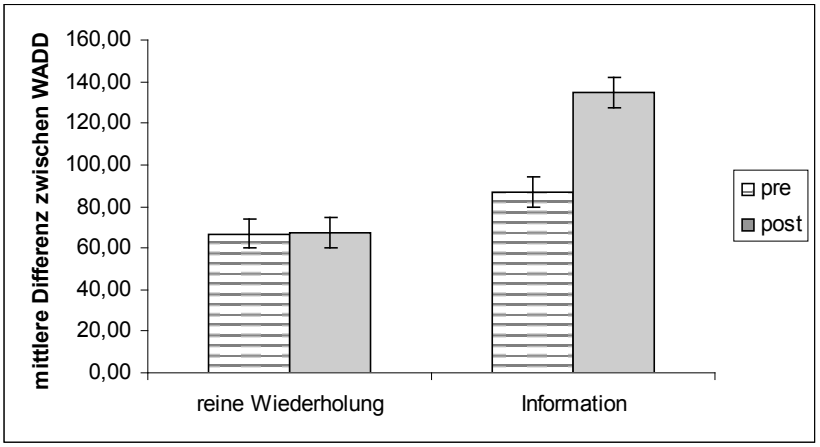


Abbildung 11: Mittlere Veränderung der Differenz zwischen den korrigierten gewichteten Summen der Optionen (WADD) vom Pre- zum Posttest pro Bedingung. Fehlerbalken zeigen Standardfehler.

Diese ANOVAs geben auch Aufschluss über die angesprochene Interaktion „Messzeitpunkt x Fall“. Auch diese ist lediglich in der Informations-Bedingung signifikant, $F(11, 693) = 1.95$, $p < .05$, $\eta_p^2 = .03$. Eine Inspektion der Mittelwerte zeigt, dass für die Fälle 4, 7 und 8 ein vergleichsweise hohes, für die Fälle 1, 2 und 9 ein vergleichsweise niedriges *spreading apart* verzeichnet werden konnte. In diesen Fällen wurden jeweils die gleichen Cues (Auto, Augenzeuge, Anruf und Ort bzw. Augenzeuge, Ort, Arbeitsweise und Gesetz)

[60] Ein mittlerer Effekt ($f = .25$, Cohen, 1988) kann in der reinen Wiederholungs-Bedingung bei $\alpha = .05$ und $n = 62$ mit einer Teststärke > .99 gefunden werden. In der Informationsbedingung kann er bei $\alpha = .05$ und $n = 64$ mit einer Teststärke von .86 ermittelt werden. Analysen erfolgten mit GPower (Faul et al., 2007).

verwendet. Dies deutet bereits auf Unterschiede in der Veränderung subjektiver Cue-Validitäten hin, denen im Folgenden nachgegangen wird.

Bevor auf der Ebene der Cues analysiert wurde, ob sich die subjektiven Cue-Validitäten verändert haben, wurde anhand von Boxplots analysiert, ob extreme Ausreißer vorlagen. Es zeigten sich in keiner Bedingung wiederholte extreme Ausreißer. Um die Veränderung auf Cue-Ebene näher zu untersuchen, wurden zunächst messwiederholte ANOVAs mit „Messzeitpunkt“ und „Cue“ als Innersubjektfaktoren pro Bedingung berechnet. Erneut wurden post-hoc Teststärkenanalysen mittels GPower (Faul et al., 2007) durchgeführt, die zunächst für die reine Wiederholungs-Bedingung dargestellt werden sollen: Das Anlegen eines Signifikanzniveaus von .05 erlaubt bei $n = 62$ das Finden eines mittleren Effekts ($f = .25$, Cohen, 1988) mit einer Teststärke > .99 (Haupteffekt „Messzeitpunkt“ und Interaktion „Messzeitpunkt x Cue“). Für die Informationsbedingung ergaben sich folgende Ergebnisse: Hier gestattet ein α-Niveau von .05 bei $n = 64$ das Finden eines mittleren Effekts ($f = .25$, Cohen, 1988) mit Teststärken von .93 (Haupteffekt „Messzeitpunkt“) bzw. > .99 (Interaktion „Messzeitpunkt x Cue“). Während der Haupteffekt „Messzeitpunkt“ nur in der reinen Wiederholungs-Bedingung, nicht aber in der Informations-Bedingung signifikant wurde, $F(1, 61) = 23.71$, $p < .001$, $\eta_p^2 = .28$ bzw. $F(1, 63) < 1$, $p = .60$, $\eta_p^2 = .00$, wurde die Interaktion „Messzeitpunkt x Cue“ lediglich in der Informations-Bedingung, nicht aber in der reinen Wiederholungs-Bedingung signifikant, $F(7, 441) = 9.03$, $p < .001$, $\eta_p^2 = .13$ bzw. $F(5.92, 361.00) = 1.86$, $p = .09$, $\eta_p^2 = .03$[61]. Gepaarte t-Tests zeigen, wie die Cues in den einzelnen Bedingungen umgewertet werden (siehe Tabelle 24). Während in der reinen Wiederholungs-Bedingung alle Cues signifikant oder zumindest tendenziell abgewertet werden, ist die Umwertung in der Informations-Bedingung systematischer: Der Cue „Augenzeuge“, der als einziger Cue für die Schuld des Verdächtigen spricht, wird wie erwartet abgewertet, wohingegen die Cues „Auto“ und „Ort“ aufgewertet werden. Tendenziell werden auch die Cues „Anruf“ und „Arbeitsweise“ aufgewertet. Die unterschiedlich starke Umwertung der Cues in dieser Bedingung erklärt auch die zuvor auf Optionen-Ebene gefundene Interaktion „Messzeitpunkt x Fall“. Die Fälle 1, 2 und 9 unterscheiden sich in zwei Cues von den Fällen 4, 7 und 8. In den erstgenannten Fällen kommen neben den Cues „Augenzeuge“ und „Ort“ die Cues „Arbeitsweise“ und „Gesetz“ vor, in den letztgenannten Fällen kommen die Cues „Auto“ und „Anruf“ vor. Die Cues „Auto“ und „Anruf“ werden signifikant bzw. tendenziell aufgewertet, was zu einer

[61] Korrektur der Freiheitsgrade nach Huynh-Feldt, da ε > .75.

größeren Differenz zwischen den korrigierten gewichteten Summen im Posttest führt. Der Cue „Arbeitsweise“ wird tendenziell aufgewertet, aber der Cue „Gesetz“ wird kaum verändert, so dass sich die Differenz zwischen den korrigierten gewichteten Summen nicht so stark vergrößert. Anzumerken ist, dass gerade die Cues signifikant umgewertet werden, die häufiger präsentiert wurden. So wurden die Cues „Augenzeuge“ und „Ort“ zwölfmal und die Cues „Auto“ und „Arbeitsweise“ sechsmal präsentiert.

Tabelle 24: Ergebnisse der gepaarten *t*-Tests pro Cue und Bedingung

	Reine Wiederholung			Information		
	t-Wert	*p*	d_z	*t*-Wert	*p*	d_z
Augenzeuge	-4.88	< .001	-.62	-6.71	< .001	-.84
Arbeitsweise	-3.05	< .01	-.39	1.59	.06	.20
Ausgaben	-1.70	< .05	-.22	-0.00	.50	-.00
Gesetz	-1.93	< .05	-.25	-0.77	.22	-.10
Fingerabdruck	-3.18	< .01	-.40	0.02	.49	.00
Anruf	-0.48	.32	-.06	1.60	.06	.20
Auto	-1.08	.14	-.14	2.39	< .05	.30
Ort	-0.23	.41	-.03	2.90	< .01	.36
n	62			64		

Anmerkungen: Die gepaarten *t*-Tests waren einseitig.

Es konnte sowohl auf Optionen- als auch auf Cue-Ebene gezeigt werden, dass bei einer Wiederholung der Entscheidungssituation die Umwertung von Cues vom Erhalt einer entscheidenden Information abhängt: Wird diese nicht erhalten, so werden alle Cues signifikant oder tendenziell abgewertet. Wird diese Information erhalten, findet eine systematischere Umwertung statt. Einschränkend ist darauf zu verweisen, dass in der Informations-Bedingung lediglich drei Cues signifikant umgewertet wurden.[62]

Konfidenz. Um zu überprüfen, ob sich die Konfidenz in die tendenziellen und finalen Entscheidungen zwischen den Bedingungen unterscheidet, wurden für beide Konfidenzurteile messwiederholte ANOVAs mit „Fall“ als Innersubjektfaktor und „Bedingung“ als Zwischensubjektfaktor berechnet. War die Konfidenz in die tendenziellen Entscheidungen die abhängige Variable, zeigten sich signifikante Ergebnisse für den Haupteffekt „Fall“,

[62] Die Ergebnisse veränderten sich weder auf Optionen- noch auf Cue-Ebene, wenn die Versuchsperson, die im Posttest stets Validitätseinschätzungen von 50% oder 100% gab, ausgeschlossen wurde.

$F(9.37, 1162.12) = 14.91$, $p < .001$, $\eta_p^2 = .11$[63], für den Haupteffekt „Bedingung“, $F(1, 124) = 9.25$, $p < .01$, $\eta_p^2 = .07$, und für die Interaktion „Fall x Bedingung“, $F(11, 1364) = 4.94$, $p < .001$, $\eta_p^2 = .04$. Der Haupteffekt „Fall“ ist damit zu begründen, dass die Konfidenzurteile je nach Fall unterschiedlich sind. Die Interaktion beruht darauf, dass die zwischen den Bedingungen auftretenden Unterschiede in den Konfidenzurteilen je nach Fall divergieren. Für die vorliegende Fragestellung ist insbesondere der Haupteffekt „Bedingung“, der auf die in der Informations-Bedingung höheren Konfidenzurteile zurückzuführen ist, interessant. Anzumerken ist hierbei, dass sich dieser Unterschied – wie deskriptiv und in Abbildung 12 zu erkennen ist – erst ab dem fünften Fall sichtlich zeigt. Ab diesem Fall sind die Konfidenzurteile in der Informations-Bedingung stets höher als in der reinen Wiederholungs-Bedingung.

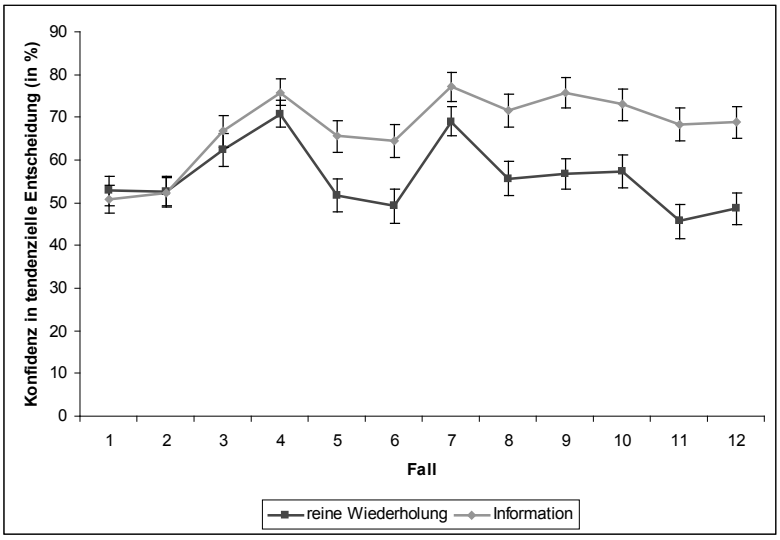


Abbildung 12: Konfidenz in die tendenzielle Entscheidung pro Fall und Bedingung. Fehlerbalken geben Standardfehler an.

Besonders deutlich zeigt sich der Bedingungsunterschied bei der Analyse der Konfidenzurteile in die finalen Entscheidungen. Erneut zeigen sich signifikante Ergebnisse für den Haupteffekt „Fall“, $F(9.20, 1140.34) = 9.03$, $p < .001$, $\eta_p^2 = .07$[64], für den Haupteffekt „Bedingung“, $F(1, 124) = 72.40$, $p < .001$, $\eta_p^2 = .37$, und für die Interaktion „Fall x Bedingung“, $F(11, 1364) = 3.40$, $p < .001$, $\eta_p^2 = .03$. Wie Abbildung 13 veranschaulicht,

[63] Die Korrektur der Freiheitsgrade erfolgte nach Huynh-Feldt.
[64] Die Korrektur der Freiheitsgrade erfolgte nach Huynh-Feldt.

sind nun die Konfidenzurteile in der Informations-Bedingung von der ersten Entscheidung an deutlich höher als in der reinen Wiederholungsbedingung.

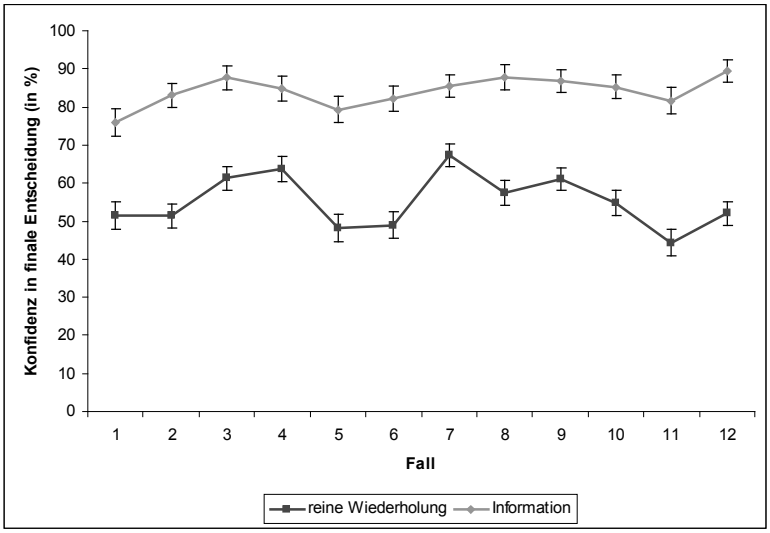


Abbildung 13: Konfidenz in die finale Entscheidung pro Fall und pro Bedingung. Fehlerbalken geben Standardfehler an.

Somit konnte die Hypothese, dass die reine Wiederholung einer Entscheidungssituation zu niedrigen Konfidenzurteilen führt, während der Erhalt einer zusätzlichen entscheidenden Information mit hohen Konfidenzurteilen einhergeht, bestätigt werden.

Berücksichtigung der „schuldig"-Entscheidungen. Da die Hypothesen auf der Annahme beruhen, dass die Versuchspersonen den jeweiligen Verdächtigen in der Mehrzahl der Fälle für unschuldig halten, galt es wie im vorherigen Experiment die Anzahl an tendenziellen bzw. finalen Entscheidungen für die Schuld des Verdächtigen in einer zusätzlichen Analyse zu berücksichtigen. Erneut wurde der *Cutoff* bei dem Wert 2 gesetzt, d.h. dass all die Versuchspersonen ausgeschlossen wurden, die sich häufiger als zweimal tendenziell bzw. final für die Schuld des Verdächtigen entschieden hatten. In den Tabellen C-6 und C-7 des Anhangs C sind die Ergebnisse separat für beide Bedingungen tabellarisch gegenübergestellt, wobei erneut ein Fokus auf die interessierenden Effekte und die signifikanten *t*-Tests gelegt wurde. An dieser Stelle sollen nur die wenigen Resultate erwähnt werden, die von den ursprünglich erhaltenen Ergebnissen abweichen. Werden die Personen ausgeschlossen, die sich tendenziell mehr als zweimal für die Schuld des Verdächtigen entschieden hatten, so zeigen sich lediglich auf der Cue-Ebene leichte Abweichungen: In der reinen Wiederholungs-Bedingung werden die Cues „Arbeitsweise" und „Gesetz" nicht mehr signifikant abgewertet. Darüber hinaus wird in der Informations-Bedingung nun auch der Cue „Anruf"

signifikant aufgewertet. Werden die Personen ausgeschlossen, die den Verdächtigen final mehr als zweimal für schuldig gehalten haben, so gibt es nur auf der Cue-Ebene eine kleine Abweichung in der reinen Wiederholungs-Bedingung: Der Cue „Gesetz“ wird nicht mehr signifikant abgewertet. Insgesamt betrachtet bleiben die Ergebnisse durch die Berücksichtigung der tendenziellen bzw. finalen „schuldig“-Entscheidungen nahezu unverändert.

Diskussion

Im vorherigen Experiment 6 führte die reine Wiederholung einer Entscheidungssituation wider Erwarten zu einer überwiegenden Abwertung von Cues. Zwei Erklärungen, die sich nicht ausschließen, wurden hierfür vorgeschlagen: Erstens könnten die Ergebnisse aus einer methodischen Schwäche resultieren. So war die für die Einschätzung der subjektiven Cue-Validitäten vorgesehene Skala so konzipiert, dass die Versuchspersonen nicht nur die Höhe der Validität, sondern auch die Richtung der Vorhersagekraft (Schuld vs. Unschuld) einschätzen konnten. Dies führte zu nicht intendierten Cue-Interpretationen und ließ die Möglichkeit zu, dass die Versuchspersonen die Cues im Kontext der Entscheidungssituation anders verstanden haben. Zweitens könnten die Resultate eine inhaltliche Ursache haben: Laut Harvey und Fischer (2005) ist das Zusammenspiel zwischen Feedback und Erwartungen zu berücksichtigen. Bei wiederholten probabilistischen Inferenzaufgaben könnten die Versuchspersonen Rückmeldungen erwartet haben, die jedoch ausblieben. Folglich konnten sie die Richtigkeit ihrer Antworten nicht einschätzen, was zu einer generellen Verunsicherung geführt haben könnte, die sich in einer Abwertung von Cues ausgedrückt hat. Es könnte zu der Verunsicherung beigetragen haben, dass zwar die Richtung, in der die Cues eine Vorhersage machten (Schuld vs. Unschuld), konstant blieb, aber die Zusammensetzung der Cues pro Fall variierte.

Im vorliegenden Experiment wurde der Frage nachgegangen, ob die Fehlinterpretationen und/oder die Verunsicherung als Ursache für die Abwertung von Cues nach wiederholten Entscheidungssituationen anzusehen ist. Hierfür wurden die Skalen zur Einschätzung der Cue-Validitäten modifiziert und es wurde *between subjects* manipuliert, ob den Versuchspersonen eine weitere Information gegeben wird, die ihnen Gewissheit über die Unschuld des Verdächtigen gibt. Darüber hinaus wurde nach jeder tendenziellen und finalen Entscheidung die Konfidenz erfasst. In den Experimenten 2 und 6 wurde die Hypothese aufgestellt, dass auch bei einer Wiederholung der Entscheidungssituation eine systematische Veränderung von Cue-Validitäten eintritt. Obgleich der im PCS-Modell zentrale Konsis-

tenzmaximierungsprozess für einmalige Entscheidungssituationen formuliert wurde, erschien es plausibel anzunehmen, dass auch nach der Wiederholung einer identisch konstruierten Entscheidungssituation eine systematische Reevaluation zu beobachten ist. Im vorliegenden Experiment wurde die diesbezügliche Hypothese vor dem Hintergrund der bisherigen Ergebnisse modifiziert. So wurde erwartet, dass die reine Wiederholung einer Entscheidungssituation die Konstruktion einer adäquaten (im Sinne des Konsistenzmaximierungsprozesses) mentalen Repräsentation beeinträchtigt, da die Versuchspersonen verunsichert sind. Das Resultat sollte eine generelle Abwertung von Cues sein. Der Erhalt einer zusätzlichen und entscheidenden Information sollte hingegen die Verunsicherung reduzieren und den Aufbau einer „passenden" mentalen Repräsentation begünstigen. Dies sollte zu einer systematischen Umwertung von Cues führen.

Insgesamt können die Hypothesen als weitgehend bestätigt angesehen werden. Lediglich in der Informations-Bedingung zeigte sich ein *spreading apart* der Alternativen, während in der reinen Wiederholungs-Bedingung keine signifikante Änderung auftrat. Folglich war aber auch auf Optionen-Ebene keine Abwertung, wie sie im vorherigen Experiment gefunden worden war, zu erkennen. Diese zeigte sich aber auf der Cue-Ebene: Wie in Experiment 6 überwog die Abwertung von Cues in der reinen Wiederholungs-Bedingung. Im Gegensatz dazu war die Umwertung in der Informations-Bedingung systematischer, da der Cue „Augenzeuge" wie erwartet abgewertet wurde und zwei Unschuld-Cues aufgewertet wurden. Bei zwei weiteren Cues war eine deutliche Tendenz zur Aufwertung zu erkennen. Einschränkend ist anzumerken, dass wie in vorherigen Experimenten nicht alle Cues systematisch umgewertet wurden und insofern keine Konsistenz*maximierung* erfolgt ist. Die Konfidenzurteile entsprachen den Erwartungen: So war die Konfidenz in die tendenziellen Entscheidungen wie auch in die finalen Entscheidungen in der Informations-Bedingung signifikant höher als in der reinen Wiederholungs-Bedingung. Somit scheinen die Versuchspersonen in der letztgenannten Bedingung tatsächlich verunsichert gewesen zu sein, während dies bei den Versuchspersonen in der erstgenannten Bedingung nicht der Fall war. Folglich wird die bei der Diskussion des sechsten Experimentes geäußerte Vermutung, dass die Wiederholung einer Entscheidungssituation einen weiteren Faktor darstellt, der die Veränderung von subjektiven Cue-Validitäten beeinflusst, gestützt. Der Wirkmechanismus könnte folgendermaßen beschrieben werden: Die reine Wiederholung führt nach den bisherigen Ergebnissen zu einer Verunsicherung, die mit einem Misslingen, eine adäquate konsistente (im Sinne des Konsistenzmaximierungsprozes-

ses) Repräsentation des Entscheidungsproblems aufzubauen, einhergeht. Wird das Entscheidungsproblem durch die Gabe einer zusätzlichen Information erleichtert, so steigt die Konfidenz, was wiederum mit dem Aufbau einer adäquaten konsistenten Repräsentation verbunden ist. Es mag der Einwand erhoben werden, dass eine solche zusätzliche Information eine Art Feedback darstellt, welches ein Ende des Entscheidungsprozesses impliziert. Dies wiederum würde die Vermutung nahelegen, dass in der entsprechenden Bedingung keine prädezisionalen, sondern postdezisionale Informationsumwertungen aufgetreten sind. Die Versuchspersonen wurden zwar darauf hingewiesen, selbst noch keine endgültige Entscheidung zu treffen, aber dennoch erscheint die Kritik insofern berechtigt, als dass natürlich nicht ausgeschlossen werden kann, dass die Versuchspersonen sich trotzdem schon entschieden hatten. Hierbei gilt aber zu beachten, dass diese Entscheidung noch nicht bindend war, was laut Festinger (1964) eine Voraussetzung für das Auftreten von Dissonanz und anschließenden Dissonanzreduktionsprozessen ist. Dass die Anzahl an tendenziellen bzw. finalen Entscheidungen für die Schuld des Verdächtigen wie in Experiment 6 keinen wesentlichen Einfluss auf die Ergebnisse hatte, ist zunächst überraschend, da ja gerade die aufkommende Entscheidung den Prozess der Konsistenzmaximierung „leitet“. Als Erklärung kommen mindestens zwei Ursachen in Betracht. Erstens ist es denkbar, dass erst eine höhere Anzahl an „schuldig“-Entscheidungen einen Einfluss auf die Ergebnisse haben könnte. Eine solch höhere Anzahl, z.B. sechs und mehr Entscheidungen für die Schuld des Verdächtigen, kam aber nur vereinzelt vor. Zweitens ist es durchaus möglich, dass die Versuchspersonen zwar dazu tendierten, sich für die Unschuld des Verdächtigen zu entscheiden, dass sie aber das wiederholte Treffen dieses Urteils misstrauisch machte und sie sich daher im Sinne eines *demand*-Effektes nur vordergründig für die Schuld des Verdächtigen entschieden. Die Daten lassen aber nicht den Schluss zu, welche der Erklärungen richtig ist. Da bei den entsprechenden Analysen hauptsächlich Personen der reinen Wiederholungs-Bedingung ausgeschlossen werden mussten, ist auch eine generelle Verunsicherung nicht auszuschließen (vgl. Experiment 6).

Weitere interessante Ergebnisse, die sich nicht auf die konkrete Fragestellung beziehen, betreffen die Klassifikation der Cues und die Präsentationshäufigkeit im Experiment. Die im vorherigen Experiment post hoc durchgeführte Klassifikation in tatnahe/eindeutig be- bzw. entlastende, tatnahe/weniger eindeutig entlastende und tatferne/entlastende Cues konnte im vorliegenden Experiment bestätigt werden. Letztlich verdeutlicht dies, dass eine entscheidungsspezifische Differenzierung von Cues stattfindet, die objektiv plausibel ist.

Interessanter ist die Frage, ob ein Zusammenhang zwischen der Kategorienzugehörigkeit und der Stärke der Umwertung besteht. Diese Frage sollte und kann mittels der vorliegenden Daten nicht beantwortet werden, stellt aber eine Anregung für folgende Arbeiten dar. Ein weiterer interessanter Aspekt betrifft die Präsentationshäufigkeit. So weisen die Ergebnisse darauf hin, dass häufiger präsentierte Cues eher umgewertet werden (in der Informations-Bedingung). So wurden die beiden umgewerteten Cues „Augenzeuge“ und „Ort“ zwölfmal präsentiert und der aufgewertete Cue „Auto“ wurde sechsmal dargeboten. Der Cue „Arbeitsweise“, der ebenfalls in sechs Fällen vorkam, wurde zumindest tendenziell aufgewertet. Ob es tatsächlich einen Zusammenhang zwischen der Präsentationshäufigkeit und der Stärke der Umwertung gibt, stellt ebenfalls eine spannende Frage für künftige Forschungsvorhaben dar.

3.7.3 Zusammenfassung und Diskussion

Dass sich eine wiederholte Konfrontation mit einer Entscheidungssituation auf die Stärke und Stabilität der prädezisionalen Informationsverzerrung auswirken kann, wurde zwar diskutiert (Glöckner, 2006; Simon, 2004), aber es mangelt an empirischer Evidenz. Folglich wurde der Einfluss der Wiederholung einer Entscheidungssituation auf die Veränderung subjektiver Cue-Validitäten in zwei Experimenten untersucht, die im Folgenden zusammenfassend diskutiert werden.

In Anlehnung an den *Jason Wells Case* von Simon, Snow et al. (2004) wurden neue, vergleichsweise komplexe probabilistische Inferenzaufgaben konstruiert. Ein Mitarbeiter eines Unternehmens wurde des Diebstahls verdächtigt, wobei ein Fakt für seine Schuld und drei andere Fakten für seine Unschuld sprachen. In Experiment 6 wurde die einmalige mit der wiederholten Darbietung einer Entscheidungssituation verglichen. Wurden die Versuchspersonen mit einer Entscheidungsaufgabe konfrontiert, waren systematische Umwertungen sowohl auf der Optionen- als auch auf der Cue-Ebene zu beobachten. Letztere waren jedoch lediglich moderat, da nur ein Cue systematisch umgewertet wurde. Bei der wiederholten Darbietung einer Entscheidungssituation wurde das überraschende Ergebnis gefunden, dass keine systematische Umwertung auftritt, sondern dass wenn Cues reevaluiert werden, diese abgewertet werden. Folglich zeigte sich auch auf der Optionen-Ebene kein *spreading apart* der Optionen, sondern eine zumindest tendenzielle Verringerung der Differenzen. Als mögliche Erklärung wurde eine Verunsicherung diskutiert. So könnten das mangelnde Feedback, in Verbindung mit der durch den Wechsel der Cue-Konstellationen

erhöhten Schwierigkeit der Aufgabe, zu selbiger und damit zu vergleichsweise moderaten Validitätseinschätzungen geführt haben. Eine Schwäche dieses Experiments bestand in den Skalen zur Validitätseinschätzung: Da die Versuchspersonen nicht nur die Höhe der Vorhersagekraft, sondern auch die Richtung der Vorhersage angeben konnten, wurden einige Cues nicht wie intendiert interpretiert. Wie auch die vorherigen Experimente gezeigt haben, ist aber eine unausgeglichene Informationslage Voraussetzung für das Auftreten von Reevaluationen, und zudem sollten die Cues in jedem Fall als Schuld-Cue oder als Unschuld-Cues gedeutet werden, um Nulleffekte zu vermeiden. Folglich konnten die Ergebnisse nicht eindeutig interpretiert werden. Im anschließenden Experiment 7 wurde daher die Hypothese bezüglich der Verunsicherung unter Verwendung modifizierter und für die Fragestellung optimierter Skalen untersucht. Während in der einen Bedingung erneut eine reine Wiederholung der Entscheidungssituation stattfand, wurde in der anderen Bedingung nach jeder Aufgabe eine zusätzliche Information gegeben, die die Schuldfrage klärte und damit Verunsicherung reduzieren sollte. Bei der reinen Wiederholung war in Übereinstimmung mit den Ergebnissen des vorherigen Experimentes keine systematische Umwertung von Cues, sondern eher eine Abwertung derselben zu beobachten. Zwar fand sich auf Optionen-Ebene keine signifikante Abnahme der Differenz zwischen den korrigierten gewichteten Summen der Optionen, doch zeigte sich auf Cue-Ebene, dass fünf von acht Cues signifikant abgewertet wurden. Folglich ist die auch im vorherigen Experiment gefundene Abwertung von Cues nicht auf die Schwäche der Skala, sondern auf die reine Wiederholung der Entscheidungssituation zurückzuführen. Wurde eine zusätzliche Information gegeben, fand eher eine systematische Umwertung statt. So ließ sich auf Optionen-Ebene ein signifikantes *spreading apart* der Alternativen beobachten. Auf Cue-Ebene wurden drei Cues signifikant und zwei weitere tendenziell systematisch umgewertet, so dass die Evidenz auch hier eher moderat ist. Ein Vergleich der Konfidenzurteile machte deutlich, dass diese in der Informations-Bedingung deutlich höher sind als in der reinen Wiederholungs-Bedingung, so dass in ersterer die Verunsicherung vergleichsweise reduziert war.

Vor dem Hintergrund dieser Ergebnisse und dem PCS-Modell gilt es zu diskutieren, welchen Einfluss die Wiederholung einer Entscheidungssituation hat und welche Rolle der Verunsicherung hierbei zukommt. Auch ist zu erörtern, weshalb die Evidenz eher moderat ist. In diesem Zusammenhang ist des Weiteren zu fragen, wovon es abhängen könnte, welche Cues umgewertet werden. Die Wiederholung einer Entscheidungsaufgabe hat sich

als weiterer interessanter Einflussfaktor bezüglich der Veränderung subjektiver Cue-Validitäten erwiesen. Dass eine reine Wiederholung eher zu einer Abwertung von Cues führt, die zusätzliche Verfügbarkeit einer wichtigen Information aber eine eher systematische Umwertung bewirkt, lässt folgenden möglichen Mechanismus plausibel erscheinen: Die reine Wiederholung einer probabilistischen Inferenzaufgabe verunsichert die Versuchspersonen, da sie kein Feedback erhalten (vgl. Harvey und Fischer, 2005) und die Entscheidungssituation zusätzlich dadurch erschwert ist, dass zwar das Informationsmuster konstant, aber die Cue-Konstellation wechselnd ist. Diese Verunsicherung drückt sich offensichtlich in niedrigen Konfidenzurteilen und zudem in erniedrigten Validitätseinschätzungen aus. Wird hingegen die Verunsicherung reduziert, indem eine entscheidende Information gegeben wird, steigen die Konfidenzurteile und die Umwertung ist systematisch. Obgleich in Experiment 2, in dem der Einfluss der Wiederholung erfolglos untersucht wurde, und 6 zunächst vermutet wurde, dass auch nach der wiederholten Darbietung eine systematische Umwertung von Cue-Validitäten eintritt, sprechen die Befunde nicht gegen das PCS-Modell. So wurde der Prozess der Konsistenzmaximierung lediglich für eine einmalige Entscheidungsaufgabe formuliert. Zwar erschien es plausibel anzunehmen, dass die Systematik in der Umwertung nicht beeinträchtigt wird, doch war diese Hypothese nicht strikt abzuleiten. Simulationsstudien könnten Aufschluss darüber geben, welche diesbezüglichen Vorhersagen das PCS-Modell macht. Es soll daher nicht von Evidenz *gegen* das PCS-Modell gesprochen werden, sondern vielmehr von Evidenz *für* einen interessanten Wirkmechanismus, der bei der Wiederholung einer Entscheidungsaufgabe eintritt. Vor dem Hintergrund des PCS-Modells und empirischer Befunde bezüglich komplexer Entscheidungsaufgaben (Simon, Snow et al., 2004) ist es aber überraschend, dass die Evidenz in allen Bedingungen eher moderat ist. Nachdem in Experiment 6 in der Wiederholungs-Bedingung nur drei von acht Cues signifikant abgewertet worden waren, lieferte Experiment 7 überzeugendere Evidenz dafür, dass eine reine Wiederholung eher zu einer Abwertung als zu gar keiner Umwertung führt. Dass bei der einmaligen Konfrontation mit einer Entscheidungsaufgabe nur ein Cue signifikant umgewertet wurde, war nach dem PCS-Modell nicht zu erwarten gewesen, zumal die Komplexität erhöht wurde, was kohärenzbasiertes Schlussfolgern begünstigt (Simon, 2004). Allerdings entspricht die Umwertung weniger Cues den Befunden der vorherigen Experimente und kann erneut mit der Möglichkeit, dass ein lokales Optimum gefunden wurde, begründet werden. Zudem gilt es Spezifika der Entscheidungsaufgabe zu beachten, auf die auch Simon, Krawczyk et al. (2004) hinweisen: Wenn eine Entscheidungsaufgabe

durch abstrakte und ambigue Informationen gekennzeichnet ist, begünstigt dies *coherence shifts*. Im vorliegenden Experiment ließen die Informationen zwar mehr Raum für Interpretationen als in den vorherigen Experimenten, doch im Vergleich zu den rechtlichen Fällen von Simon und Kollegen waren die Informationen wenig widersprüchlich. Damit werden erneut die Aspekte der Ambiguität und der Konfidenz aufgegriffen, die im abschließenden Kapitel 4 zu erörtern sind. Zuvor soll aber diskutiert werden, wovon es abhängen könnte, welche Cues umgewertet werden. Auf Grundlage der Ergebnisse in der Informations-Bedingung wurde vermutet, dass insbesondere die Cues, die häufig präsentiert werden, systematisch umgewertet werden. Dies gilt es in nachfolgenden Experimenten zu prüfen. Über alle Bedingungen beider Experimente hinweg betrachtet, ist die Umwertung eines Cues konstant: So wird der Cue „Augenzeuge“ in jeder Bedingung signifikant abgewertet. Vor dem Hintergrund der bisherigen Ausführungen erscheint es plausibel anzunehmen, dass dies auf die Cue-Konstellation zurückzuführen ist. So ist „Augenzeuge“ der einzige Cue, der in jedem Fall und bezogen auf die Vorhersage in der gleichen Konstellation auftrat, da er der einzige Schuld-Cue war. Zwar wurde auch „Ort“ in jedem Fall präsentiert, aber in wechselnden Konstellationen. In jedem Fall sprach also die Mehrheit der Cues gegen den Cue „Augenzeuge“, was zu seiner Abwertung geführt haben könnte. Die Frage, inwiefern die Wiederholung die Stärke der Veränderung beeinflusst, bleibt künftigen Forschungsarbeiten vorbehalten.

Eine offene Frage stellt erneut das Auftreten sogenannter Switcher dar. In den Experimenten 6 und 7 wurden die Personen als Switcher klassifiziert, deren tendenzielle sich von der finalen Entscheidung unterschied. Das Auftreten solcher Switcher wurde auch von Simon und Kollegen in einigen Experimenten beobachtet (Holyoak und Simon, 1999; Simon et al., 2001; Simon, Snow et al., 2004). Zwar räumen sie ebenfalls ein, dass die Ursache für diese Umkehrung unklar ist, doch erachten sie dies als Beleg für die Annahme, dass die entsprechenden Prozesse prädezisional und folglich Änderungen der Entscheidung möglich sind. Diese Erklärung kann zwar für die Experimente 6 und 7 herangezogen werden, obgleich hier auch *demand*-Effekte nicht auszuschließen sind. Eine Untersuchung des Auftretens von Switchern war nicht Gegenstand dieser Arbeit, da ein gewisses Umschwenken einer Minorität von Personen aber scheinbar fast immer zu beobachten ist, ist die Aufforderung von Holyoak und Simon (1999) zu wiederholen: “An important issue for future research will be to investigate the causes and consequences of such decision reversals” (S. 10).

4. DISKUSSION

Die allgemeine Zielsetzung der vorliegenden Arbeit bestand darin, einen Beitrag zur Erforschung des Strebens nach Konsistenz in der Entscheidungsforschung zu leisten. Die Forschung der letzten Jahre hat wichtige Fortschritte in diesem Bereich bewirkt: Auf empirischer Ebene konnte die breite Gültigkeit des Prinzips der Konsistenzmaximierung nachgewiesen werden, auf theoretischer Ebene haben sich PCS-Modelle als geeignete Erklärungsansätze erwiesen. Ausgehend von diesen Erkenntnissen wurde eine Arbeitsdefinition von Konsistenz vorgeschlagen, die das hier vertretene Verständnis von Konsistenz präzisiert, und damit wichtig für die empirischen Arbeiten ist: *Konsistenz bezeichnet einen Zustand der Einheitlichkeit, des Zusammenhangs und der Widerspruchsfreiheit von Elementen in einer bestimmten Situation, der im Zuge des Informationsverarbeitungsprozesses automatisch durch Reevaluationen hergestellt wird.* Um das Verständnis des Konsistenzstrebens zu vertiefen, erschien insbesondere die Erforschung von Charakteristika der Situation und der Person sinnvoll, die einen Einfluss auf die Stärke der Reevaluationen haben. Diese Faktoren könnten unter Rückgriff auf verschiedenste Entscheidungstypen untersucht werden, doch wurden bewusst probabilistische Inferenzentscheidungen verwendet. Dies ist vor allem damit zu begründen, dass Glöckner und Betsch (2008b) ein entsprechendes PCS-Modell vorgeschlagen haben, das sowohl der Bedeutung der Konsistenzmaximierung Rechnung trägt als auch eine formale Beschreibung dieses Prozesses ermöglicht. Die damit verbundenen Annahmen, dass Menschen eine Vielzahl an Informationen schnell und kompensatorisch verarbeiten können und dass eine Entscheidungsregel zur Beschreibung des Entscheidungsverhaltens ausreicht, konnten in zahlreichen Experimenten bestätigt werden. Dies unterstreicht die Erklärungskraft dieses Modells. Außerdem ist die Verwendung probabilistischer Inferenzentscheidungen besonders interessant, da Cues mit zwei Ausprägungen (vorhanden vs. nicht vorhanden) und die Einschätzung ihrer Validität im Vergleich zu anderen Informationen weniger Interpretationsspielraum bieten und damit besonders deutlich auf das Wirken des Konsistenzmaximierungsprozesses verweisen. Damit kann die Zielsetzung der vorliegenden Arbeit dahingehend spezifiziert werden, dass Einflussfaktoren auf die Konsistenzmaximierung in probabilistischen Inferenzentscheidungen untersucht wurden. Zur Untersuchung dieser Frage wurden insgesamt sieben Experimente und drei Simulationsstudien durchgeführt.

In diesem abschließenden Kapitel werden sowohl die empirischen Befunde als auch übergreifende theoretische Aspekte und Implikationen diskutiert. Zunächst wird eine zusammenfassende und kritische Erörterung der empirischen Befunde vor dem Hintergrund der Konsistenzforschung im Allgemeinen und des PCS-Modells von Glöckner und Betsch (2008b) im Speziellen vorgenommen (Kapitel 4.1). Hierauf aufbauend werden allgemeine Aspekte und offene Fragen bezüglich der Konsistenzforschung, bezogen auf den Bereich der Entscheidungspsychologie, besprochen (Kapitel 4.2). Die Perspektive weitet sich dann zu Implikationen und Perspektiven für die Entscheidungs-, die Konsistenzforschung und die Praxis (Kapitel 4.3). Eine Schlussbetrachtung in Kapitel 4.4 rundet nicht nur dieses Kapitel, sondern auch die vorliegende Arbeit ab.

4.1 Zusammenfassende Darstellung der empirischen Befunde

In den empirischen Arbeiten wurde eine Vertiefung des Verständnisses des Konsistenzmaximierungsprozesses in probabilistischen Inferenzaufgaben angestrebt, indem insbesondere der Einfluss situationaler und personaler Faktoren auf die Stärke der prädezisionalen Informationsumwertung untersucht wurde. Ausgehend von dem wichtigen Ergebnis, dass die systematische Umwertung von Cue-Validitäten tatsächlich auf die Reflexion über eine Entscheidungssituation zurückzuführen ist und kein maßgeblicher oder sensitivierender Einfluss des Pretests vorliegt, wurden situationale und personale Faktoren hinsichtlich ihres Einflusses auf die Konsistenzmaximierung untersucht. Bezüglich situationaler Faktoren konnte gezeigt werden, dass die Ausgeglichenheit der Informationslage von Relevanz ist. Während bei einer unausgeglichenen Informationslage überwiegend deutliche systematische Umwertungen zu verzeichnen waren, war dies bei einer ausgeglichenen Informationslage nicht der Fall. Des Weiteren konnte die Wiederholung einer Entscheidungssituation als Einflussfaktor identifiziert werden. Eine wiederholte Darbietung einer Entscheidungssituation führte nicht zu einer systematischen Umwertung von Cue-Validitäten, wie sie bei einmaliger Darbietung beobachtet werden konnte, sondern eher zu einer Abwertung. Eine systematische Umwertung fand dann statt, wenn eine zusätzliche Information gegeben wurde. Da in diesem Fall die Konfidenz in die Entscheidung signifikant höher war als bei einer reinen Wiederholung, deutet dies auf die Rolle der Verunsicherung hin. Ein replizierbarer Zusammenhang von personalen Faktoren und der Stärke der Umwertung konnte nicht nachgewiesen werden. Ferner zeigte sich, dass die

Umwertungen größtenteils moderat waren, da nicht jeder Cue umgewertet wurde, und dass die Ergebnisse zum Teil nicht ohne weiteres durch das PCS-Modell vorhergesagt werden konnten. Eine zusammenfassende Interpretation wird daher zunächst nur auf der inhaltlichen Ebene vorgenommen: So sprechen die Befunde insgesamt dafür, dass die systematische Umwertung von Cue-Validitäten eine Folge der Reflexion über die Entscheidungsaufgabe und damit der Informationsverarbeitung ist. Ob eine systematische Umwertung auftritt, ist allerdings eine Frage der Entscheidungssituation. Als Moderatoren konnten die Ausgeglichenheit der Informationslage und die Wiederholung identifiziert werden. Personale Einflussfaktoren konnten nicht ermittelt werden. Unter Berücksichtigung bisheriger Forschungsergebnisse scheint das Streben nach Konsistenz folglich ein fundamentaler Prozess zu sein, sofern gewisse Voraussetzungen erfüllt sind.

In den folgenden Kapiteln sollen diese Ausführungen näher erörtert werden. Neben der inhaltlichen Interpretation der Daten ist dann auch die Vorhersagekraft des PCS-Modells zu thematisieren. Folglich findet eine kritische Diskussion sowohl vor dem theoretischen Hintergrund der Konsistenzforschung im Allgemeinen als auch des PCS-Modells im Speziellen statt.

4.1.1 Der Einfluss des Pretests

Die Veränderung subjektiver Cue-Validitäten *within subjects* in einem Pretest-Posttest-Design zu erfassen bietet den Vorteil, dass erstens der Unterschiedlichkeit der Validitätseinschätzungen Rechnung getragen wird und zweitens die Umwertung in Abhängigkeit von der individuellen Entscheidung betrachtet werden kann. Da aber der mögliche Einfluss der Bearbeitung des Pretests auf nachfolgende Einschätzungen der Cue-Validitäten bislang ungeklärt war, galt es diese Frage zu untersuchen. Es konnte unter Verwendung eines Solomon-Viergruppenplans der wichtige Nachweis erbracht werden, dass weder ein maßgeblicher noch ein sensitivierender Einfluss des Pretests vorliegt. Damit kann geschlussfolgert werden, dass die systematische Reevaluation von Cue-Validitäten tatsächlich auf das Treatment, nämlich die systematische Reflexion über eine Entscheidungsaufgabe, zurückgeführt werden kann. Dieser Befund rechtfertigt nicht nur die Wahl des Untersuchungsdesigns, sondern stützt auch die Annahme, dass das Erhöhen von Konsistenz, das durch die systematische Informationsumwertung erzielt wird, eine natürliche Folge des Informationsverarbeitungsprozesses ist.

4.1.2 Charakteristika der Situation

Die überzeugende Evidenz für das Auftreten prädezisionaler Informationsverzerrungen veranlasste Brownstein (2003) zu der Aufforderung, nun moderierende Variablen zu untersuchen, um das Verständnis dieser Reevaluationen zu vertiefen. Diesem Aufruf folgend wurden Charakteristika der Situation und der Person (siehe Kapitel 4.1.3) hinsichtlich ihres Einflusses auf die Umwertung von Informationen – als Ausdruck des Erhöhens von Konsistenz (z.B. Glöckner und Betsch, 2008b; Russo et al., 2008; Simon, Snow et al., 2004) – in probabilistischen Inferenzentscheidungen untersucht. Vor dem Hintergrund des bisherigen Forschungsstandes erschien es sinnvoll, die Ausgeglichenheit der Informationslage zu variieren bzw. ihren Einfluss dann gezielt zu untersuchen und dem Einfluss der Wiederholung nachzugehen. Bezüglich der Informationslage zeigte sich insgesamt, dass bei einer unausgeglichenen Informationslage systematische Umwertungen von Cue-Validitäten auftreten, was bei einer ausgeglichenen Informationslage nicht der Fall war. Zudem bestand ein positiver Zusammenhang zwischen der Größe der Differenz zwischen den gewichteten Summen im Posttest, die sowohl durch die Informationslage als auch die Umwertung beeinflusst wird, und der Konfidenz. Die wiederholte Darbietung einer Entscheidungssituation führte im Gegensatz zu einer einmaligen Darbietung nicht zu einer systematischen Umwertung von Cues. Zu beobachtende Reevaluationen gingen vielmehr in die Richtung einer Abwertung. Eine zusätzliche Information, die auf die richtige Lösung verwies, führte hingegen wieder zu systematischen Umwertungen. Diese gingen mit höheren Konfidenzwerten als im Fall der reinen Wiederholung einher, und zwar auch bereits vor der finalen Entscheidung. Der Versuch, diese Befunde mit dem im PCS-Modell für probabilistische Inferenzaufgaben implementierten automatischen Prozess der Konsistenzmaximierung überzeugend zu erklären, stößt an Grenzen: Da hier dieser Mechanimus als Kern eines jeden Entscheidungsprozesses aufgefasst wird, sollten systematische Umwertungen in jedem Muster auftreten. Derartige Grenzen nimmt Glöckner (2008b) vorweg, indem er darauf hinweist, dass “the degree of contradiction in the network can be so high that it cannot be sufficiently resolved by PCS processes” (S. 273). Auf der einen Seite verdeutlicht diese Einschränkung, dass die fehlende Umwertung von Informationen nicht per se gegen PCS-Prozesse spricht. Auf der anderen Seite besteht die Schwierigkeit, dass nicht spezifiziert ist, *wann* ein Widerspruch zu hoch ist. Folglich kann diese Aussage zwar post-hoc herangezogen werden, vorhersagbar war der Einfluss der Ausgeglichenheit der Informationslage aber deshalb nicht. Nüchtern betrachtet bedeutet dies, dass das bisherige Verständnis des

Konsistenzmaximierungsprozesses beschränkt ist. Positiv formuliert leisten die Ergebnisse bezüglich der Informationslage damit einen Beitrag zum Verständnis des Konsistenzmaximierungsprozesses. Die Diskussion der Ergebnisse vor dem Hintergrund des PCS-Modells findet in Kapitel 4.1.4 statt. Für die inhaltliche Erklärung der Daten bedarf es nun aber eines Rückgriffs auf die breitere Basis der Konsistenzforschung. Dies gilt auch für den Einfluss der Wiederholung, für den zudem die obige Aussage hinsichtlich des Grades an Widerspruch nicht zutreffend erscheint. Da das PCS-Modell für einmalige Entscheidungssituationen formuliert wurde, konnte der Einfluss der Wiederholung zwar zunächst nur explorativ untersucht werden, doch erschien es angesichts des Mechanismus der Konsistenzmaximierung plausibel anzunehmen, dass die wiederholte Darbietung einer Entscheidungssituation auch zu systematischen Umwertungen führt. Dass dies nicht der Fall war und stattdessen die Verunsicherung eine Rolle gespielt zu haben scheint, macht ebenfalls eine Erweiterung der Perspektive erforderlich. Die Notwendigkeit, post-hoc Erklärungen zu generieren, wird hierbei als Nachteil und als Chance zugleich begriffen: Der Nachteil besteht darin, dass diese Erklärungen (noch) nicht empirisch überprüft sind und damit nur eine plausible Interpretations*möglichkeit* darstellen. Die Chance besteht in dem Versuch, unter Berücksichtigung verschiedener Befunde ein allgemeineres Bild der Informationsumwertung bzw. Konsistenzmaximierung zu zeichnen, sowie in der damit verbundenen Anregung weiterer Forschungsarbeiten.

Vor dem Hintergrund verschiedener empirischer Befunde sowie intuitiv hätte erwartet werden können, dass (1) eine ausgeglichene Informationslage zu stärkeren Umwertungen führt, da sie von stärkerer Ambiguität gekennzeichnet ist als eine unausgeglichene Informationslage (vgl. Russo et al., 1998; Simon, 2004) und damit schwieriger zu lösen ist (vgl. Brownstein, 2003), und dass (2) die Wiederholung einer Entscheidungssituation mindestens zu systematischen Umwertungen im Sinne einer Stabilisierung (vgl. Hoeffler und Ariely, 1999; Simon et al., 2008), wenn nicht gar zu stärkeren Umwertungen führt. In den dargestellten Experimenten wurden diese Erwartungen nicht bestätigt, vielmehr wurde das Gegenteil gefunden! Berücksichtigt man allerdings weitere Befunde aus der entscheidungspsychologischen Konsistenzforschung sowie zudem Spezifika der Entscheidungsaufgabe, kann – obige Einschränkung bezüglich der nachträglichen Generierung dieser Interpretation beachtend – eine mögliche Erklärung vorgeschlagen werden. Zunächst ist anzuführen, dass auf der einen Seite eine höhere Konfidenz in die *leading option* zu stärkerer Informationsumwertung führt (Russo et al., 2008) und auf der anderen Seite die Erhöhung

von Konsistenz das Treffen einer konfidenten Entscheidung (z.B. Simon, Snow et al., 2004) erlaubt. Konfidenz und Reevaluationen bedingen sich somit wechselseitig. Stellt man sich die Frage, wie initial eine *leading option* identifiziert werden kann, von der man zusätzlich überzeugt ist, so finden sich je nach Entscheidungstyp unterschiedliche Erklärungen. Während etwa bei Präferenzentscheidungen bestehende Vorlieben relevant sind, sind in probabilistischen Inferenzentscheidungen die Ausprägung der Cues und die Höhe der Cue-Validitäten von Bedeutung. In anderen Entscheidungen mögen persönliche Überzeugungen eine maßgebliche Rolle spielen. An dieser Stelle könnten zwei Spezifika probabilistischer Inferenzentscheidungen zu beachten sein: Es gibt eine objektiv richtige Lösung *und* die Informationen selbst sind nicht widersprüchlich. Zwar schließen beide Aspekte nicht das Auftreten prädezisionaler Umwertungen aus, doch stellen sie erschwerte Bedingungen dar, da eine richtige Lösung angestrebt wird *und* der Interpretationsspielraum eingeschränkt ist. Dies behindert eventuell die wechselseitige Einflussnahme von Konfidenz und Konsistenz: Erstens ist zwar die Identifikation einer *leading option* möglich, aber bei schwierigen Aufgaben eventuell mit weniger Konfidenz verbunden als bei anderen Entscheidungstypen. Zweitens könnte die Umwertung durch das Vorliegen vergleichsweise eindeutiger Informationen zumindest erschwert sein, weshalb gerade bei dem Fehlen einer deutlich überlegenen Alternative hierfür ein höherer kognitiver Aufwand erforderlich wäre. Beides ist vor allem dann der Fall, wenn die Informationslage keinen deutlichen Hinweis darauf gibt, welche Option richtig ist, oder wenn die Wiederholung einer Entscheidungssituation ohne Feedback und mit wechselnden Cue-Konstellationen zu Verunsicherung führt. In diesen Fällen könnte es eines erhöhten kognitiven Aufwandes für die systematische Umwertung oder einer Reduktion der Verunsicherung bedürfen, damit eine konfidente Entscheidung getroffen werden kann. Letzteres wurde bezüglich der Wiederholung gezeigt, indem eine zusätzliche und entscheidende Information gegeben wurde. Trotz fehlender Umwertung konnte in den dargestellten Experimenten eine Entscheidung getroffen werden, wobei aber eine andere Alternative auch nicht möglich war. Denkbar ist jedoch, dass die Konfidenz in die Entscheidung dann vergleichsweise niedriger ist, wofür es erste Evidenz gibt. Einschränkend ist aber zu erwähnen, dass der Zusammenhang zwischen der Höhe der Differenz zwischen den korrigierten gewichteten Summen im Posttest und der Konfidenz nicht vollends überzeugend war, da sich die Konfidenz nach einer Entscheidung mit einer unausgeglichenen Informationslage nur signifikant von der Bedingung, in der ein 3:3-Muster präsentiert wurde, unterschied. Allerdings waren die Umwertungen und die Konfidenz in der

Bedingung mit der unausgeglichenen Informationslage ohnehin vergleichsweise gering. Zudem konnte – deskriptiv betrachtet – der positive Zusammenhang zwischen der Konfidenz und der Differenz zwischen den gewichteten Summen im Posttest nicht auf bereits im Pretest bestehende Unterschiede zurückgeführt werden. Folglich sind diesbezüglich weitere Studien wünschenswert, doch wird die Annahme einer Wechselbeziehung zwischen der Konfidenz und Reevaluationen in probabilistischen Inferenzentscheidungen gestützt. Welche alternativen Möglichkeiten es in einer Situation, in der keine konsistente Lösung gefunden werden kann, geben kann, wird in Kapitel 4.2.1 thematisiert.

Insgesamt betrachtet impliziert die vorgeschlagene Interpretation, dass die Maximierung von Konsistenz an gewisse Mindestvoraussetzungen gebunden ist. Vergleicht man die Experimente zu probabilistischen Inferenzen mit denjenigen zu anderen Entscheidungsaufgaben, wird deutlich, dass die Maximierung von Konsistenz im ersten Fall tatsächlich unter erschwerten Bedingungen untersucht wird. Selbst in den Experimenten von Dekay et al. (2009), Bond et al. (2007) sowie Kostopoulou et al. (2009), in denen die Konsistenzmaximierung ebenfalls unter vergleichsweise schweren Bedingungen untersucht wurde, wurden Mindestvoraussetzungen geschaffen, und zwar durch die Manipulation der *leading option* (Bond et al., 2007), durch die Ambiguität der Informationen (Kostopoulou et al., 2009) oder durch die Komplexität der Aufgabe und die Erfassung von Präferenzen (DeKay et al., 2009). Die wechselseitige Beeinflussung von Konfidenz und Konsistenz bedarf eines Anstoßes: Entweder wird etwa eine Option prominent präsentiert oder die Informationen lassen genügend Spielraum, damit ein tendenzieller Favorit relativ leicht Stützung erfahren kann. Abbildung 14 fasst die vorgestellten Überlegungen schematisch zusammen.

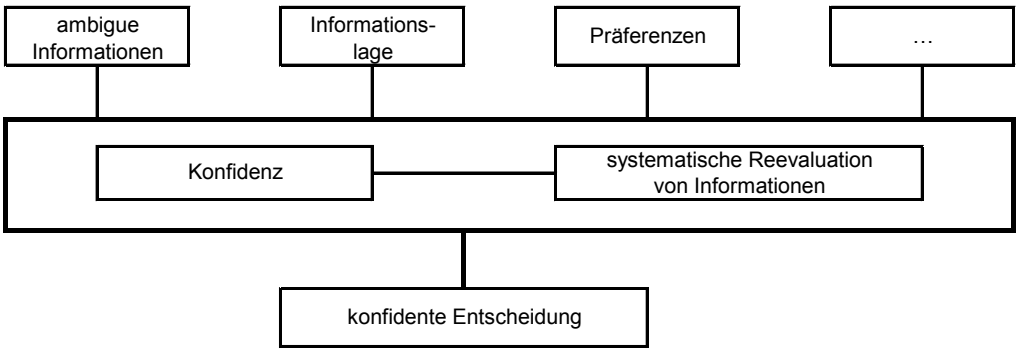


Abbildung 14: Schematischer Überblick über die möglichen Wirkbeziehungen

Demnach bedarf es gewisser Rahmenbedingungen bzw. Mindestvoraussetzungen, damit ein initiales Vertrauen in eine favorisierte Alternative besteht und systematische Reevaluationen einsetzen können, so dass schließlich eine konfidente Entscheidung getroffen werden kann. Ist dieser Schritt etwa durch die Beschaffenheit der Aufgabe oder der Informationen erschwert, könnte der Prozess der Konsistenzmaximierung behindert sein. Wichtig ist, dass dieser Erklärungsansatz weder in Frage stellt, dass das Erhöhen von Konsistenz eine automatische Folge der Informationsverarbeitung ist, noch dass es dem Treffen einer konfidenten Entscheidung dient. Es wird lediglich postuliert, dass eine erfolgreiche Konsistenzmaximierung minimaler Voraussetzungen bedarf. Interessanterweise wird dieser Aspekt (noch) nicht explizit in der entsprechenden Literatur diskutiert. Bezogen auf die vorgestellten Befunde bedeutet dies, dass im Falle einer ausgeglichenen Informationslage die Konfidenz in die führende Alternative nicht hoch ist und die Umwertung zudem durch die Art der Informationen (bedingte Wahrscheinlichkeiten) erschwert ist. Im Falle der reinen Wiederholung kann die Verunsicherung zu groß gewesen sein, als dass eine *leading option* mit Konfidenz hätte identifiziert werden können und systematische Reevaluationen hätten einsetzen können.

Abschließend sei noch einmal explizit darauf hingewiesen, dass diese Ausführungen das Ziel hatten, einen *möglichen* Erklärungsansatz für die Ergebnisse vorzustellen. Dass dieser vorläufig ist und der empirischen Bestätigung bzw. Revision bedarf, ist unvermeidlich. Allerdings wurde bewusst der Versuch unternommen bzw. die Herausforderung angenommen, nicht erwartete Ergebnisse zu erklären und so zu einem vertieften Verständnis des Konsistenzmaximierungsprozesses beizutragen. In jedem Fall ist zu hoffen, dass weitere diesbezügliche Bemühungen unternommen werden. Einige Vorschläge werden in Kapitel 4.3.1 unterbreitet.

4.1.3 Charakteristika der Person

Die Existenz individueller Unterschiede bezüglich des Strebens nach Konsistenz und deren Relevanz für die Informationsverarbeitung wurden sowohl in der Konsistenzforschung im Allgemeinen (z.B. Abelson und Rosenberg, 1958; Festinger, 1957; Rosenberg, 1960) als auch im speziellen Kontext der Entscheidungspsychologie (z.B. Simon, 2004; Svenson, 1992) diskutiert. Wird zudem berücksichtigt, dass unter Verwendung des gleichen Materials (Experiment 1/Bedingung 1, Experiment 3/Bedingung 2, Experiment 5) zwar stets systematische Umwertungen zu beobachten waren, aber in unterschiedlicher Stärke,

erscheint der Einfluss von Persönlichkeitscharakteristika umso plausibler. Systematisch untersucht wurde der Zusammenhang zwischen bestimmten Persönlichkeitseigenschaften und der prädezisionalen Informationsverzerrung als Ausdruck der Erhöhung von Konsistenz aber bisher nur vereinzelt (Beckmann und Kuhl, 1984; Meloy, 2000; Russo et al., 1998; Russo et al., 2000).

In der vorliegenden Arbeit wurden verschiedene Maße verwendet, für die ein Einfluss vermutet oder bereits nachgewiesen werden konnte. Insbesondere eine gewisse Abneigung gegenüber Ambiguität bzw. ein Streben nach Konsistenz erschien als potenzieller Einflussfaktor. Dieser wurde mehr oder weniger direkt über die Erfassung der Präferenz für Konsistenz (Cialdini et al., 1995), des *Need for Cognitive Closure* (Webster und Kruglanski, 1994) sowie der Tendenz zu maximieren (Schwartz et al., 2002) untersucht. Des Weiteren wurde aufgrund vorheriger Evidenz (Beckmann und Kuhl, 1984) ein Einfluss der Handlungsorientierung, genauer gesagt der Handlungsorientierung bei der Handlungsplanung, angenommen. Der Zusammenhang mit den *Big Five* als fundamentale Persönlichkeitsdimensionen wurde explorativ mit dem BFI-10 (Rammstedt und John, 2007) untersucht. Es konnte kein überzeugender Zusammenhang mit einem der Persönlichkeitsfaktoren nachgewiesen werden. Gefundene signifikante Korrelationen sind eher mit Vorsicht zu interpretieren: Der signifikante Zusammenhang mit der Präferenz für Konsistenz konnte nicht repliziert werden, die Skala zur Messung der Verträglichkeit erwies sich als problematisch und das Ergebnis, dass nur die Handlungsorientierung bei der Tätigkeitsausführung signifikant mit der Stärke der Umwertung korrelierte, während kein Zusammenhang mit der Handlungsorientierung bei der Handlungsplanung gefunden werden konnte, war nicht zu erwarten und bedarf der Replikation. Auf der einen Seite ist das Fehlen eines überzeugenden Zusammenhangs mit einer der Persönlichkeitsvariablen überraschend, da die Relevanz individueller Unterschiede häufig diskutiert wird. Auf der anderen Seite konnte auch bei der Mehrzahl der bisher untersuchten Faktoren kein entsprechender Zusammenhang gefunden werden (Meloy, 2000; Russo et al., 1998; Russo et al., 2000). Gewiss kann damit nicht ausgeschlossen werden, dass andere oder extreme Persönlichkeitsausprägungen einen Einfluss haben bzw. spezifischere Tests von Nöten sind. Allerdings stellt sich die Frage, welche weiteren Eigenschaften von Relevanz sein könnten. Gerade das naheliegendste Charakteristikum, nämlich eine Abneigung gegenüber Ambiguität, scheint – wie die vorgestellten und die bisherigen Experimente gezeigt haben – keinen Einfluss zu haben. Weitere Variablen, die fundamentale Persönlichkeitseigenschaften ausdrücken, hatten

ebenfalls keinen überzeugenden Einfluss. Folglich wird der Einfluss personaler Charakteristika als zu vernachlässigend erachtet. Vielmehr erscheint eine situative und damit vergängliche Anpassung des Bedürfnisses nach Konsistenz plausibel, veranlasst etwa durch eine hohe persönliche Relevanz der Aufgabe – eine Sichtweise, die bereits vertreten wird (z.B. Betsch, 2005a; Glöckner und Betsch, 2008b; Svenson, 1992). Dies gilt es in folgenden Experimenten zu prüfen (vgl. Kapitel 4.3.1). Dies unterstreicht erneut, dass es sich bei der systematischen und zu Konsistenz führenden Verzerrung von Informationen um einen basalen Prozess zu handeln scheint, der allerdings durch Aufgabenfaktoren beeinflusst werden kann.

4.1.4 Vorhersage der Ergebnisse durch das PCS-Modell

Eine wichtige Annahme im PCS-Modell von Glöckner und Betsch (2008b) besteht in der Auffassung, dass Menschen viele Informationen in einer parallelen, schnellen und kompensatorischen Weise verarbeiten können. Die empirische Stützung für diese Fähigkeit zur schnellen und kompensatorischen Informationsintegration stellt für Glöckner (2008a) eine in Modellen wie der *adaptive toolbox* (Gigerenzer et al., 1999) zentrale Annahme infrage, nämlich diejenige der begrenzten Rationalität (*bounded rationality*). Für die vorliegende Arbeit von besonderem Interesse ist aber, dass nach Glöckner (2008a) auch eine zweite Auffassung von Ansätzen, die die Verwendung rationaler oder aber einfacher Entscheidungsstrategien postulieren, anzuzweifeln ist: die Unveränderbarkeit von Informationen. Entsprechende empirische Belege berichten Glöckner et al. (im Druck) ausführlich. Dass Cue-Validitäten im Entscheidungsprozess modifiziert werden, konnte auch in den Experimenten dieser Arbeit bestätigt werden, allerdings war diese systematische Veränderung an Charakteristika der Entscheidungsaufgabe gebunden und betraf nicht immer alle Cues gleichermaßen. Bevor diese Einschränkung erörtert wird, soll zunächst hervorgehoben werden, dass die Entdeckung von Reevaluationen insofern beachtlich ist, als dass das Untersuchungsdesign auch hinderliche Bedingungen enthielt. Betrachtet man etwa die von Brownstein (2003) aufgelisteten Moderatoren, so fällt auf, dass bezogen auf die systematische Reevaluation (im Gegensatz zur selektiven Informationssuche) zwar die Ähnlichkeit der Alternativen, die Kenntlichmachung als Entscheidungsaufgabe und die fehlende Aufforderung zur Rechtfertigung der Entscheidung günstige Bedingungen für das Auftreten von prädezisionalen Umwertungen darstellen. Allerdings führen fehlende Konsequenzen, fehlender Zeitdruck sowie die Erwartung weiterer Informationen eher zu weniger

Umwertungen. Darüber hinaus ist aber insbesondere die im Vergleich zu anderen Informationen niedrigere Ambiguität (DeKay et al., 2009; Simon, 2004) zu beachten, die eher zu schwächeren Informationsumwertungen führen sollte. Einen derartigen Effekt hat vermutlich auch das Vorhandensein einer objektiv richtigen Lösung. Damit erfährt die Annahme prädezisionaler Informationsverzerrungen als Bestandteil des Entscheidungsprozesses zwar eindrücklich Stützung, inwiefern die Ergebnisse aber mit der im PCS-Modell implementierten Auffassung der automatischen Konsistenzmaximierung vereinbar sind, ist separat zu beurteilen. Hier gilt es insbesondere zwei Aspekte zu erörtern: (1) die Evidenz für systematische Umwertungen ist zum Teil lediglich moderat, (2) entsprechend dem Konsistenzmaximierungsprozess hätten Umwertungen in weiteren Bedingungen erwartet werden können. Beide Gesichtspunkte wurden in der Diskussion der jeweiligen Experimentalergebnisse bereits angesprochen und sollen nun ausführlicher erörtert werden.

Traten systematische Umwertungen auf, waren diese überwiegend moderat, da nicht jeder Cue systematisch und signifikant verändert wurde. Dies war mit dem PCS-Modell aber nicht vorherzusagen. Zwar wird immer wieder darauf hingewiesen, dass der in einem konnektionistischen Netzwerk implementierte Konsistenzmaximierungsprozess nicht zwangsläufig zu einem globalen Minimum an Energie (bzw. Maximum an Organisation) führt, sondern dass auch lokale Minima (bzw. Maxima) erreicht werden können (z.B. Read et al., 1997). Kritisch betrachtet stellt dies allerdings eine post-hoc-Erklärung dar. Ob ein globales oder lokales Minimum erreicht wird, konnte nicht antizipiert werden. Demzufolge weisen die Ergebnisse nun auf die Notwendigkeit vorhersagen zu können, welche Cues umgewertet werden. Die dargestellten Experimente zum Einfluss der Informationslage bieten keinen eindeutigen Anhaltspunkt für entsprechende inhaltliche Begründungen: So wurden bei gleichem Design entweder alle Cues (Experiment 3/Bedingung 2) oder ein Teil der Cues (Experiment 1/Bedingung 1, Experiment 5) oder kein Cue (Experiment 4/Bedingung 2) umgewertet. Selbst bei gleichem Material (Experiment 1/Bedingung 1, Experiment 3/Bedingung 2, Experiment 5) waren die Ergebnisse uneinheitlich. Allerdings wurde hier stets der Cue „Bild“ aufgewertet, der im Pretest eine sehr niedrige Validitätseinschätzung hatte. Damit hat dieser Cue zwar eine höhere Chance zur Aufwertung, allerdings wurde er auch bei einer ausgeglichenen Informationslage verwendet, in der er nicht umgewertet wurde. Folglich scheint die Cue-Konstellation in Zusammenhang mit der initialen Einschätzung zu beachten zu sein. Einen weiteren Anhaltspunkt geben die Experimente zum Einfluss der Wiederholung: Traten systematische Umwertungen bei

wiederholter Darbietung auf (Experiment 7/Bedingung 2), zeigte sich deskriptiv ein Zusammenhang zwischen der Wiederholungshäufigkeit und der Umwertung. Die Frage, ob diese Beobachtung bestätigt werden kann, ist weiteren Forschungsarbeiten vorbehalten. Die plausible Vermutung, dass stets der Cue signifikant abgewertet wird, der als einziger gegen eine Mehrheit von Cues spricht, konnte – betrachtet über alle Experimente – in der Mehrzahl der entsprechenden Bedingungen bestätigt werden. Obgleich also die eher moderate Evidenz nicht zu erwarten war, spricht sie nicht per se gegen das Wirken von Konsistenzmaximierungsprozessen. Dennoch bedarf es nun gezielter Forschungsarbeiten, um die Frage, welche Cues umgewertet werden, zu klären.

Der zweite Aspekt, den es zu erörtern gilt, betrifft das Ergebnis, dass systematische Umwertungen nicht in jeder Bedingung zu verzeichnen waren. Da der Konsistenzmaximierungsprozess als Kernmechanismus des Entscheidungsprozesses verstanden wird (Glöckner und Betsch, 2008b), war eine systematische Reevaluation zumindest auch in ausgeglichenen Informationslagen zu erwarten. Das PCS-Modell wurde für einmalige Entscheidungssituationen formuliert, so dass eine Hypothese bezüglich der Informationsumwertung bei wiederholter Darbietung einer Entscheidungssituation nicht stringent abgeleitet werden konnte. Es war aber plausibel anzunehmen, dass die direkt aufeinander folgende Darbietung mehrerer identisch konzipierter Entscheidungssituationen zumindest zu einer gleich starken systematischen Reevaluation führen sollte. Diese Erwartungen konnten aber nicht bestätigt werden. Bei einer ausgeglichenen Informationslage fand keine deutliche systematische Umwertung statt, vereinzelt auftretende systematische Reevaluationen eines Cues ließen sich nicht replizieren. Die Simulationsstudien stützten diese Beobachtung, da in den entsprechenden Bedingungen die mit dem PCS-Modell vorhergesagten finalen Aktivationen keinen signifikanten Beitrag zur Vorhersage der Validitätseinschätzungen im Posttest machten. Wie bereits in Kapitel 4.1.2 angesprochen, kann post-hoc vermutet werden, dass aufgrund der Ausgeglichenheit der Informationslage der Widerspruch zu hoch ist, so dass im primären Netzwerk keine konsistente Lösung automatisch gefunden werden kann (Glöckner, 2008b). Kritisch anzumerken ist, dass es der Spezifikation bedarf, *wann* der Widerspruch zu hoch ist – vor allem vor dem Hintergrund der bisherigen Ergebnisse der Konsistenzforschung, nach denen ein gewisses Maß an Widerspruch eher förderlich für eine Umwertung ist. Ansonsten könnte – überspitzt formuliert – der nachträgliche Hinweis auf einen zu hohen Widerspruch willkürlich werden. Zu einer ersten Klarifizierung tragen die vorgestellten Befunde bei, da sie auf einen Indikator für Widerspruch, nämlich die Ausgeglichenheit der Informationslage,

hinweisen. Hierbei gilt es aber auch die Cue-Dispersion zu beachten, die in den vorliegenden Experimenten relativ gering war (siehe Anhang B für einen deskriptiven Überblick über die Validitätseinschätzungen). Bezüglich des Einflusses der Wiederholung einer identisch konzipierten Entscheidungssituation konnte keine eindeutige Hypothese generiert, sondern lediglich eine plausible Erwartung formuliert werden. Diese gründete auf der Überlegung, dass die Wiederholung einer Entscheidungssituation als wiederholtes Wirken des Konsistenzmaximierungsprozesses verstanden werden kann. Dass diese Erwartung nicht bestätigt werden konnte, spricht damit nicht gegen, aber auch nicht für die Vorhersagekraft des PCS-Modells. Vielmehr widersprechen die Ergebnisse der obigen Überlegung und verweisen auf die Notwendigkeit, neben dem basalen Prozess der Konsistenzmaximierung weitere Variablen wie die Verunsicherung zu berücksichtigen. Einschränkend ist zu erwähnen, dass im vorliegenden Kapitel nur die Vorhersagekraft des PCS-Modells für den spezifischen Aspekt der Konsistenzmaximierung diskutiert werden konnte. Gesichtspunkte wie die Rolle deliberativer Prozesse oder die Modifikation der Konsistenzschwelle, die in dem umfassenden integrativen Modell von Glöckner und Betsch (2008b) angesprochen werden und Ansatzpunkte für Erklärungen darstellen, waren nicht Bestandteil dieser Arbeit und sind daher eher unter einer erweiterten Perspektive (siehe Kapitel 4.2 und 4.3) zu erörtern.

Zusammenfassend ist festzuhalten, dass die Vorhersagekraft des PCS-Modells für die vorliegenden Befunde beschränkt ist. Dies liegt zum einen daran, dass der Konsistenzmaximierungsprozess bisher nur für einmalige Entscheidungssituationen formuliert wurde. Zum anderen ist die begrenzte Vorhersagekraft fehlender Präzisierungen geschuldet: So kann die Maximierung von Konsistenz zwar als Kernmechanismus jedes Entscheidungsprozesses verstanden werden, allerdings ist unter bestimmten Bedingungen wie einem zu hohen Widerspruch das automatische Finden einer konsistenten Lösung nicht möglich. Zudem ist keine optimale Lösung garantiert, was sich empirisch in der systematischen Umwertung von lediglich einzelnen Cues ausdrückt. Problematisch ist, dass die genannten Ansatzpunkte zwar Bestandteile des Modells sind, aber dennoch nur post-hoc generiert werden konnten. Es war nicht möglich vorherzusagen, wann genau der Widerspruch so hoch ist, dass der automatische Konsistenzmaximierungsprozess erfolglos verläuft, oder wann nur ein lokales Energie-Minimum erreicht werden kann. Eine negative Interpretation dieser Befundlage könnte in dem Vorwurf münden, man könne post-hoc vieles erklären, aber wenig vorhersagen. Dies zu vermeiden kann durch eine konstruktive Sichtweise, die der

Wissenschaft eine aktive Rolle beimisst, gelingen: So ist zu beachten, dass PCS-Modelle wie das von Glöckner und Betsch (2008b) ihr Potential hinsichtlich der Abbildung der menschlichen Informationsverarbeitung unter Beweis gestellt haben. Das PCS-Modell für probabilistische Inferenzentscheidungen bietet nun Anhaltspunkte für Erklärungen hinsichtlich des spezifischen Aspekts der Konsistenzmaximierung, die allerdings einer Präzisierung bzw. näheren Untersuchung bedürfen. Ob diese dann der empirischen Prüfung standhalten, ist eine Frage für künftige Forschungsarbeiten.

4.1.5 Kritische Aspekte

Zum Abschluss dieses Kapitels soll kurz auf einige kritische Aspekte hingewiesen werden, die es grundsätzlich zu berücksichtigen gilt. Dies betrifft erstens die explizite Erfassung der subjektiven Cue-Validitäten. In Kapitel 3.2 wurde begründet, weshalb diese Vorgehensweise gerechtfertigt ist. Die Beibehaltung der Messmethode ermöglichte es zudem, die Experimente dieser Arbeit zu vergleichen. Dennoch erscheint es im Sinne eines multimethodalen Ansatzes angebracht, zukünftig mehrere Methoden anzuwenden. Eine Registrierung der Blickbewegung könnte zusätzlich Aufschluss darüber geben, ob die Auf- und Abwertung von Cues mit unterschiedlichen Blickbewegungen einhergehen. Entsprechende Erkenntnisse könnten zu einem vertieften Verständnis des prädezisionalen Informationsverarbeitungsprozesses beitragen. Ebenfalls in Kapitel 3.2 diskutiert wurde die Abgrenzung prä- und postdezisionaler Prozesse. Das in den dargestellten Experimenten gewählte Vorgehen erscheint vor dem Hintergrund der bisherigen Forschungsergebnisse angemessen, um die Informationsumwertung *im* Entscheidungsprozess zu untersuchen. Obgleich es unwahrscheinlich ist, kann natürlich dennoch nicht ausgeschlossen werden, dass sich die Versuchspersonen nicht nur tendenziell, sondern final entschieden. Da diese Entscheidung aber nicht bindend war, können postdezisionale Prozesse der Dissonanzreduktion oder Konsolidierung ausgeschlossen werden (vgl. auch Glöckner et al., im Druck; Simon et al., 2001). Interessanter als die Abgrenzung erscheint aber die Frage, wie sich das Zusammenwirken prä- und postdezisionaler Prozesse (vgl. Svenson, 1992; Svenson, 1996) darstellt. Dieser Aspekt konnte aufgrund der vorgenommenen Schwerpunktsetzung nicht untersucht werden, stellt aber ein interessantes künftiges Forschungsvorhaben dar. Der dritte kritische Punkt, der angesprochen werden soll, betrifft die Charakterisierung prädezisionaler Informationsumwertungen als Methode der Konsistenzmaximierung. Dass diese zutrifft, wurde in Kapitel 2 ausführlich dargelegt. Umkehrt ist aber zu fragen, welche alternativen

Wege es zur Erhöhung der Konsistenz gibt. Die in anderen Studien untersuchte inhaltliche Umwertung oder selektive Erinnerung von Informationen konnten nicht betrachtet werden. In diesem Zusammenhang ist abschließend auch darauf hinzuweisen, dass nur ein spezifischer Aspekt des integrativen Modells von Glöckner und Betsch (2008b), nämlich die automatische Konsistenzmaximierung im primären Netzwerk, betrachtet wurde. Welche Rolle deliberate Prozesse oder das Absenken der Konsistenzschwelle bei der Erhöhung von Konsistenz spielen, konnte und sollte nicht betrachtet werden. Letztlich sind die angesprochenen Aspekte Resultat einer notwendigen und gerechtfertigten Fokussierung. Dass sie an dieser Stelle angesprochen wurden, soll aber die Grenzen dieser und Möglichkeiten künftiger Untersuchungen aufzeigen.

4.2 ALLGEMEINE DISKUSSIONSPUNKTE UND OFFENE FRAGEN

Nachdem im vorangegangenen Kapitel die Ergebnisse vor dem Hintergrund der Konsistenzforschung im Allgemeinen und des PCS-Modells für probabilistische Inferenzen im Speziellen diskutiert worden sind, soll im folgenden Kapitel die Perspektive erweitert werden. Ausgehend von den Befunden dieser Arbeit werden allgemeinere Diskussionspunkte und offene Fragen thematisiert, die für die Untersuchung und Charakterisierung der Konsistenzmaximierung in der Entscheidungspsychologie von Relevanz sind.

4.2.1 Die Grenzen der Konsistenzmaximierung?

Das Streben nach Konsistenz stößt an Grenzen. Die Arbeiten der letzten Jahre haben darauf fokussiert zu zeigen, dass prädezisionale Informationsverzerrungen unter verschiedenen Bedingungen auftreten. Diese Schwerpunktsetzung ist aus zwei Gründen verständlich: Erstens galt es den überzeugenden Nachweis zu erbringen, dass neben postdezisionaler Dissonanzreduktion auch prädezisionale Informationsumwertungen auftreten. Zweitens waren vermutlich die weite Verbreitung und der fundamentale Charakter dieses Phänomens erstaunlich und weckten Interesse. Auf diesem Wege konnte herausgearbeitet werden, dass die prädezisionale Reevaluation von Informationen als Mechanismus zur Erhöhung bzw. Maximierung von Konsistenz verstanden werden kann, der eine automatische Folge des Informationsverarbeitungsprozesses ist. Nachdem sich die Auffassung durchgesetzt hatte, dass Informationen bereits im Entscheidungsprozess systematisch umgewertet werden, galt es moderierende Faktoren zu identifizieren und zu untersuchen. Ein entsprechender Versuch

wurde auch in dieser Arbeit unternommen. Hierbei zeigte sich unerwarteterweise, dass unter bestimmten Bedingungen keine systematischen Validitätsumwertungen (oder wenige) als Ausdruck des Konsistenzmaximierungsprozesses zu beobachten waren. Was bedeutet dies für die Charakterisierung des Konsistenzmaximierungsprozesses? Angesichts der relativ geringen Anzahl an Experimenten sind eindeutige Aussagen verfrüht. Doch unter dem Vorbehalt, dass weitere empirische Stützung von Nöten ist, lassen sich erste Vermutungen anstellen. So könnte es fruchtbar sein, den Blick auch auf etwaige Minimalvoraussetzungen zu richten, die für ein erfolgreiches Wirken bzw. das Auftreten von Konsistenzmaximierungsprozessen erforderlich sind. Mögliche Kandidaten wurden bereits in Kapitel 4.1.2 vorgeschlagen. Es wurde argumentiert, dass wenn zum Beispiel die Widersprüchlichkeit der Aufgabe so hoch ist, dass keine Option mit Konfidenz präferiert wird, und zudem der Interpretationsspielraum der Informationen aus verschiedenen Gründen begrenzt ist, eventuell keine erfolgreiche Konsistenzmaximierung möglich ist. Diese Einschätzung ist vorläufig, und zwar sowohl aufgrund der nachträglichen Generierung als auch aufgrund der mangelnden Präzision. Dennoch wird die Auffassung vertreten, dass die Untersuchung minimaler Voraussetzungen zu einem Verständnis des Konsistenzmaximierungsprozesses beitragen kann. Explizit sei darauf hingewiesen, dass weder das prinzipielle Streben nach Konsistenz noch die Charakterisierung des Konsistenzmaximierungsprozesses als natürliche Folge des Informationsverarbeitungsprozesses in Frage gestellt werden. Es wird lediglich postuliert, dass eine Erhöhung von Konsistenz durch systematische Informationsumwertungen nicht immer möglich ist. Wie wir etwa ein Objekt nicht sehen können, wenn es durch ein Hindernis verdeckt ist, können wir eventuell keine konsistente Lösung generieren, wenn gewisse Minimalkonditionen nicht vorliegen. Interessant ist also, wann Konsistenzmaximierungsprozesse *nicht* auftreten. Damit kann ein von Cialdini et al. (1995) formuliertes Problem als Chance betrachtet werden:

> Except in their most fundamental forms, consistency effects are not easily obtained or replicated. (…) Despite the large number of studies that supported dissonance predictions, certain investigators complained privately of the difficulty of building coherent, dissonance-based research programs because of a vexing inability to produce or reproduce predicted effects with sufficient power or frequency. (S. 318)

Dies ist für Cialdini et al. ein Grund für die Abkehr von der Untersuchung von Konsistenzphänomenen und verweist für sie auf die Möglichkeit der Einflussnahme individueller

Unterschiede. Ob auch die Untersuchung prädezisionaler Informationsverzerrungen mit der Schwierigkeit, stabile Effekte zu finden, zu kämpfen hat, kann an dieser Stelle nicht beurteilt werden – doch verdeutlicht das Zitat die Notwendigkeit, auch im Sinne der ursprünglichen Intention gescheiterte Experimente zu berichten. Das Streben nach Konsistenz stößt hinsichtlich eines weiteren Aspektes an Grenzen: So wurde oftmals nur ein Teil der Cues umgewertet. Dass einige Informationen resistent gegenüber einer Umwertung sind, wurde bereits früh festgehalten (z.B. Asch, 1946) und in einigen Konsistenztheorien berücksichtigt (z.B. Festinger, 1957; Osgood und Tannenbaum, 1955). Hinsichtlich der prädezisionalen Informationsumwertung ist aber noch unklar, welche Informationen eher der Veränderung unterliegen. Interessanterweise fanden etwa Russo et al. (2000) einen positiven Zusammenhang zwischen der Wichtigkeit von Attributen und der Stärke der Umwertung, wobei dies aber nur auf die Berufsgruppe der Verkäufer zutraf. Ob es hier tatsächlich einen Zusammenhang gibt, ist weiter zu untersuchen. Einflüsse der Art der Erfahrung sind hier ebenso denkbar. Dass nicht immer alle Informationen umgewertet werden, verdeutlicht noch einmal, dass die Konsistenzmaximierung vor allem als Prozess zu verstehen ist. Es ist aber – wie bereits angesprochen wurde – nicht garantiert, dass eine maximal konsistente Lösung resultiert. Warum dies auch manchmal gar nicht erwünscht ist, wird in Kapitel 4.2.2 thematisiert.

Geht man nun von den Entscheidungssituationen aus, in denen Konsistenz zwar angestrebt wird, aber nicht zufriedenstellend im Zuge der prädezisionalen Informationsumwertung erreicht werden kann, stellt sich die Frage, wie der Entscheidungsprozess weiter verläuft und damit die Grenzen der Konsistenzmaximierung überschritten werden können. Letztlich wird hiermit angesprochen, welches Ziel mit der Maximierung von Konsistenz erreicht werden soll. Dieser Aspekt wurde in der vorliegenden Arbeit weitestgehend ausgeklammert, so dass lediglich einige Überlegungen dargestellt werden sollen. Wie aus den Ausführungen des Kapitels 2 deutlich geworden ist, ist die Erhöhung von Konsistenz einerseits selbst erstrebenswert, sie kann aber auch – je nach Entscheidungstyp – dazu verhelfen, eine ausreichend konfidente Entscheidung zu treffen oder konkurrierende Handlungstendenzen abzuwehren. In der vorliegenden Arbeit war das Treffen einer ausreichend konfidenten Entscheidung wichtig. Nach dem Modell von Glöckner und Betsch (2008b; siehe auch Betsch, 2005a) kann nun entweder die Konsistenzschwelle θ gesenkt werden oder es können deliberative Prozesse einsetzen, damit eine Entscheidung getroffen werden kann. Nach Betsch (2005a) ist ersteres besonders dann der Fall, wenn zeitliche oder

kognitive Beschränkungen vorliegen. Dient die Konsistenzmaximierung etwa der Gewährleistung der Handlungsfähigkeit, so scheint gerade das Absenken der Konsistenzschwelle nicht auszureichen. Es ermöglicht zwar das Treffen einer Entscheidung, jedoch sind postdezisionale Konflikte nicht auszuschließen. Statt die Abgrenzung zwischen prädezisionalen und postdezisionalen Prozessen zu betonen, erscheint es sinnvoller, ihr Zusammenwirken zu betrachten. Eine solche Perspektive nimmt Svenson (1992; Svenson, 1996) in seiner Diff Con-Theorie ein. Inwiefern die Berücksichtigung postdezisionaler Prozesse von Relevanz ist, hängt gewiss – wie bereits angesprochen – vom Entscheidungstyp ab und kann hier nur als Ausblick dargestellt werden. Außerdem ist – unabhängig vom Entscheidungstyp – das Verhindern einer Entscheidung denkbar, das weder in den vorliegenden Experimenten noch in den betrachteten Studien eine Alternative darstellte, obgleich diese Möglichkeit des Öfteren erwähnt wird (z.B. Anderson, 2003; Montgomery und Willén, 1999). Hier könnte auch gefragt werden, unter welchen Umständen eher eine Entscheidung vermieden wird anstatt die anderen Mechanismen anzuwenden.

4.2.2 Vor- und Nachteile der Konsistenzmaximierung

Ausgehend von der Beobachtung, dass in bestimmten Entscheidungssituationen keine prädezisionalen Informationsverzerrungen auftreten, stellt sich die Frage, ob dies eher vorteilhaft oder nachteilig ist. Die hiermit angesprochene Diskussion über die positiven und die negativen Aspekte systematischer Umwertungen als Ausdruck der Konsistenzmaximierung wird in der entsprechenden Literatur oft aufgegriffen. Diese baut auf einer prinzipiell positiven Sichtweise, in der die Umwertung von Informationen nicht als „Fehler“, sondern vielmehr als funktional und als natürliche Folge des Informationsverarbeitungsprozesses betrachtet wird (z.B. Read et al., 1997; Russo et al., 1996; Simon, 2004), auf. Als spezifischer Vorteil ist zunächst die Vermeidung eines unangenehmen Zustands der Dissonanz (Elliot und Devine, 1994) zu nennen. Folglich wird allein das Erreichen von Konsistenz als erstrebenswertes Ziel erachtet (Russo et al., 2008). Besonders häufig wird erwähnt, dass die systematische Reevaluation von Informationen adaptiv ist, da sie die Entscheidungsfindung vereinfacht (z.B. Bond et al., 2007; DeKay et al., 2009; Simon, 2004; Simon, Snow et al., 2004). In diesem Zusammenhang wird ebenfalls darauf hingewiesen, dass auf diesem Wege eine *konfidente* Entscheidung getroffen werden kann, und zwar selbst bei widersprüchlichen und komplexen Informationen (z.B. Simon, 2004; Simon, Snow et al., 2004). Nach Bond et al. (2007) kann die Wahrnehmung von Konsistenz

– direkt oder vermittelt über die Konfidenz in eine Entscheidung – auch zum persönlichen Wohlergehen beitragen. Sogar die Aufhellung einer negativen Stimmungslage (Meloy, 2000) erscheint ebenso möglich wie Selbstbestätigung (Meloy, 2000; Russo et al., 1996). Beckmann und Kuhl (1984) grenzen sich von den kognitiven Konsistenzansätzen ab und betonen vielmehr die Gewährleistung bzw. Aufrechterhaltung der eigenen Handlungsfähigkeit als Vorteil der systematischen Informationsumwertung. Neben diesen Vorteilen werden aber auch einige Nachteile aufgelistet. Zunächst ist hier der offensichtliche Verlust an Genauigkeit zu nennen (z.B. Bond et al., 2007; DeKay et al., 2009). Dass die resultierenden Entscheidungen auf verzerrten Informationen basieren, wird insbesondere im Hinblick auf rechtliche Entscheidungen als weiterer problematischer Aspekt diskutiert (z.B. Hope et al., 2004; Simon, 2004). Ebenso wird die Gefahr der Manipulation gesehen, indem etwa Reihenfolge-Effekte zur Überzeugung einer Zuhörerschaft genutzt werden oder ein produzierendes Unternehmen positiv dargestellt wird, um die Kaufbereitschaft beim Konsumenten zu erhöhen (z.B. Bond et al., 2007; Russo et al., 2006). Einen wichtigen Nachteil sprechen Shultz und Lepper (1996) zwar speziell hinsichtlich der postdezisionalen Dissonanzreduktion an, doch er trifft ebenso auf die prädezisionale Konsistenzmaximierung zu: Die systematische Umwertung von Informationen verhindert Lernen. Wird etwa im Zuge einer Konsumentenentscheidung eine bestimmte Herstellermarke aufgewertet, so kann dies dazu führen, dass das Produkt stets von dieser Marke gekauft wird, obgleich es ein besseres oder günstigeres Produkt gibt. Damit muss die Entscheidung nicht falsch sein, doch je nach Interesse des Konsumenten kann sie suboptimal sein. Einschränkend ist zwar darauf hinzuweisen, dass Informationsumwertungen lediglich temporär (Simon, 2004; Simon et al., 2008) sind, doch deuten Befunde zum Einfluss der Wiederholung von Präferenzentscheidungen (Hoeffler und Ariely, 1999) auf die Möglichkeit dauerhafter Effekte hin. Zudem konnten Transfereffekte nachgewiesen werden, die ebenfalls den Entscheidungsprozess maßgeblich beeinflussen (Holyoak und Simon, 1999). Aus diesen Ausführungen ist deutlich geworden, dass bei der Beurteilung, wann eher die Vorteile und wann eher die Nachteile der prädezisionalen Informationsverzerrung im Vordergrund stehen, die jeweilige Entscheidungsaufgabe zu berücksichtigen ist. Während in Präferenzentscheidungen die Konfidenz wichtiger sein kann als die Genauigkeit, wird letztere eher hinsichtlich riskanter (DeKay et al., 2009) oder rechtlicher Entscheidungen (Simon, 2004) hervorgehoben. Interessant ist zudem die Auffassung, dass auch Inkongruenz ihren Reiz hat (McGuire, 1968). In diesem Sinne hielt bereits Zajonc (1960) fest: “People like to make sense of their world, but they

also seek out the magician to be entertained by incongruity" (S. 280). Darüber hinaus kann sie sogar förderlich für den Entscheidungsprozess sein: Svenson (1996) berichtet, dass weniger wichtige Attribute durchaus inkompatibel mit der gewählten Alternative sein können, da dies zu höherer Stabilität führe, die wiederum für die Umsetzung der Entscheidung wichtig sei. Allerdings erscheint hier stets ein geringes Maß an Inkonsistenz tolerierbar, wohingegen prinzipiell ein Streben nach Konsistenz vorherrscht – inklusive der Vor- und Nachteile.

4.2.3 Methoden des (De-)Biasing

Die Beobachtung prädezisionaler Informationsverzerrungen und deren potenziell negative Auswirkungen haben zu der Frage geführt, welche Methoden des *Debiasing* es gibt. In Kapitel 2.4.2 wurde bereits dargestellt, dass diesbezüglich zwar einige Vorschläge gemacht wurden, aber kaum erfolgreiche empirische Untersuchungen stattgefunden haben. Da gerade das mangelnde Bewusstsein der Reevaluationen als problematisch erachtet wird (z.B. Russo et al., 2000), sehen etwa Carlson und Russo (2001) einen ersten Ansatzpunkt darin, dieses Phänomen bewusst zu machen. In ähnlicher Weise hat sich Nickerson (1998) zum verwandten Phänomen des *confirmation bias* geäußert: Wird darauf hingewiesen, welche Entscheidungsfehler mit dem *confirmation bias* vermieden werden sollen, könnte dies der Verzerrungstendenz entgegenwirken. Auch die *consider the opposite*-Technik (Lord et al., 1984) wird von verschiedenen Autoren als sinnvoll erachtet (z.B. Simon, 2004). Aus organisationaler Perspektive schlagen Larson und King (1996) die Etablierung einer Lernkultur vor. Empirische Stützung fanden bisher die Technik des *consider the opposite* (Simon, 2004) sowie eine vor dem eigentlichen Entscheidungsprozess stattfindende Evaluation einzelner Attributskomponenten (Carlson und Pearo, 2004). Vor dem Hintergrund der insbesondere von der Forschergruppe um Russo erzielten Ergebnisse, dass Reihenfolge-Effekte auftreten (z.B. Russo et al., 2006; Russo et al., 1998), erscheint zudem eine simultane Informationspräsentation denkbar. Obgleich diese nicht in allen Kontexten umsetzbar ist, könnte sie etwa für den Bereich der Konsumentenentscheidungen eine Möglichkeit darstellen (vgl. Payne et al., 1988). Prinzipiell gilt zu beachten, dass es schwierig erscheint, effektive Strategien des *Debiasing* zu finden (Carlson und Pearo, 2004; Simon, 2004). Nach den Ausführungen des Kapitels 4.2.2 erscheint es angebracht, bei der Bewertung dieser Beobachtung eine differenzierte Sicht einzunehmen. Schließlich stehen je nach Entscheidungsaufgabe bzw. Zielsetzung die Vor- oder die Nachteile der prädezisiona-

len Informationsverzerrung im Vordergrund. Demnach ist die fehlende Möglichkeit des *Debiasing* dann problematisch, wenn genaue und objektive Entscheidungen zu treffen sind. Besteht das Ziel aber im Treffen einer konfidenten Entscheidung und ist diese nicht mit hohen Kosten verbunden, ist die angesprochene Schwierigkeit zu vernachlässigen. In einem solchen Fall kann *Debiasing* sogar nachteilig sein: "For consumers who value feeling confident in their choice, such debiasing may sacrifice more in confidence than it gains in accuracy" (Carlson und Pearo, 2004, S. 57). An dieser Stelle stellt sich sogar die Frage, ob eine gezielte Verstärkung der systematischen Informationsverzerrung manchmal sinnvoll sein kann. Dies ist zunächst durch eine Berücksichtigung der in Kapitel 2.4 vorgestellten Moderatoren möglich. Wird aber das Treffen einer konfidenten Entscheidung etwa durch Faktoren wie die Informationslage und eine hohe Diagnostizität der Informationen behindert, könnte ein direktes *Biasing* dem Entscheidungsprozess förderlich sein. Cialdini et al. (1995) schlagen für eine Erhöhung der Konsistenzmotivation z.B. vor, diese salient zu machen und Konsistenz als erstrebenswertes Ziel darzustellen. Ebenso ist es denkbar, die vorher als problematisch diskutierten Reihenfolge-Effekte gezielt zu nutzen. Damit rückt man aber eher in die Nähe einer bewussten Informationssuche, die in dieser Arbeit nicht thematisiert wird. Wann ein *Debiasing* und wann ein *Biasing* sinnvoll ist, hängt von der Entscheidungsaufgabe ab. Sowohl die prädezisionale Informationsverzerrung als auch Strategien des *(De-)Biasing* sind somit hinsichtlich ihrer Adaptivität zu beurteilen. Auch bezüglich anderer kognitiver Urteilstendenzen kann nicht unabhängig von der Situation beurteilt werden, ob sie vorteilhaft oder nachteilig sind bzw. ob ihnen entgegengewirkt werden soll oder nicht (Nickerson, 1998).

4.3 IMPLIKATIONEN UND PERSPEKTIVEN

Ausgehend von den bisherigen Ausführungen werden im folgenden Kapitel Implikationen für die Entscheidungs- und Konsistenzforschung sowie perspektivisch Forschungsanregungen aufgezeigt. Des Weiteren wird der Versuch unternommen, die Bedeutung prädezisionaler Informationsverzerrungen für verschiedene praktische Kontexte darzustellen.

4.3.1 Für die Entscheidungsforschung

Die Erkenntnisse dieser Arbeit leisten einen wichtigen Beitrag für die Entscheidungsforschung. Der Befund, dass subjektive Cue-Validitäten im Entscheidungsprozess verändert

werden (Glöckner et al., im Druck), konnte nicht nur bestätigt werden, sondern es konnte darüber hinaus gezeigt werden, dass diese Veränderung tatsächlich auf das Treatment zurückzuführen ist. Damit wird für probabilistische Inferenzentscheidungen die Annahme bidirektionaler Prozesse gestützt. Dass Cue-Validitäten prädezisional modifiziert werden, ist insofern beachtenswert, als dass die Annahme einer unidirektionalen Schlussfolgerung von Cues auf Optionen lange Zeit dominierte. Interpretiert man die systematische Umwertung von Cues als "the natural consequence of the normal mechanisms of cognitive processing" (Simon, 2004, S. 546), so ist ihr Auftreten aber nicht überraschend. Damit ist die Maximierung von Konsistenz auch nicht als kognitiver „Fehler" zu erachten. Vielmehr gilt es je nach Entscheidungssituation bzw. Zielsetzung die Adaptivität zu beurteilen. In jedem Fall gilt es in entscheidungspsychologischen Experimenten zu bedenken, dass Cue-Validitäten umgewertet werden können. Berücksichtigt man, dass Cue-Validitäten subjektiv sind und damit Entscheidungssituationen bzw. Informationslagen ganz unterschiedlich interpretiert werden können, erscheint eine Erfassung subjektiver Validitäten sinnvoll. Darüber hinaus ist zu bedenken, dass scheinbar das Wirken von Konsistenzmaximierungsprozessen durch Charakteristika der Situation beeinflusst wird, weshalb die Beachtung bidirektionaler Prozesse eventuell dazu beitragen kann, empirische Ergebnisse zu erklären. So können sich Entscheidungen und Konfidenzurteile je nach dem, ob und wie stark Umwertungen erfolgt sind, unterscheiden. Zwar sind diese zum Teil gering, doch kann die Auswirkung dann relevant sein, wenn sie im Entscheidungsverhalten oder aber in der Einschätzung der Konfidenz zum Ausdruck kommen. Letztlich ist in Untersuchungen zu fragen, ob Umwertungen möglich sind und wie relevant diese für die Ergebnisse sein könnten. Auf Basis dieser Beurteilung können dann eventuell die subjektiven Validitäten erhoben werden oder Maßnahmen ergriffen werden, um die Umwertungen zu reduzieren. Insbesondere die bisher identifizierten Moderatoren geben hierbei Anhaltspunkte. Erst wenn die Möglichkeit prädezisionaler Informationsverzerrungen stets zumindest in Betracht gezogen wird, hat sich die konstruktive Sichtweise in der Entscheidungsforschung etabliert. Eine weitere Implikation für die Entscheidungsforschung resultiert aus der Beobachtung, dass mittels bidirektionaler Prozesse, genauer gesagt mittels des PCS-Modells für probabilistische Inferenzen, nicht alle Ergebnisse zufriedenstellend erklärt werden konnten. Dies stellt die Entscheidungsforschung vor eine Herausforderung: Auf der einen Seite ist der mit diesem Modell gesetzte Anspruch, ein übergreifendes Modell des Entscheidungsverhaltens zu entwickeln, durchaus zu begrüßen. Zahlreiche empirische Belege für die Gültigkeit des

Modells stützen des Weiteren diesen Ansatz. Folglich sollte dieser Anspruch beibehalten werden. Auf der anderen Seite bedarf es aber auch der Spezifikation, zumal es sich bei dem Konsistenzmaximierungsprozess um den zentralen Mechanismus des Modells handelt. Diesen eingehend zu verstehen wird damit als besonders relevant erachtet, auch für die Überzeugungskraft des allgemeinen Modells. Folglich bedarf es gezielter Untersuchungen, unter welchen Bedingungen der Konsistenzmaximierungsprozess erfolgreich wirken kann und durch welche Faktoren er beeinflusst wird. Die vorliegenden Ergebnisse liefern hierfür erste Anhaltspunkte, weshalb im Folgenden Perspektiven für die Forschung aufgezeigt werden sollen. Es wurde die Vermutung geäußert, dass die Konsistenzmaximierung an Minimalvoraussetzungen gebunden ist, die die Generierung einer *leading option* und die damit verbundene Stützung durch systematische Umwertungen ermöglichen. Bezüglich des Einflusses der Informationslage wurde die These aufgestellt, dass es die Widersprüchlichkeit der Informationslage in Verbindung mit dem geringen Interpretationsspielraum der Informationen nicht zuließe, dass initial entweder eine *leading option* mit hoher Konfidenz vorliegt oder durch Umwertungen generiert wird. Das Zusammenspiel von Konfidenz und Umwertungen war folglich behindert. Um dies zu überprüfen, sind verschiedene Experimente denkbar. Erstens könnte man den Widerspruch, der aus der Ausgeglichenheit der Informationslage resultiert, erniedrigen, indem die Cue-Dispersion erhöht wird. Dies sollte die Konfidenz in die *leading option* erhöhen und damit auch den kognitiven Aufwand, der für eine Umwertung erforderlich ist, senken. Zweitens könnte man direkt an den Optionen ansetzen und die Wahl einer bestimmten Option fördern (vgl. Boiney, Kennedy & Nye, 1997; Simon, Snow et al., 2004), indem etwa vorher eine positive Assoziation mit dieser Option aufgebaut wurde. Drittens könnte vor dem Hintergrund der Arbeiten von Russo und Kollegen (z.B. Russo et al., 1998; Russo et al., 2000) auch eine sequentielle Darbietung der Informationen aufschlussreich sein. Das Präsentationsformat könnte sich als ein interessanter Moderator für die Leichtigkeit, mit der eine *leading option* generiert und durch weitere Umwertungen gestützt wird, herausstellen. Wichtig ist, dass hier lediglich die Informationsdarbietung verändert wird, während nicht in Frage gestellt wird, dass die Informationen parallel verarbeitet werden. Eine vierte Möglichkeit wurde bereits in Kapitel 4.2.3 angesprochen: die Erhöhung der Konsistenzmotivation. Diese könnte den mit systematischen Umwertungen verbundenen kognitiven Aufwand rechtfertigen. Hier gilt es geeignete Wege zu finden, wie dies bewerkstelligt werden kann. Es eröffnet sich dadurch die prinzipielle Möglichkeit, das Zusammenspiel von Kognition und Motivation zu untersuchen. Hinsicht-

lich des Einflusses der Wiederholung konnte gezeigt werden, dass die Reduktion der Verunsicherung zu einer systematischen Umwertung führt. Hier gilt es allerdings die Rolle der wechselnden Cue-Konstellationen zu untersuchen, da nicht auszuschließen ist, dass diese einen Beitrag zur Verunsicherung geleistet haben. Folglich sollten nicht nur identisch konzipierte Entscheidungssituationen dargeboten werden, sondern diese sollten auch stets die gleichen Cues beinhalten (vgl. Experiment 2). Des Weiteren stellt es eine interessante Forschungsfrage dar, die Veränderung der Cue-Validitäten im Verlauf zu verfolgen, etwa indem Cue-Validitäten nach jeder Entscheidungssituation erfasst werden. Allerdings sind hier andere Erhebungsmethoden von Nöten, um Erinnerungs- oder auch Ermüdungseffekte auszuschließen. Ein Rückgriff auf die Pferderennen-Metapher von Russo und Kollegen (z.B. Meloy und Russo, 2004; Russo et al., 2008; Russo et al., 1998) oder die Erfassung von Konfidenzintervallen gemäß der Vorgehensweise von Hoeffler und Ariely (1999) stellen erste Anhaltspunkte dar. Die bisherigen Ausführungen zielten direkt auf die Fragestellung bzw. Befunde dieser Arbeit ab. Eine weitere Forschungsperspektive, die auf den vorgestellten Befunden aufbaut, besteht in der Untersuchung der Frage, welche Cues eher umgewertet werden. Welche Bedeutung etwa die Valenz einer Erfahrung oder die Sicherheit bezüglich der Validitätseinschätzungen haben, ist bis heute ungeklärt. Die gezielte Verwendung bestimmter, positiv bzw. negativ eingeschätzter Cues, könnte aufschlussreich sein. Die Sicherheit in die Validitätseinschätzungen könnte in einem Lernexperiment manipuliert werden, indem die Validitäten einiger (unbekannter) Cues einfacher und die anderer schwerer zu lernen sind. Die Bearbeitung dieser Fragestellung war keines der Ziele dieser Arbeit, doch soll der Hinweis auf diese Thematik zeigen, welche weiteren Fragen aus den empirischen Befunden abzuleiten sind. Im Folgenden sollen nun unter dieser breiteren Perspektive weitere Anregungen gegeben werden, die dann letztlich ebenfalls zum Ziel dieser Arbeit, nämlich das Verständnis des Konsistenzmaximierungsprozesses zu vertiefen, beitragen können. Besonders wichtig erscheint die Frage, wie der Entscheidungsprozess weiter verläuft, wenn keine konsistente Lösung gefunden werden kann. Es wurde bereits in Kapitel 4.2.1 erörtert, dass dann mehrere Möglichkeiten denkbar sind: das Einsetzen deliberativer Prozesse, die Herabsetzung der Konsistenzschwelle, die Erhöhung von Konsistenz durch postdezisionale Prozesse oder das Vermeiden einer Entscheidung. Zunächst sollte aber ermittelt werden, ob diese Strategien überhaupt zu einer zufriedenstellenden Lösung führen können. Hinsichtlich der ersten drei Strategien ist zu klären, ob diese zu einer ausreichend konsistenten Entscheidung führen, und hinsichtlich der letzten Strategie

ist zu überprüfen, ob dies tatsächlich eine Option darstellt. Die experimentelle Umsetzung ist je nach der zu untersuchenden Möglichkeit unterschiedlich schwierig. So ist die Option, eine Entscheidung zu vermeiden, einfach zu ergänzen. Auch die Rolle postdezisionaler Prozesse kann *within subjects* untersucht werden, indem eine postdezisionale Reflexion angeboten wird. Ob bzw. wann deliberative Prozesse einsetzen, ist hingegen schwer zu ermitteln. Allerdings könnte man eine Experimentalgruppe instruieren, dass sie nach der Angabe der tendenziellen Entscheidung noch einmal über die Situation nachdenken soll. Das heißt, man würde nachträglich Deliberation induzieren, wofür es verschiedene Möglichkeiten gibt (für einen Überblick siehe Horstmann, Hausmann & Ryf, 2010). Durch den Vergleich der Validitätseinschätzungen und der Konfidenzurteile *between* oder auch *within subjects* könnte die Rolle deliberativer Prozesse ergründet werden. Auch die Untersuchung der Frage, ob die Konsistenzschwelle herabgesetzt wird, stellt eine Herausforderung dar. Denkbar ist eine Vorgehensweise gemäß der von Hausmann und Läge (2008) zur Erfassung des *Desired Level of Confidence* (DLC) vorgeschlagenen: Nachdem ermittelt wird, wie groß die Differenz zwischen zwei Optionen (gemessen an den korrigierten gewichteten Summen) sein muss, damit eine Entscheidung getroffen werden kann, können nachfolgend verschiedene Entscheidungssituationen, die diese Differenz unterschreiten, präsentiert werden. Es könnte dann die Wahl gelassen werden, ob entschieden wird oder nicht. Ersteres könnte als ein Indiz für ein Absenken der Konsistenzschwelle interpretiert werden. Welche der vier Möglichkeiten wann und eventuell von wem bevorzugt wird, ist eine empirische Frage. Gewiss werden hier auch Persönlichkeitsunterschiede zu untersuchen sein, wobei die Empfehlung von Musch und Wagner (2007), neben einer reliablen Messung der Umwertungen auch eine ausreichende Höhe der Teststärke der Verfahren zu gewährleisten, an dieser Stelle zu erwähnen ist. Des Weiteren sind situative Anpassungen denkbar, die mit *within subjects*-Designs untersucht werden können. Insgesamt gilt es auch bezüglich dieser Prozesse, die nach einer erfolglosen (automatischen) Konsistenzmaximierung einsetzen, moderierende Faktoren zu untersuchen. Natürlich trifft dies auch weiterhin auf den Prozess der Konsistenzmaximierung selbst zu, zumal – wie deutlich geworden ist – Spezifika der Entscheidungsaufgaben zu berücksichtigen sind. Dies führt letztlich zu der Frage, an welche minimalen Voraussetzungen das erfolgreiche Wirken von Konsistenzmaximierungsprozessen gebunden ist. Die Befunde dieser Arbeit haben gezeigt, dass diese bisher vernachlässigte Frage in Zukunft zu untersuchen ist.

Die vorliegenden Befunde liefern einen wichtigen Beitrag zur Etablierung des Prozesses der Konsistenzmaximierung in der Entscheidungsforschung. Die Bedeutung dieses Mechanismus kommt zum einen in dessen zentraler Stellung in einem übergeordneten Entscheidungsmodell zum Ausdruck und macht zum anderen auch seine eingehendere Untersuchung erforderlich. Einige Vorschläge wurden in diesem Kapitel gemacht, so dass das Verständnis des Konsistenzmaximierungsprozesses weiter vertieft werden kann, nachdem in dieser Arbeit ein wichtiger Schritt in diese Richtung unternommen wurde.

4.3.2 Für die Konsistenzforschung

In der vorliegenden Arbeit wurde das Streben nach Konsistenz unter einem spezifischen Blickwinkel untersucht: Im entscheidungspsychologischen Kontext wurde das Erhöhen von Konsistenz anhand der systematischen Umwertung von Cue-Validitäten im Entscheidungsprozess untersucht. Folglich erscheint es als ein zu großer Schritt, aus den vorgestellten Befunden direkte Implikationen für die Konsistenzforschung abzuleiten. Vielmehr soll die Erforschung prädezisionaler Informationsverzerrungen als Beispiel verstanden werden, anhand dessen einige generelle Trends in der Konsistenzforschung und Perspektiven für diese aufgezeigt werden können. Bezüglich der Trends ist zunächst die Etablierung von PCS-Modellen zu nennen (vgl. Kapitel 2.2.3). Damit geht eine Charakterisierung des Erhöhens von Konsistenz als natürliche Folge des Informationsverarbeitungsprozesses einher. Diese Interpretation stellt eine Weiterentwicklung im Bereich der Konsistenzforschung dar, die allerdings bereits von McGuire (1968) ähnlich formuliert wurde. In jüngster Zeit ist ein Trend zu übergreifenden Modellen zu erkennen, wie sie etwa von Monroe und Read (2008) für den Bereich der Einstellungen sowie von Read et al. (2010) für den Bereich der Persönlichkeit konzipiert wurden. Diese Modelle zeichnen sich auf der einen Seite durch ihren integrativen Charakter und auf der Seite durch die Fähigkeit, spezifische Aspekte des Forschungsbereichs zu simulieren, aus. Dieses Potential besitzt auch das dieser Arbeit zugrunde liegende PCS-Modell von Glöckner und Betsch (2008b), das es daher künftig weiter zu erforschen gilt. Als weiterer Trend ist zu beobachten, dass der grundlegende kognitive Prozess der Konsistenzmaximierung um verschiedene Aspekte erweitert wird. Hier sind insbesondere die Berücksichtigung des Lernens (Monroe und Read, 2008; Read et al., 2010), affektiver Reaktionen (Betsch, 2005a; Thagard, 2003) sowie deliberater Prozesse (Betsch, 2005a; Glöckner und Betsch, 2008b) zu nennen. Diese Aspekte gilt es in den betreffenden Modellen zu vertiefen bzw. eventuell auch in anderen

Modellen zu berücksichtigen. Darüber hinaus bieten sich weitere Perspektiven. So scheinen z.B. die Anforderungen an PCS-Modelle gestiegen zu sein. Es ist nicht mehr nur die erfolgreiche Simulation bekannter Phänomene wünschenswert, sondern es wird auch die Forderung geäußert, spezifische Vorhersagen zu generieren (z.B. Monroe und Read, 2008). Entsprechende Anstrengungen wurde u.a. von Shultz und Lepper (1996) erfolgreich unternommen. Des Weiteren gilt es neben dem Prozess der automatischen Konsistenzmaximierung als den zentralen Bestandteil von PCS-Modellen eventuell weitere Prozesse bzw. Systeme zu berücksichtigen. Monroe und Read (2008) halten etwa fest: "Finally, it cannot be ignored that there are other psychological systems that play an important part in social cognition" (S. 754). Angesichts der in dieser Arbeit ermittelten Befunde, nach denen prädezisionale Informationsverzerrungen als Ausdruck der Konsistenzmaximierung nicht in jeder Entscheidungssituation auftreten, scheint diese Aussage auch auf den Bereich der Entscheidungen übertragbar zu sein. Hier sollte künftig das Zusammenwirken mit z.B. motivationalen Prozessen ergründet werden. Interessant wäre es auch, den Zusammenhang zwischen dem Organisationsgrad eines Systems, der sich nach der in Kapitel 2.2.3 dargestellten Formel berechnen lässt, und der Anspannung einer Person zu untersuchen. Physiologische Maße, wie sie etwa von Topolinski, Likowski, Weyers und Strack (2009) eingesetzt wurden, können hier hilfreich sein. Die Autoren erfassten faziale Muskelreaktionen mittels EMG und konnten zeigen, dass semantisch kohärente Worttriaden mit spezifischen Muskelaktivationen einhergehen, die – grob gesprochen – positiven Affekt ausdrücken. Folglich ist auch eine Erweiterung des Methodenrepertoires wünschenswert. Abgesehen von grundlagenorientierten Experimenten ist auch ein Blick auf praktische Kontexte lohnenswert, wie insbesondere die folgenden Ausführungen zeigen.

4.3.3 Für praktische Kontexte

Die weite Verbreitung des Phänomens der prädezisionalen Informationsverzerrung lässt eine Erweiterung der Perspektive auf praktische Kontexte sinnvoll erscheinen. Hierbei werden bestehende empirische Befunde sowie die bisherigen Ausführungen dieses Kapitels berücksichtigt. Da praktische Settings nicht in der vorliegenden Arbeit betrachtet wurden, dient die Darstellung weniger dazu konkrete Implikationen, sondern vielmehr die prinzipielle Relevanz aufzuzeigen. Zudem ist abgesehen von der Plausiblität der Aussagen doch ihr spekulativer und damit vorläufiger Charakter stets zu bedenken.

Dass prädezisionale Informationsverzerrungen in praktischen Settings von Relevanz sind, wurde bereits vielfach diskutiert. So weisen Bond et al. (2007) u.a. auf ihre Relevanz für die Bereiche des Verkaufs und der Politik hin. Beachtlich ist auch das Ergebnis von Kostopoulou et al. (2009), nach dem Ärzte bei medizinischen Entscheidungen einer systematischen Reevaluation von Informationen unterliegen. Unabhängig von der Diagnose werden die Entscheidungen mit hoher Konfidenz getroffen. Simon (2004) erörtert ausführlich Implikationen für rechtliche Entscheidungen, u.a. die Vor- und Nachteile einer früheren Instruktion der Geschworenen und der Implementierung einer Debiasing-Strategie. Aus der (fiktiven) Personalauswahl berichtet Timmermans (1993) ein mit den hier fokussierten Reevaluationen verwandtes Phänomen. So werden die gewählten Alternativen häufiger und dabei häufiger positiv beurteilt. Larson und King (1996) erweitern die Perspektive auf größere organisationale Einheiten und machen auf die interessante Beobachtung systemischer Reevaluationen aufmerksam. Sie untermauern ihre Ausführungen mit prominenten Beispielen wie der Watergate-Affäre Anfang der 70er Jahre und der Kuba-Krise aus dem Jahr 1962, was eindrücklich das Ausmaß möglicher Folgen verdeutlicht. Auf dieses weist auch Olsen (1997) bei der Diskussion des verwandten Phänomens des *desirability bias* bei Investment-Managern hin. Zudem weist er bezüglich der Stärke der Informationsverzerrung neben der Bedeutung der Wichtigkeit eines Ereignisses auf die Bedeutung seiner Ambiguität hin. Wie in der Diskussion der empirischen Befunde deutlich geworden ist, scheint diese Variable tatsächlich eine wichtige Rolle zu spielen. In Entscheidungssituationen mit ambiguen Informationen ist die Wahrscheinlichkeit, dass systematische Reevaluationen auftreten, daher besonders hoch. Hierbei gilt es zu beachten, dass Ambiguität hier bedeutet, dass ein gewisser Interpretationsspielraum gegeben ist, so dass auch vermeintlich objektive Informationen wie Wahrscheinlichkeitsangaben je nach Kontext der Veränderung unterliegen können. Hinzu kommt, dass selbst wenn dieser Spielraum eingeschränkt ist, wie dies in den vorliegenden Experimenten der Fall war, in bestimmten Entscheidungssituationen Umwertungen auftreten. Als relevante Variable wurde hierbei die Leichtigkeit, eine *leading option* zu ermitteln und nachfolgend durch Reevaluationen zu stützen, erörtert. Es sind etwa im Falle von Entscheidungen in Unternehmen viele Situationen denkbar, in denen Bedingungen, die Verzerrungen begünstigen, vorliegen – man berücksichtige nur die Möglichkeit systemischer Reevaluationen oder das Auftreten von Reihenfolge-Effekten. Verstärkt wird das Problem dadurch, dass das Auftreten prädezisionaler im Gegensatz zu postdezisionalen Informationsverzerrungen weniger bekannt ist. Im

Kontext insbesondere der Werbe- und Konsumentenpsychologie werden letztere äußerst häufig erwähnt und sind auch unter Laien bekannt. Das kann bedeuten, dass sich Personen zwar prinzipiell der Tatsache bewusst sind, dass Informationen systematisch verzerrt werden. Sie gehen aber davon aus, dass dies nach einer Entscheidung auftritt, weshalb sie in dem irrtümlichen Glauben entscheiden, die Informationen unverzerrt bewertet zu haben. Es mag der Einwand erhoben werden, dass die Effekte derart klein sind, dass sie nicht von Relevanz für Entscheidungen in praktischen Settings sind. Dem sei entgegnet, dass kleine Effekte eine beträchtliche Wirkung haben können, wie dies insbesondere von Larson und King (1996) sowie von Olsen (1997) herausgearbeitet wurde. Dadurch erhält die Schwierigkeit, derartige Verzerrungen überhaupt zu entdecken, ein noch größeres Gewicht. Hier kann es hilfreich sein, zumindest das Bewusstsein für das Auftreten dieser Effekte zu schärfen. Doch selbst dann bleibt das Problem, diesen entgegenzuwirken, bestehen (vgl. Kapitel 4.3.3). Hinzu kommt, dass die Bedeutung zentraler situativer Größen bisher unklar ist. Während etwa Russo et al. (2000) die Verflachung von Hierarchien insofern negativ beurteilen, als dass aufgrund der reduzierten Kontrolle Verzerrungen eher unentdeckt bleiben, befürworten Larson und King (1996) diese mit der Begründung, dass Informationen weniger häufig aus Überzeugungsgründen verzerrt werden müssen. Hier eröffnen sich interessante und relevante Forschungsaufgaben.

Damit kann – die einleitende Einschränkung berücksichtigend – zusammenfassend geschlussfolgert werden, dass in praktischen Kontexten prädezisionale Informationsverzerrungen von großer Relevanz sind. Es werden wichtige und folgenreiche Entscheidungen in Situationen getroffen, die die Gefahr bergen, auf verzerrten Informationen zu basieren. Die Schwierigkeit, diese Verzerrungen zu entdecken und dann zu beheben, stellt eine Herausforderung für die Zukunft dar.

4.4 SCHLUSSBETRACHTUNG

Sie haben Ihre Wahl getroffen: Sie entscheiden sich für Wohnung A, die sich durch eine attraktive Wohnlage und eine gute Erreichbarkeit mittels öffentlicher Verkehrsmittel auszeichnet. Die Nachteile, dass die Wohnung weit von Ihrem Arbeitsplatz entfernt liegt und renovierungsbedürftig ist, nehmen Sie in Kauf. Zufällig hat ein Freund von Ihnen zeitgleich eine Wohnung gesucht und gefunden. Diese ist zwar ebenfalls renovierungsbedürftig, liegt aber in der Nähe seines Arbeitsplatzes, in einem schönen Viertel und ist gut an das

öffentliche Verkehrsnetz angebunden. Ihnen ist aufgefallen, dass Ihr Freund noch bevor er sich final für diese Wohnung entschieden hatte, die Bedeutung der positiven Attribute zunehmend hervorhob. Der Zustand der Wohnung, der ihm stets besonders wichtig war, fiel für ihn kaum noch ins Gewicht. Zudem wirkt Ihr Freund deutlich konfidenter hinsichtlich seiner Entscheidung als Sie. Hat Ihr Freund bewusst die Informationslage beschönigt? Wahrscheinlicher ist, dass bei Ihrem Freund der unbewusste und automatische Prozess der Konsistenzmaximierung, der sich in einer systematischen Umwertung der Informationen ausdrückt, wirkungsvoll war. Dadurch konnte er zudem eine vergleichsweise konfidentere Entscheidung treffen.

Dieses Beispiel und die empirischen Befunde dieser Arbeit zeigen auf der einen Seite eindrücklich, dass Informationen nicht stabil sind. Sogar vermeintlich objektive Informationen wie Wahrscheinlichkeitsangaben können der Veränderung unterliegen. Auf der anderen Seite machen das Beispiel und die vorgestellten empirischen Ergebnisse auch potenzielle Grenzen der automatischen Konsistenzmaximierung deutlich. So werden nicht in jeder Entscheidungssituation Informationen systematisch umgewertet. Dies weist auf einen Aspekt hin, den bereits Zajonc (1960) angesprochen hat: “But already today there are exceptions to consistency and balance” (S. 295). Dieser Gesichtspunkt wurde allerdings bislang vernachlässigt. Daher können bisher auch lediglich Vermutungen angestellt werden, warum Konsistenz durch systematische Reevaluationen nicht immer erhöht werden kann. Plausibel ist, dass für ein erfolgreiches Wirken dieses Prozesses gewisse Mindestvoraussetzungen vorliegen müssen, die das effektive Zusammenwirken von Konfidenz und Informationsumwertungen ermöglichen. In künftigen Arbeiten ist daher erstens gezielt zu untersuchen, warum in bestimmten Entscheidungssituationen keine erfolgreiche Konsistenzmaximierung zu beobachten ist. Zweitens ist zu ergründen, was das Scheitern der Erhöhung von Konsistenz für Konsequenzen hat. Als Möglichkeiten wurden das Einsetzen postdezisionaler oder deliberater Prozesse, das Absenken der Konsistenzschwelle sowie das Vermeiden einer Entscheidung in Betracht gezogen. Im Laufe der Konsistenzforschung haben sich PCS-Modelle als vielversprechende Ansätze erwiesen. In der Entscheidungsforschung haben Glöckner und Betsch (2008b) ein integratives Modell der automatischen und deliberaten Entscheidungsfindung vorgeschlagen, in dem der PCS-Prozess der automatischen Konsistenzmaximierung den Kernmechanismus darstellt. In zahlreichen empirischen Arbeiten konnten zentrale Annahmen dieses Modells gestützt werden. In der vorliegenden Arbeit wurde auf den Kernmechanismus der Konsistenzmaximierung fokussiert. Es zeigte

sich, dass nicht alle Ergebnisse mit dem PCS-Modell vorhergesagt oder eindeutig erklärt werden konnten. Ansatzpunkte sind zwar vorhanden, doch bedarf es der näheren Präzisierung, um gezielte Vorhersagen machen zu können. Die Erforschung der Ursachen, warum die Konsistenzmaximierung nicht immer erfolgreich ist, und der damit verbundenen Konsequenzen kann somit mit einer eventuellen Präzisierung des Modells einhergehen.

Die vorliegende Arbeit leistet einen Beitrag zur Erforschung des Strebens nach Konsistenz in der Entscheidungsforschung. Es konnten Faktoren identifiziert werden, die einen oder aber auch keinen Einfluss auf den Prozess der Konsistenzmaximierung in probabilistischen Inferenzaufgaben ausüben, und dabei Grenzen dieses Prozesses aufgezeigt werden. Die Ermittlung der Gründe für die angesprochenen Grenzen und die Identifikation weiterer Einflussfaktoren stellen spannende Herausforderungen für künftige Arbeiten dar.

LITERATURVERZEICHNIS

Abelson, R. P. (1968). Psychological Implication. In Abelson, R. P.; Aronson, E.; McGuire, W. J.; Newcomb, T. M.; Rosenberg, M. J. und Tannenbaum, P. H. (Hrsg.), *Theories of cognitive consistency: A sourcebook* (S. 112-139). Chicago: Rand McNally & Company.

Abelson, R. P.; Aronson, E.; McGuire, W. J.; Newcomb, T. M.; Rosenberg, M. J. und Tannenbaum, P. H. (Hrsg.) (1968): Theories of cognitive consistency: A sourcebook, Rand McNally and Company, Chicago.

Abelson, R. P. und Rosenberg, M. J. (1958): Symbolic psycho-logic: A model of attitudinal cognition, *Behavioral Science* (Band 3), S. 1-13.

Anderson, C. J. (2003): The psychology of doing nothing: Forms of decision avoidance result from reason and emotion, *Psychological Bulletin* (Band 129), S. 139-167.

Aronson, E. (1968). Dissonance theory: Progress and problems. In Abelson, R. P.; Aronson, E.; McGuire, W. J.; Newcomb, T. M.; Rosenberg, M. J. und Tannenbaum, P. H. (Hrsg.), *Theories of cognitive consistency: A sourcebook* (S. 5-27). Chicago: Rand McNally & Company.

Aronson, E. und Carlsmith, J. M. (1963): Effect of the severity of threat on the devaluation of forbidden behavior, *Journal of Abnormal and Social Psychology* (Band 66), S. 584-588.

Aronson, E. und Mills, J. (1959): The effect of severity of initiation on liking for a group, *Journal of Abnormal and Social Psychology* (Band 59), S. 177-181.

Asch, S. E. (1946): Forming impressions of personality, *Journal of Abnormal and Social Psychology* (Band 41), S. 258-290.

Beckmann, J. und Kuhl, J. (1984): Altering information to gain action control: Functional aspects of human information processing in decision making, *Journal of Research in Personality* (Band 18), S. 224-237.

Betsch, T. (2005a). Preference theory: An affect-based approach to recurrent decision making. In Betsch, T. und Haberstroh, S. (Hrsg.), *The routines of decision making* (S. 39-65). Mahwah, NJ: Lawrence Erlbaum Associates.

Betsch, T. (2005b): Wie beeinflussen Routinen das Entscheidungsverhalten?, *Psychologische Rundschau* (Band 56), S. 261-270.

Betsch, T.; Haberstroh, S.; Molter, B. und Glöckner, A. (2004): Oops, I did it again - relapse errors in routinized decision making, *Organizational Behavior and Human Decision Processes* (Band 93), S. 62-74.

Boiney, L. G.; Kennedy, J. und Nye, P. (1997): Instrumental bias in motivated reasoning: More when more is needed, *Organizational Behavior and Human Decision Processes* (Band 72), S. 1-24.

Bond, S. D.; Carlson, K. A.; Meloy, M. G.; Russo, J. E. und Tanner, R. J. (2007): Information distortion in the evaluation of a single option, *Organizational Behavior and Human Decision Processes* (Band 102), S. 240-254.

Brehm, J. W. (1956): Postdecision changes in the desirability of alternatives, *Journal of Abnormal and Social Psychology* (Band 52), S. 384-389.

Bröder, A. (2000): "Take The Best - Ignore The Rest" - Wann entscheiden Menschen begrenzt rational?, Pabst Science Publishers, Lengerich.

Brownstein, A. L. (2003): Biased predecision processing, *Psychological Bulletin* (Band 129), S. 545-568.

Brownstein, A. L.; Read, S. J. und Simon, D. (2004): Bias at the racetrack: Effects of individual expertise and task importance on predecision reevaluation of alternatives, *Personality and Social Psychology Bulletin* (Band 30), S. 891-904.

Brunswik, E. (1955): Representative design and probabilistic theory in a functional psychology, *Psychological Review* (Band 62), S. 193-217.

Busemeyer, J. R. und Johnson, J. G. (2004). Computational models of decision making. In Koehler, D. J. und Harvey, N. (Hrsg.), *Blackwell handbook of judgment and decision making* (S. 133-154). Malden, MA: Blackwell.

Busemeyer, J. R. und Townsend, J. T. (1993): Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment, *Psychological Review* (Band 100), S. 432-459.

Campbell, D. T. und Stanley, J. C. (1966): Experimental and quasi-experimental designs for research, Rand McNally, Chicago.

Carlson, K. A. und Pearo, L. K. (2004): Limiting predecisional distortion by prior valuation of attribute components, *Organizational Behavior and Human Decision Processes* (Band 94), S. 48-59.

Carlson, K. A. und Russo, J. E. (2001): Biased interpretation of evidence by mock jurors, *Journal of Experimental Psychology: Applied* (Band 7), S. 91-103.

Cartwright, D. und Harary, F. (1956): Structural balance: A generalization of Heider's theory, *Psychological Review* (Band 63), S. 277-293.

Christensen-Szalanski, J. J. J. und Willham, C. F. (1991): The hindsight bias: A meta-analysis, *Organizational Behavior and Human Decision Processes* (Band 48), S. 147-168.

Cialdini, R. B.; Trost, M. R. und Newsom, J. T. (1995): Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications, *Journal of Personality and Social Psychology* (Band 69), S. 318-328.

Cohen, J. (1988): Statistical power analysis for the behavioral sciences, 2. Auflage, Lawrence Erlbaum Associates, Mahwah, NJ.

Cooper, J. und Fazio, R. H. (1984). A new look at dissonance theory. In Berkowitz, L. (Hrsg.), *Advances in experimental social psychology* (S. 229-264). Orlando, FL: Academic Press.

DeKay, M. L.; Patiño-Echeverri, D. und Fischbeck, P. S. (2009): Distortion of probability and outcome information in risky decisions, *Organizational Behavior and Human Decision Processes* (Band 109), S. 79-92.

Elliot, A. J. und Devine, P. G. (1994): On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort, *Journal of Personality and Social Psychology* (Band 67), S. 382-394.

Evans, J. St. B. T.; Clibbens, J.; Cattani, A.; Harris, A. und Dennis, I. (2003): Explicit and implicit processes in multicue judgment, *Memory & Cognition* (Band 31), S. 608-618.

Faul, F.; Erdfelder, E.; Buchner, A. und Lang, A.-G. (2009): Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses, *Behavior Research Methods* (Band 41), S. 1149-1160.

Faul, F.; Erdfelder, E.; Lang, A.-G. und Buchner, A. (2007): G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences, *Behavior Research Methods* (Band 39), S. 175-191.

Fazio, R. H. und Olson, M. A. (2003): Implicit measures in social cognition research: Their meaning and use, *Annual Review of Psychology* (Band 54), S. 297-327.

Festinger, L. (1957): A theory of cognitive dissonance, Stanford University Press, Stanford, CA.

Festinger, L. (1964): Conflict, decision, & dissonance, Tavistock Publications, London.

Festinger, L. und Carlsmith, J. M. (1959): Cognitive consequences of forced compliance, *Journal of Abnormal and Social Psychology* (Band 58), S. 203-210.

Fischbacher, U. (2007): z-Tree: Zurich toolbox for ready-made economic experiments, *Experimental Economics* (Band 10), S. 171-178.

Fischhoff, B. (1975): Hindsight≠ foresight: The effect of outcome knowledge on judgment under uncertainty, *Journal of Experimental Psychology: Human Perception and Performance* (Band 1), S. 288-299.

Fischhoff, B. (1982). Debiasing. In Kahneman, D.; Slovic, P. und Tversky, A. (Hrsg.), *Judgment under uncertainty: Heuristics and biases* (S. 422-444). Cambridge, UK: Cambridge University Press.

Fremdwörterbuch, Langenscheidt (o.J.): Konsistenz, http://services.langenscheidt.de/fremdwb/fremdwb.html, 15. April 2010

Gigerenzer, G.; Hoffrage, U. und Kleinbölting, H. (1991): Probabilistic mental models: A Brunswikian theory of confidence, *Psychological Review* (Band 98), S. 506-528.

Gigerenzer, G.; Todd, P. M. und the ABC Research Group (1999): Simple heuristics that make us smart, Oxford University Press, New York.

Glöckner, A. (2006): Automatische Prozesse bei Entscheidungen, Kovac, Hamburg.

Glöckner, A. (2008a). Does intuition beat fast and frugal heuristics? A systematic empirical analysis. In Plessner, H.; Betsch, C. und Betsch, T. (Hrsg.), *Intuition in judgment and decision making* (S. 309-325). Mahwah, NJ: Lawrence Erlbaum.

Glöckner, A. (2008b). How evolution outwits bounded rationality: The efficient interaction of automatic and deliberate processes in decision making and implications for institutions. In Engel, C. und Singer, W. (Hrsg.), *Better than conscious? Decision making, the human mind, and implications for institutions* (S. 259-284). Cambridge, MA: MIT Press.

Glöckner, A. und Betsch, T. (2008a): Do people make decisions under risk based on ignorance? An empirical test of the priority heuristic against cumulative prospect theory, *Organizational Behavior and Human Decision Processes* (Band 107), S. 75-95.

Glöckner, A. und Betsch, T. (2008b): Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making, *Judgment and Decision Making* (Band 3), S. 215-228.

Glöckner, A. und Betsch, T. (2008c): Multiple-reason decision making based on automatic processing, *Journal of Experimental Psychology: Learning, Memory, and Cognition* (Band 34), S. 1055-1075.

Glöckner, A. und Betsch, T. (im Druck): Accounting for critical evidence while being precise and avoiding the strategy selection problem in a parallel constraint satisfac-

tion approach - a reply to Marewski (in press), *Journal of Behavioral Decision Making*.

Glöckner, A.; Betsch, T. und Schindler, N. (im Druck): Coherence shifts in probabilistic inference tasks, *Journal of Behavioral Decision Making*.

Glöckner, A. und Bröder, A. (2010): Processing of recognition information and additional cues: A model-based analysis of choice, confidence, and response time, *Zur Veröffentlichung eingereichtes Manuskript*.

Glöckner, A. und Herbold, A.-K. (im Druck): An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes, *Journal of Behavioral Decision Making*.

Greiner, B. (2004). An online recruitment system for economic experiments. In Kremer, K. und Macho, V. (Hrsg.), *Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63* (S. 79-93). Göttingen: Gesellschaft für Wissenschaftliche Datenverarbeitung.

Grossberg, S. (1978). A theory of visual coding, memory, and development. In Leeuwenberg, E. L. J. und Buffart, H. F. J. M. (Hrsg.), *Formal theories of visual perception*. New York: Wiley.

Guilbault, R. L.; Bryant, F. B.; Brockway, J. H. und Posavac, E. J. (2004): A meta-analysis of research on hindsight bias, *Basic and Applied Social Psychology* (Band 26), S. 103-117.

Häcker, H. O. und Stapf, K.-H. (Hrsg.) (2009): Dorsch psychologisches Wörterbuch, 15. Auflage, Verlag Hans Huber, Bern.

Harmon-Jones, E. und Harmon-Jones, C. (2007): Cognitive dissonance theory after 50 years of development, *Zeitschrift für Sozialpsychologie* (Band 38), S. 7-16.

Harvey, N. und Fischer, I. (2005). Development of experience-based judgment and decision making: The role of outcome feedback. In Betsch, T. und Haberstroh, S. (Hrsg.), *The routines of decision making* (S. 119-137). Mahwah, NJ: Lawrence Erlbaum Associates.

Hausmann, D. und Läge, D. (2008): Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition, *Judgment and Decision Making* (Band 3), S. 229-243.

Hawkins, S. A. und Hastie, R. (1990): Hindsight: Biased judgments of past events after the outcomes are known, *Psychological Bulletin* (Band 107), S. 311-327.

Heider, F. (1946): Attitudes and cognitive organization, *Journal of Psychology* (Band 21), S. 107-112.

Heider, F. (1958): The psychology of interpersonal relations, Lawrence Erlbaum Associates, Hillsdale, NJ.

Herkner, W. (1991): Lehrbuch Sozialpsychologie, Verlag Hans Huber, Bern.

Hinton, G. E.; McClelland, J. L. und Rumelhart, D. E. (1986). Distributed Representations. In Rumelhart, D. E.; McClelland, J. L. und The PDP Research Group (Hrsg.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations* (S. 77-109). Cambridge, MA: MIT Press/Bradford.

Hoeffler, S. und Ariely, D. (1999): Constructing stable preferences: A look into dimensions of experience and their impact on preference stability, *Journal of Consumer Psychology* (Band 8), S. 113-139.

Hoffrage, U.; Hertwig, R. und Gigerenzer, G. (2000): Hindsight bias: A by-product of knowledge updating?, *Journal of Experimental Psychology: Learning, Memory, and Cognition* (Band 26), S. 566-581.

Holyoak, K. J. und Simon, D. (1999): Bidirectional reasoning in decision making by constraint satisfaction, *Journal of Experimental Psychology: General* (Band 128), S. 3-31.

Holyoak, K. J. und Thagard, P. (1989): Analogical mapping by constraint satisfaction, *Cognitive Science* (Band 13), S. 295-355.

Hope, L.; Memon, A. und McGeorge, P. (2004): Unterstanding pretrial publicity: Predecisional distortion of evidence by mock jurors, *Journal of Experimental Psychology: Applied* (Band 10), S. 111-119.

Hopfield, J. J. (1982): Neural networks and physical systems with emergent collective computational abilities, *Proceedings of the National Academy of Sciences, USA* (Band 79), S. 2554-2558.

Horstmann, N.; Hausmann, D. und Ryf, S. (2010). Methods for inducing intuitive and deliberate processing modes. In Glöckner, A. und Witteman, C. L. M. (Hrsg.), *Foundations for tracing intuition: Challenges and methods* (S. 219-237). London: Psychology Press.

Janis, I. L. und Mann, L. (1977): Decision making: A psychological analysis of conflict, choice, and commitment, The Free Press, New York.

Jarvis, W. B. G. und Petty, R. E. (1996): The need to evaluate, *Journal of Personality and Social Psychology* (Band 70), S. 172-194.

Klauer, K. C. (2006). Kognitive Balance und Konsistenz. In Bierhoff, H.-W. und Frey, D. (Hrsg.), *Handbuch der Sozialpsychologie und Kommunikationspsychologie* (S. 380-387). Göttingen: Hogrefe.

Koffka, K. (1935): Principles of gestalt psychology, Routledge and Kegan Paul, London.

Köhler, W. (1947): Gestalt psychology: An introduction to new concepts in modern psychology, Liveright, New York.

Kostopoulou, O.; Mousoulis, C. und Delaney, B. (2009): Information search and information distortion in the diagnosis of an ambiguous presentation, *Judgment and Decision Making* (Band 4), S. 408-418.

Kuhl, J. (1994). Action and state orientation: Psychometric properties of the action control scales (ACS-90). In Kuhl, J. und Beckmann, J. (Hrsg.), *Volition and personality: Action versus state orientation* (S. 47-59). Göttingen: Hogrefe.

Kuhl, J. (ohne Jahresangabe): Kurzanweisung zum Fragebogen HAKEMP 90. Handlungskontrolle nach Erfolg, Misserfolg und prospektiv, Universität, Fachbereich Psychologie, Abteilung für Differentielle Psychologie und Persönlichkeitsforschung, Osnabrück.

Kunda, Z. und Thagard, P. (1996): Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory, *Psychological Review* (Band 103), S. 284-308.

Langer, E. (1994). The illusion of calculated decisions. In Schank, R. C. und Langer, E. (Hrsg.), *Beliefs, reasoning, and decision making. Psycho-Logic in honor of Bob Abelson* (S. 33-53). Hillsdale, NJ: Lawrence Erlbaum Associates.

Larson, E. W. und King, J. B. (1996): The systemic distortion of information: An ongoing challenge to management, *Organizational Dynamics* (Band 24), Nr. 3, S. 49-61.

Lord, C. G.; Lepper, M. R. und Preston, E. (1984): Considering the opposite: A corrective strategy for social judgment, *Journal of Personality and Social Psychology* (Band 47), S. 1231-1243.

Mann, L.; Janis, I. L. und Chaplin, R. (1969): Effects of anticipation of forthcoming information on predecisional processes, *Journal of Personality and Social Psychology* (Band 11), S. 10-16.

Marewski, J. (im Druck): On the theoretical precision and strategy selection problem of a single-strategy approach: A comment on Glöckner, Betsch, and Schindler (in press), *Journal of Behavioral Decision Making*.

McClelland, J. L. und Rumelhart, D. E. (1981): An interactive activation model of context effects in letter perception: Part 1. An account of basic findings, *Psychological Review* (Band 88), S. 375-407.

McClelland, J. L. und Rumelhart, D. E. (1986). Preface. In Rumelhart, D. E.; McClelland, J. L. und The PDP Research Group (Hrsg.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations* (S. IX-XIII). Cambridge, MA: MIT Press/Bradford.

McClelland, J. L.; Rumelhart, D. E. und Hinton, G. E. (1986). The appeal of parallel distributed processing. In Rumelhart, D. E.; McClelland, J. L. und The PDP Research Group (Hrsg.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations* (S. 3-45). Cambridge, MA: MIT Press/Bradford.

McClelland, J. L.; Rumelhart, D. E. und The PDP Research Group (Hrsg.) (1986): Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models, MIT Press/Bradford, Cambridge, MA.

McGuire, W. J. (1960a): Cognitive consistency and attitude change, *Journal of Abnormal and Social Psychology* (Band 60), S. 345-353.

McGuire, W. J. (1960b): Direct and indirect persuasive effects of dissonance-producing messages, *Journal of Abnormal and Social Psychology* (Band 60), S. 354-358.

McGuire, W. J. (1968). Theory of the structure of human thought. In Abelson, R. P.; Aronson, E.; McGuire, W. J.; Newcomb, T. M.; Rosenberg, M. J. und Tannenbaum, P. H. (Hrsg.), *Theories of cognitive consistency: A sourcebook* (S. 140-162). Chicago: Rand McNally & Company.

Meloy, M. G. (2000): Mood-driven distortion of product information, *Journal of Consumer Research* (Band 27), S. 345-359.

Meloy, M. G. und Russo, J. E. (2004): Binary choice under instructions to select versus reject, *Organizational Behavior and Human Decision Processes* (Band 93), S. 114-128.

Metzger, W. (1966). Figural-Wahrnehmung. In Metzger, W. (Hrsg.), *Handbuch der Psychologie. 1. Band: Der Aufbau des Erkennens* (S. 693-744). Göttingen: Hogrefe.

Mills, J. (1968). Interest in supporting and discrepant information. In Abelson, R. P.; Aronson, E.; McGuire, W. J.; Newcomb, T. M.; Rosenberg, M. J. und Tannenbaum, P. H. (Hrsg.), *Theories of cognitive consistency: A sourcebook* (S. 771-776). Chicago: Rand McNally and Company.

Mills, J. und O'Neal, E. (1971): Anticipated choice, attention, and halo effect, *Psychonomic Science* (Band 22), S. 231-233.

Monroe, B. M. und Read, S. J. (2008): A general connectionist model of attitude structure and change: The ACS (attitudes as constraint satisfaction) model, *Psychological Review* (Band 115), S. 733-759.

Montgomery, H. (1983). Decision rules and the search for a dominance structure: Towards a process model of decision making. In Humphreys, P.; Svenson, O. und Vari, A. (Hrsg.), *Analyzing and aiding decision processes* (S. 343-369). Amsterdam/Budapest: North-Holland and Hungarian Academic Press.

Montgomery, H. (1989). From cognition to action: The search for dominance in decision making. In Montgomery, H. und Svenson, O. (Hrsg.), *Process and structure in human decision making* (S. 23-49). New York: John Wiley & Sons Ltd.

Montgomery, H. und Willén, H. (1999). Decision making and action: The search for a good structure. In Juslin, P. und Montgomery, H. (Hrsg.), *Judgment and decision making: Neo-Brunswikian and process-tracing approaches* (S. 147-173). Mahwah, NJ: Lawrence Erlbaum Associates.

Musch, J. und Wagner, T. (2007): Did everybody know it all along? A review of individual differences in hindsight bias, *Social Cognition* (Band 25), S. 64-82.

Müsseler, J. und Prinz, W. (Hrsg.) (2002): Allgemeine Psychologie, Spektrum Akademischer Verlag GmbH, Heidelberg.

Nenkov, G. Y.; Morrin, M.; Ward, A.; Schwartz, B. und Hulland, J. (2008): A short form of the Maximization Scale: Factor structure, reliability and validity studies, *Judgment and Decision Making* (Band 3), S. 371-388.

Neuberg, S. L. und Newsom, J. T. (1993): Personal need for structure: Individual differences in the desire for simple structure, *Journal of Personality and Social Psychology* (Band 65), S. 113-131.

Newcomb, T. M. (1968a). Interpersonal Balance. In Abelson, R. P.; Aronson, E.; McGuire, W. J.; Newcomb, T. M.; Rosenberg, M. J. und Tannenbaum, P. H. (Hrsg.), *Theories of cognitive consistency: A sourcebook* (S. 28-51). Chicago: Rand McNally and Company.

Newcomb, T. M. (1968b). Introduction. In Abelson, R. P.; Aronson, E.; McGuire, W. J.; Newcomb, T. M.; Rosenberg, M. J. und Tannenbaum, P. H. (Hrsg.), *Theories of cognitive consistency: A sourcebook* (S. XV-XVII). Chicago Rand McNally & Company.

Newell, B. R. und Bröder, A. (2008): Cognitive processes, models and metaphors in decision research, *Judgment and Decision Making* (Band 3), S. 195-204.

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

Nisbett, R. E. und Wilson, T. D. (1977): Telling more than we can know: Verbal reports on mental processes, *Psychological Review* (Band 84), S. 231-259.

O'Neal, E. (1971): Influence of future choice importance and arousal upon the halo effect, *Journal of Personality and Social Psychology* (Band 19), S. 334-340.

Olsen, R. A. (1997): Desirability bias among professional investment managers: Some evidence from experts, *Journal of Behavioral Decision Making* (Band 10), S. 65-72.

Osgood, C. E.; Suci, G. J. und Tannenbaum, P. H. (1957): The measurement of meaning, University of Illinois Press, Urbana, IL.

Osgood, C. E. und Tannenbaum, P. H. (1955): The principle of congruity in the prediction of attitude change, *Psychological Review* (Band 62), S. 42-55.

Payne, J. W.; Bettman, J. R.; Coupey, E. und Johnson, E. J. (1992): A constructive process view of decision making: Multiple strategies in judgment and choice, *Acta Psychologica* (Band 80), S. 107-141.

Payne, J. W.; Bettman, J. R. und Johnson, E. J. (1988): Adaptive strategy selection in decision making, *Journal of Experimental Psychology: Learning, Memory, and Cognition* (Band 14), S. 534-552.

Payne, J. W.; Bettman, J. R. und Johnson, E. J. (1992): Behavioral decision research: A constructive processing perspective, *Annual Review of Psychology* (Band 43), S. 87-131.

Pennington, N. und Hastie, R. (1986): Evidence evaluation in complex decision making, *Journal of Personality and Social Psychology* (Band 51), S. 242-258.

Pennington, N. und Hastie, R. (1988): Explanation-based decision making: Effects of memory structure on judgment, *Journal of Experimental Psychology: Learning, Memory, and Cognition* (Band 14), S. 521-533.

Pennington, N. und Hastie, R. (1992): Explaining the evidence: Tests of the story model for juror decision making, *Journal of Personality and Social Psychology* (Band 62), S. 189-206.

Polman, E. (2010): Information distortion in self-other decision making, *Journal of Experimental Social Psychology* (Band 46), S. 432-435.

Rammstedt, B. und John, O. P. (2007): Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German, *Journal of Research in Personality* (Band 41), S. 203-212.

Rasch, B.; Friese, M.; Hofmann, W. und Naumann, E. (2006a): Quantitative Methoden. Band 1, 2. Auflage, Springer, Heidelberg.

Rasch, B.; Friese, M.; Hofmann, W. und Naumann, E. (2006b): Quantitative Methoden. Band 2, 2. Auflage, Springer, Heidelberg.

Read, S. J. und Marcus-Newhall, A. (1993): Explanatory coherence in social explanations: A parallel distributed processing account, *Journal of Personality and Social Psychology* (Band 65), S. 429-447.

Read, S. J. und Miller, L. C. (1993): Rapist or "regular guy": Explanatory coherence in the construction of mental models of others, *Personality and Social Psychology Bulletin* (Band 19), S. 526-540.

Read, S. J. und Miller, L. C. (1994). Dissonance and balance in belief systems: The promise of parallel constraint satisfaction processes and connectionist modeling approaches. In Schank, R. C. und Langer, E. (Hrsg.), *Beliefs, reasoning, and decision making. Psycho-logic in honor of Bob Abelson* (S. 209-235). Hillsdale, NJ: Lawrence Erlbaum Associates.

Read, S. J. und Miller, L. C. (1998a). On the dynamic construction of meaning: An interactive activation and competition model of social perception. In Read, S. J. und

Miller, L. C. (Hrsg.), *Connectionist models of social reasoning and social behavior* (S. 27-68). Mahwah, NJ: Lawrence Erlbaum Associates.

Read, S. J. und Miller, L.C. (Hrsg.) (1998b): Connectionist models of social reasoning and social behavior, Lawrence Erlbaum Associates, Mahwah, NJ.

Read, S. J.; Monroe, B. M.; Brownstein, A. L.; Yang, Y.; Chopra, G. und Miller, L.C. (2010): A neural network model of the structure and dynamics of human personality, *Psychological Review* (Band 117), S. 61-92.

Read, S. J.; Vanman, E. J. und Miller, L. C. (1997): Connectionism, parallel constraint satisfaction processes, and gestalt principles: (Re)Introducing cognitive dynamics to social psychology, *Personality and Social Psychology Review* (Band 1), S. 26-53.

Rogers, W. H. (1993): Regression standard errors in clustered samples, *Stata Technical Bulletin* (Band 13), S. 19-23.

Rosenberg, M. J. (1956): Cognitive structure and attitudinal affect, *Journal of Abnormal and Social Psychology* (Band 53), S. 367-372.

Rosenberg, M. J. (1960): A structural theory of attitude dynamics, *Public Opinion Quarterly* (Band 24), S. 319-340.

Rosenberg, M. J. (1968). Hedonism, inauthenticity, and other goals toward expansion of a consistency theory. In Abelson, R. P.; Aronson, E.; McGuire, W. J.; Newcomb, T. M.; Rosenberg, M. J. und Tannenbaum, P. H. (Hrsg.), *Theories of cognitive consistency: A sourcebook* (S. 73-111). Chicago: Rand McNally & Company.

Rumelhart, D. E.; Hinton, G. E. und McClelland, J. L. (1986). A general framework for parallel distributed processing. In Rumelhart, D. E.; McClelland, J. L. und The PDP Research Group (Hrsg.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations* (S. 45-76). Cambridge, MA: MIT Press/Bradford.

Rumelhart, D. E.; McClelland, J. L. und The PDP Research Group (Hrsg.) (1986): Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1. Foundations, MIT Press/Bradford, Cambridge, MA.

Rumelhart, D. E.; Smolensky, P.; McClelland, J. L. und Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In McClelland, J. L.; Rumelhart, D. E. und The PDP Research Group (Hrsg.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models* (S. 7-57). Cambridge, MA: MIT Press/Bradford.

Russo, J. E.; Carlson, K. A. und Meloy, M. G. (2006): Choosing an inferior alternative, *Psychological Science* (Band 17), S. 899-904.

Russo, J. E.; Carlson, K. A.; Meloy, M. G. und Yong, K. (2008): The goal of consistency as a cause of information distortion, *Journal of Experimental Psychology: General* (Band 137), S. 456-470.

Russo, J. E.; Medvec, V. H. und Meloy, M. G. (1996): The distortion of information during decisions, *Organizational Behavior and Human Decision Processes* (Band 66), S. 102-110.

Russo, J. E.; Meloy, M. G. und Medvec, V. H. (1998): Predecisional distortion of product information, *Journal of Marketing Research* (Band 35), S. 438-452.

Russo, J. E.; Meloy, M. G. und Wilks, T. J. (2000): Predecisional distortion of information by auditors and salespersons, *Management Science* (Band 46), S. 13-27.

Schlink, S. und Walther, E. (2007): Kurz und gut: Eine deutsche Kurzskala zur Erfassung des Bedürfnisses nach kognitiver Geschlossenheit, *Zeitschrift für Sozialpsychologie* (Band 38), S. 153-161.

Schmitt, N. (1996): Uses and abuses of coefficient alpha, *Psychological Assessment* (Band 8), S. 350-353.

Schwartz, B.; Ward, A.; Monterosso, J.; Lyubomirsky, S.; White, K. und Lehman, D. R. (2002): Maximizing versus satisficing: Happiness is a matter of choice, *Journal of Personality and Social Psychology* (Band 83), S. 1178-1197.

Shultz, T. R. und Lepper, M. R. (1996): Cognitive dissonance reduction as constraint satisfaction, *Psychological Review* (Band 103), S. 219-240.

Simon, D. (2004): A third view of the black box: Cognitive coherence in legal decision making, *The University of Chicago Law Review* (Band 71), S. 511-586.

Simon, D. und Holyoak, K. J. (2002): Structural dynamics of cognition: From consistency theories to constraint satisfaction., *Personality and Social Psychology Review* (Band 6), S. 283-294.

Simon, D.; Krawczyk, D. C.; Bleicher, A. und Holyoak, K. J. (2008): The transience of constructed preferences, *Journal of Behavioral Decision Making* (Band 21), S. 1-14.

Simon, D.; Krawczyk, D. C. und Holyoak, K. J. (2004): Construction of preferences by constraint satisfaction, *Psychological Science* (Band 15), S. 331-336.

Simon, D.; Pham, L. B.; Le, Q. A. und Holyoak, K. J. (2001): The emergence of coherence over the course of decision making, *Journal of Experimental Psychology: Learning, Memory, and Cognition* (Band 27), S. 1250-1260.

Simon, D.; Snow, C. J. und Read, S. J. (2004): The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction, *Journal of Personality and Social Psychology* (Band 86), S. 814-837.

Smith, E. R. (1996): What do connectionism and social psychology offer each other?, *Journal of Personality and Social Psychology* (Band 70), S. 893-912.

Solomon, R. L. (1949): An extension of control group design, *Psychological Bulletin* (Band 46), S. 137-150.

Spellman, B. A. und Holyoak, K. J. (1992): If Saddam is Hitler then who is George Bush? Analogical Mapping between systems of social roles, *Journal of Personality and Social Psychology* (Band 62), S. 913-933.

Spellman, B. A.; Ullman, J. B. und Holyoak, K. J. (1993): A coherence model of cognitive consistency: Dynamics of attitude change during the Persian Gulf War, *Journal of Social Issues* (Band 49), S. 147-165.

Stahlberg, D. und Frey, D. (1997). Konsistenztheorien. In Frey, D. und Greif, S. (Hrsg.), *Sozialpsychologie: Ein Handbuch in Schlüsselbegriffen* (S. 214-221). Weinheim: Psychologie Verlags Union.

Svenson, O. (1992): Differentiation and consolidation theory of human decision making: A frame of reference for the study of pre- and post-decision processes, *Acta Psychologica* (Band 80), S. 143-168.

Svenson, O. (1996): Decision making and the search for fundamental psychological regularities: What can be learned from a process perspective?, *Organizational Behavior and Human Decision Processes* (Band 65), S. 252-267.

Svenson, O.; Salo, I. und Lindholm, T. (2009): Post-decision consolidation and distortion of facts, *Judgment and Decision Making* (Band 4), S. 397-407.

Tannenbaum, P. H. (1968). The congruity principle: Retrospective reflections and recent research. In Abelson, R. P.; Aronson, E.; McGuire, W. J.; Newcomb, T. M.; Rosenberg, M. J. und Tannenbaum, P. H. (Hrsg.), *Theories of cognitive consistency: A sourcebook* (S. 52-72). Chicago: Rand McNally & Company.

Thagard, P. (1989): Explanatory coherence, *Behavioral and Brain Sciences* (Band 12), S. 435-502.

Thagard, P. (2000): Coherence in thought and action, MIT Press, Cambridge, MA.

Thagard, P. (2003): Why wasn't O.J. convicted? Emotional coherence in legal inference, *Cognition and Emotion* (Band 17), S. 361-383.

Thagard, P. und Kunda, Z. (1998). Making sense of people: Coherence mechanisms. In Read, S. J. und Miller, L. C. (Hrsg.), *Connectionist models of social reasoning and social behavior* (S. 3-26). Mahwah, NJ: Lawrence Erlbaum.

Thagard, P. und Millgram, E. (1995). Inference to the best plan: A coherence theory of decision. In Ram, A. und Leake, D. B. (Hrsg.), *Goal-driven learning* (S. 439-454). Cambridge, MA: MIT Press.

Thagard, P. und Verbeurgt, K. (1998): Coherence as constraint satisfaction, *Cognitive Science* (Band 22), S. 1-24.

Timmermans, D. (1993): The impact of task complexity on information use in multi-attribute decision making, *Journal of Behavioral Decision Making* (Band 6), S. 95-111.

Topolinski, S.; Likowski, K. U.; Weyers, P. und Strack, F. (2009): The face of fluency: Semantic coherence automatically elicits a specific pattern of facial muscle reactions, *Cognition and Emotion* (Band 23), S. 260-271.

von Ehrenfels, C. (1890): Ueber "Gestaltqualitäten", *Vierteljahresschrift für wissenschaftliche Philosophie* (Band 14), S. 249-292.

Walton Braver, M. C. und Braver, S. L. (1988): Statistical treatment of the Solomon four-group design: A meta-analytic approach, *Psychological Bulletin* (Band 104), S. 150-154.

Wang, H.; Johnson, T. R. und Zhang, J. (1998). UEcho: A model of uncertainty management in human abductive reasoning. In Gernsbacher, M. A. und Derry, S. J. (Hrsg.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (S. 1113-1118). Hillsdale, NJ: Erlbaum.

Webster, D. M. und Kruglanski, A. W. (1994): Individual differences in need for cognitive closure, *Journal of Personality and Social Psychology* (Band 67), S. 1049-1062.

Wertheimer, M. (1923): Untersuchungen zur Lehre von der Gestalt. II, *Psychologische Forschung* (Band 4), S. 301-350.

Wertheimer, M. (1924): Über Gestalttheorie, http://gestalttheory.net/gta/Dokumente/gestalttheorie.html, 25. Februar 2009

Zajonc, R. B. (1960): The concepts of balance, congruity, and dissonance, *Public Opinion Quarterly* (Band 24), S. 280-296.

ABBILDUNGSVERZEICHNIS

Abbildung 1: Das PCS-Modell für probabilistische Inferenzentscheidungen (aus Glöckner und Betsch, 2008b, S. 218) 55

Abbildung 2: Schematischer Überblick über die empirischen Arbeiten 82

Abbildung 3: Schema des Solomon-Viergruppenplans 90

Abbildung 4: Differenz der subjektiven Cue-Validitäten zwischen den beiden Messzeitpunkten t1 (Pretest) und t2 (Posttest) 95

Abbildung 5: Differenz zwischen den korrigierten gewichteten Summen (WADD) der Optionen zum Pretest (pre) und zum Posttest (post) 108

Abbildung 6: Differenz der subjektiven Cue-Validitäten zwischen den beiden Messzeitpunkten t1 (Pretest) und t2 (Posttest) 109

Abbildung 7: Differenz zwischen den korrigierten gewichteten Summen (WADD) der Optionen zum Pretest (pre) und zum Posttest (post) 119

Abbildung 8: Konfidenz in die Entscheidung 122

Abbildung 9: Differenz der subjektiven Cue-Validitäten zwischen den beiden Messzeitpunkten t1 (Pretest) und t2 (Posttest) 129

Abbildung 10: Skala zur Einschätzung der subjektiven Cue-Validitäten 162

Abbildung 11: Mittlere Veränderung der Differenz zwischen den korrigierten gewichteten Summen der Optionen (WADD) vom Pre- zum Posttest pro Bedingung 182

Abbildung 12: Konfidenz in die tendenzielle Entscheidung pro Fall und Bedingung 185

Abbildung 13: Konfidenz in die finale Entscheidung pro Fall und pro Bedingung 186

Abbildung 14: Schematischer Überblick über die möglichen Wirkbeziehungen 200

TABELLENVERZEICHNIS

Tabelle 1: Überblick über die veränderte Interpretation prädezisionaler Reevaluationen 76

Tabelle 2: Ergebnisse der *t*-Tests pro Bedingung 110

Tabelle 3: Korrelationen zwischen der Cue-Umwertung (absolut) und der Häufigkeit der Erfahrung 111

Tabelle 4: Interaktionseffekte der messwiederholten Varianzanalysen pro Bedingung 120

Tabelle 5: Korrelationen der Persönlichkeitsmaße mit der mittleren Cue-Umwertung 121

Tabelle 6: Beispielitems der verwendeten Persönlichkeitsmaße 126

Tabelle 7: Überblick über interne Konsistenzen, Mediane und Anzahl der Items der verwendeten Persönlichkeitsmaße 129

Tabelle 8: Korrelationen der Persönlichkeitsmaße mit der mittleren Cue-Umwertung 130

Tabelle 9: Überblick über die Art der Verbindungen zwischen Cues und Optionen 140

Tabelle 10: Korrelationen zwischen den empirischen Posttest-Validitäten und den finalen Aktivationen 145

Tabelle 11: Ergebnisse der Regressionsanalysen 147

Tabelle 12: Korrelationen zwischen den empirischen Posttest-Validitäten und den finalen Aktivationen 150

Tabelle 13: Ergebnisse der Regressionsanalysen für das erste Set an Simulationsstudien 151

Tabelle 14: Ergebnisse der Regressionsanalysen für das zweite Set an Simulationsstudien 152

Tabelle 15: Korrelation zwischen den empirischen Posttestwerten und den finalen Aktivationen 156

Tabelle 16: Ergebnisse der Regressionsanalysen 157

Tabelle 17: Die einzelnen Fakten sowie ihre Abkürzung 163

Tabelle 18: Tendenzielle und finale Entscheidungen bei wiederholter Darbietung der Entscheidungsaufgabe 167

Tabelle 19: Ergebnisse der *t*-Tests für jeden Cue bei wiederholter Darbietung der Entscheidungsaufgabe 170

Tabelle 20: Mittlere Pretest-Validitätswerte (*M*) und Standardabweichungen (*SD*) pro Bedingung 179

Tabelle 21: Tendenzielle und finale Entscheidungen in der Bedingung „reine Wiederholung“ 180

Tabelle 22: Tendenzielle Entscheidungen in der Bedingung „Information“ 180

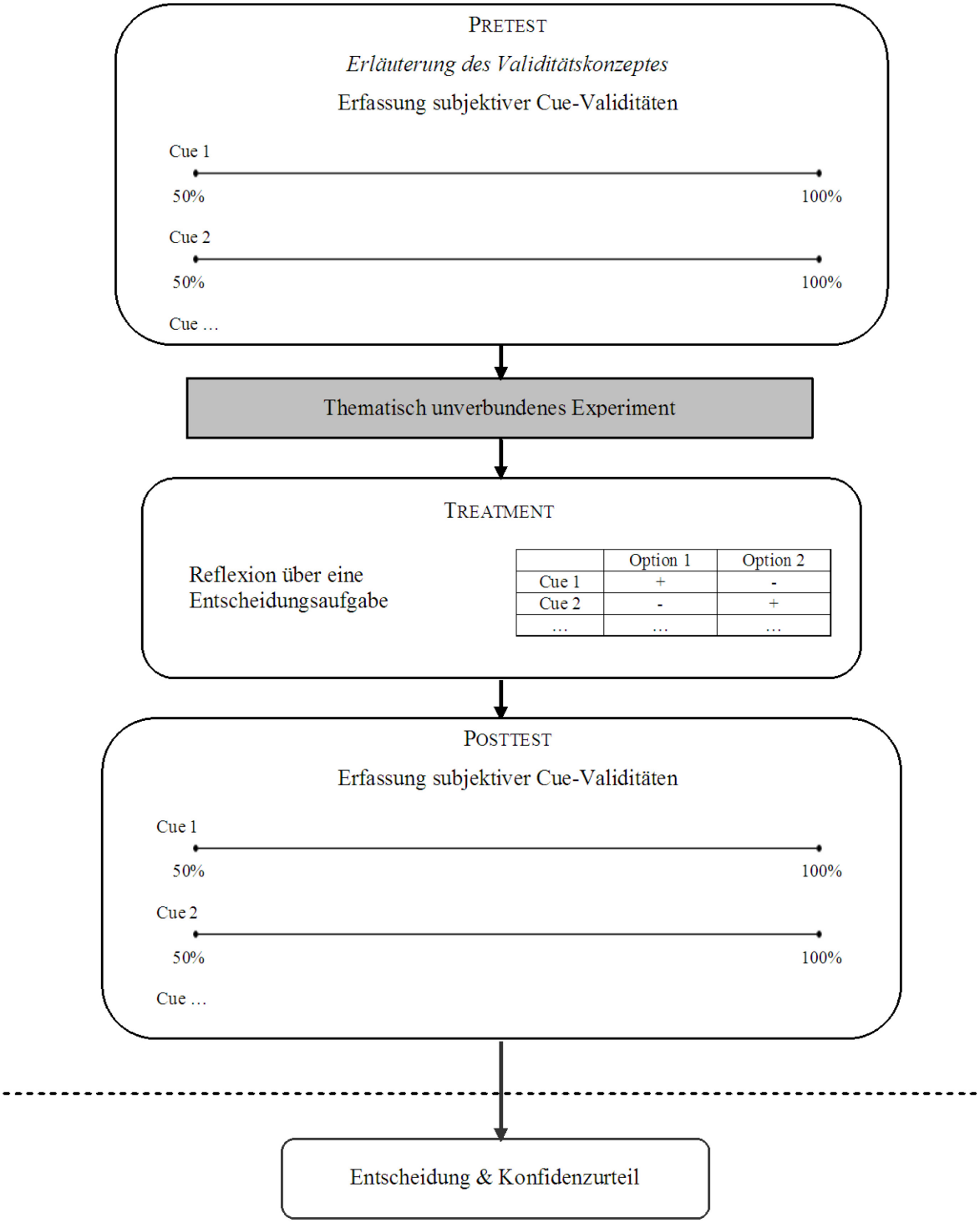
Tabelle 23: Ergebnisse der messwiederholten Varianzanalyse 181

Tabelle 24: Ergebnisse der gepaarten *t*-Tests pro Cue und Bedingung 184

ANHANG

ANHANG A: MATERIAL UND INSTRUKTIONEN

Anhang A-1: Schematische Abbildung des allgemeinen Versuchsaufbaus



Anhang A-2: Gemeinsame Instruktionen für die Experimente 1, 3 und 5

Bezüglich der Experimente 1 und 3 werden die Instruktionen nur für Bedingung 1 (Exp.1) bzw. für Bedingung 2 (Exp. 3) wiedergegeben. Die Instruktionen für die übrigen Bedingungen ergeben sich durch das Weglassen einzelner Experimentalteile (Exp. 1) bzw. durch die Modifikation der Cue-Anzahl (Exp. 3). Aus Gründen der Übersichtlichkeit werden die Experimentalteile hier wie auch im Folgenden gekennzeichnet.

Einführung:

Herzlich Willkommen!

Im Folgenden möchten wir Sie um eine genaue Einschätzung bitten.[65] So sollen Sie angeben, inwiefern ein Wetterbericht einer bestimmten Quelle das Wetter richtig vorhersagt.

Die Vorhersagekraft eines Wetterberichtes wird als dessen VALIDITÄT bezeichnet.

Die Validität eines Wetterberichtes ist definiert als die Wahrscheinlichkeit, mit der das Wetter an einem Ort, für den der Wetterbericht eine bestimmte Vorhersage macht, eher dieser Vorhersage entspricht als an einem Ort, für den der Wetterbericht diese Vorhersage nicht macht.

Ein solcher Wetterbericht könnte z.B. von einem Wahrsager stammen, der für einen Ort Regen vorhersagt oder nicht.

Da es nur zwei Arten von Vorhersagen gibt (richtig/ falsch), hat ein Wetterbericht mit einer Validität von 50% keine Vorhersagekraft.

Nun sollen Sie die Vorhersagekraft (Validität) von vier Wetterberichten einschätzen.

Bitte achten Sie hierbei auf eine möglichst gute und differenzierte Einschätzung.[66]

Pretest:

Bitte schätzen Sie die Vorhersagekraft (Validität) der folgenden Wetterberichte für das Wetter an einem Ferienort in sieben Tagen möglichst gut und differenziert ein.

Zur Erinnerung: Die Validität eines Wetterberichtes ist definiert als die Wahrscheinlichkeit, mit der das Wetter an einem Ort, für den der Wetterbericht eine bestimmte Vorhersage macht, eher dieser Vorhersage entspricht als an einem Ort, für den der Wetterbericht diese Vorhersage nicht macht.

---

[65] Exp. 4: „Im Folgenden möchten wir Sie zunächst um eine genaue Einschätzung bitten“. Exp. 5: „Im Folgenden möchten wir Sie zunächst um genaue Einschätzungen bitten.“

[66] Exp. 5: „Bitte achten Sie hierbei auf möglichst gute und differenzierte Einschätzungen.

Wetterbericht im Internet unter www.wetteronline.de

50% 100%

Wetterbericht von SAT.1

50% 100%

Wetterbericht der Bild-Zeitung

50% 100%

Wetterbericht des ZDF am Ende des heute journals

50% 100%

Treatment:

Bitte stellen Sie sich vor, Sie möchten Urlaub machen und haben zwei Ferienorte A und B in der engeren Auswahl.

Sie möchten gerne an den Ort fahren, an dem Sie eher sonniges Wetter erwarten.

Deshalb informieren Sie sich anhand verschiedener Wetterberichte, ob sie sonniges Wetter an diesen Orten vorhersagen (+) oder nicht (-).

Der Urlaub beginnt in sieben Tagen und bis heute besteht die letzte Möglichkeit, ein Sonderangebot Ihres Reisebüros wahrzunehmen. Die Sonderangebote sind für beide Ferienorte gleich gut. Jedoch liegen die Orte weit voneinander entfernt, so dass das Wetter unterschiedlich sein kann.

Ihnen liegen folgende Wetterberichte vor:

- Wetterbericht im Internet unter www.wetteronline.de
- Wetterbericht von SAT.1
- Wetterbericht der Bild-Zeitung
- Wetterbericht des ZDF am Ende des heute journals

Da Sie noch den Wetterbericht der ARD am Ende der Tagesthemen abwarten wollen, treffen Sie noch keine endgültige Entscheidung darüber, an welchem Ort Sie eher sonniges Wetter erwarten.

Sehen Sie sich nun in der folgenden Matrix an, ob die jeweiligen Wetterberichte sonniges Wetter vorhersagen (+) oder nicht (-), und überlegen Sie, an welchem Ort Sie aufgrund der vorliegenden Informationen tendenziell eher sonniges Wetter erwarten.

	Ferienort A	Ferienort B
www.wetteronline.de	-	+
Sat.1	+	-
Bild	-	+
ZDF	-	+

Posttest:

siehe Pretest

Anhang A-3: Material und Instruktionen für Experiment 2

PRETEST/POSTTEST IN EXPERIMENT 2

Einführung:

Herzlich Willkommen!

Im Folgenden möchten wir Sie zunächst um eine Einschätzung bitten.

So sollen Sie angeben, inwiefern ein bestimmtes Merkmal einer Stadt deren Größe (gemessen an der Einwohnerzahl) vorhersagt.

Die Vorhersagekraft eines Merkmals für die Größe einer Stadt wird als dessen VALIDITÄT bezeichnet.

Die Validität eines Merkmals ist definiert als die Wahrscheinlichkeit, mit der eine Stadt, die dieses Merkmal hat, größer ist als eine Stadt, die dieses Merkmal nicht hat.

Ein solches Merkmal könnte z.B. sein, dass eine Stadt einen bedeutenden Dom hat oder nicht.

Da es nur zwei Arten von Vorhersagen gibt (größer/kleiner), hat ein Merkmal mit einer Validität von 50% keine Vorhersagekraft.

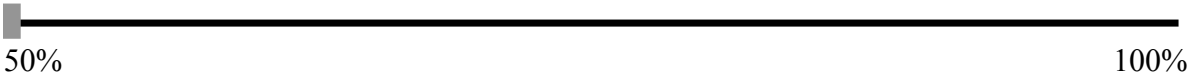
Nun sollen Sie die Vorhersagekraft (Validität) von zwei Merkmalen einschätzen.

Bitte achten Sie hierbei auf eine möglichst gute und differenzierte Einschätzung.

Pretest/Posttest:

Bitte schätzen Sie die Vorhersagekraft (Validität) der folgenden Merkmale für die Einwohnerzahl von brasilianischen Städten mit mehr als 100.000 Einwohnern möglichst gut und differenziert ein.

Validität: Die Stadt hat eine Universität.



50% 100%

Validität: Die Stadt hat einen Flughafen (dieser kann auch regional sein).



50% 100%

TREATMENT FÜR BEIDE BEDINGUNGEN VON EXPERIMENT 2

Bedingung 1:

Herzlich Willkommen!

In der folgenden Untersuchung sollen Sie sich eine bestimmte Situation vorstellen.

Bitte stellen Sie sich vor, Sie möchten zusammen mit einem Freund in den Urlaub fahren. Als Reiseziel haben Sie sich für Brasilien entschieden. Nun möchten Sie entscheiden, in welche Stadt Sie beide fahren.

Ihr Reisebüro bietet Ihnen für zwei brasilianische Städte ein identisches Sonderangebot an.

Da dieses Sonderangebot nur noch bis heute Abend gilt, möchten Sie Ihre Planung heute abschließen. Ihnen beiden ist es sehr wichtig, in die größere der beiden zur Wahl stehenden Städte zu fahren, da Sie sich in größeren Städten mehr Freizeitangebote erhoffen.

Die Angaben über die Größe der Städte sind in Ihrer Landkarte von Brasilien allerdings sehr vage, so dass Sie auf Basis anderer Informationen auf die Stadtgrößen schließen müssen.

Ihre Landkarte gibt Ihnen Informationen darüber, ob die Stadt eine Universität hat und ob die Stadt einen Flughafen, der auch regional sein kann, hat.

Im Folgenden werden die beiden Städte sowie deren Merkmale in Form einer Matrix dargestellt. Ein + steht für das Vorhandensein eines Merkmals, ein - steht für die Abwesenheit eines Merkmals.

Beispiel:

	Stadt A	Stadt B
Merkmal 1	+	-
Merkmal 2	+	-

Wenn Sie noch Fragen haben, melden Sie sich bitte beim Versuchsleiter.

Sehen Sie sich nun die folgenden brasilianischen Städte sowie deren Merkmale gut an und überlegen Sie sich, welche Stadt Sie aufgrund der vorliegenden Informationen tendenziell für größer halten würden.

Bitte denken Sie daran, noch keine Entscheidung zu treffen, da Ihr Freund und Sie später im Internet nach weiteren Informationen suchen möchten.

Präsentation einer Entscheidungsmatrix der Form (siehe Material)

	A	B
Universität	+	-
Flughafen	-	+

Nach dem Posttest:

Leider kann aufgrund technischer Störungen keine Verbindung zum Internet aufgebaut werden. Sie und Ihr Freund entscheiden sich also doch auf Basis der vorliegenden Informationen für ein Reiseziel.

Bedingung 2:

Herzlich Willkommen!

In der folgenden Untersuchung sollen Sie sich eine bestimmte Situation vorstellen.

Bitte stellen Sie sich vor, Sie möchten zusammen mit einem Freund in den Urlaub fahren. Als Reiseziel haben Sie sich für Brasilien entschieden. Startort und Zielort sind bereits festgelegt. Nun möchten Sie Ihre Reiseroute planen.

An verschiedenen Stellen der Planung stehen Sie vor der Wahl, ob Sie in die Richtung einer Stadt (Stadt A) oder in die Richtung einer anderen Stadt (Stadt B) fahren möchten. Sie möchten gerne stets in die größere der beiden Städte fahren, da Sie sich in den größeren Städten mehr Freizeitangebote erhoffen. Da ein Sonderangebot Ihres Reisebüros nur noch bis heute Abend gilt, möchten Sie Ihre Planung heute abschließen. Die Angaben über die Größe der Städte sind in Ihrer Landkarte von Brasilien allerdings sehr vage, so dass Sie auf Basis anderer Informationen auf die Stadtgrößen schließen müssen.

Ihre Landkarte gibt Ihnen Informationen darüber, ob die Stadt eine Universität hat und ob die Stadt einen Flughafen, der auch regional sein kann, hat.

Ihr Freund und Sie legen Ihre Route aber noch nicht endgültig fest, da Sie später zuhause im Internet nach weiteren Informationen suchen möchten.

Im Folgenden werden die Städtepaare sowie deren Merkmale in Form einer Matrix dargestellt. Ein + steht für das Vorhandensein eines Merkmals, ein - steht für die Abwesenheit eines Merkmals.

Beispiel:

	Stadt A	Stadt B
Merkmal 1	+	-
Merkmal 2	+	-

Wenn Sie noch Fragen haben, melden Sie sich bitte beim Versuchsleiter.

Sehen Sie sich nun die folgenden brasilianischen Städte sowie deren Merkmale nacheinander gut an und überlegen Sie sich, welche Stadt Sie aufgrund der vorliegenden Informationen tendenziell für größer halten würden.

Bitte denken Sie daran, noch keine Entscheidung zu treffen, da Ihr Freund und Sie später im Internet nach weiteren Informationen suchen möchten.

Präsentation von zwölf Entscheidungsmatrizen der Form (siehe Material)

	A	B
Universität	+	-
Flughafen	-	+

Nach dem Posttest:

Leider kann aufgrund technischer Störungen keine Verbindung zum Internet aufgebaut werden. Sie und Ihr Freund entscheiden sich also doch auf Basis der vorliegenden Informationen für ein Reiseziel.

MATERIAL FÜR EXPERIMENT 2

Überblick über das Material:

Name	UNIVERSITÄT	FLUGHAFEN	Einwohnerzahl
Maringa (PR)	+	-	325.968
Piracicaba (SP)	+	-	358.108
Anapolis (GO)	+	-	325.544
Passo Fundo (RS)	+	-	183.300
Mogi das Cruzes (SP)	+	-	362.991
Santa Maria (RS)	+	-	263.403
Sao Carlos (SP)	+	-	212.956
Santos (SP)	+	-	418.288
Juiz de Fora (MG)	+	-	513.348
Taubate (SP)	+	-	265.514
Blumenau (SC)	+	-	292.972
Marilia (SP)	+	-	218.113
Macapa (AP)	-	+	344.194
Ilheus (BA)	-	+	220.144
Varzea Grande (MT)	-	+	230.466
Corumba (MS)	-	+	96.343
Parauapebas (PA)	-	+	133.261
Foz do Iguacu (PR)	-	+	311.336
Uruguaiana (RS)	-	+	123.781
Aracaju (SE)	-	+	520.207
Palmas (TO)	-	+	179.707
Sao Luis (MA)	-	+	957.899
Juazeiro do Norte (CE)	-	+	242.139
Petrolina (PE)	-	+	268.339

Anmerkungen: Dargestellt ist, ob die Stadt eine Universität bzw. einen Flughafen hat (+) oder nicht (-). AC = Acre, AP = Amapa, AM = Amazonas, BA = Bahia, CE = Ceara, GO = Goias, MA = Maranhao, MT = Mato Grosso, MS = Mato Grosso do Sul, MG = Minas Gerais, PA = Para, PR = Parana, PE = Pernambuco, RS = Rio Grande do Sul, SC = Santa Catarina, SP = Sao Paulo, SE = Sergipe, TO = Tocantins

Quellen:
1) http://www.universidades.com.br/brasil.htm (10/2007)
2) http://www.infraero.gov.br/aero.php (10/2007)
3) http://www.ibge.gov.br (10/2007)

Entscheidungsmatrix in Bedingung 1:

	Marilia	Aracaju
Universität	+	-
Flughafen	-	+

Entscheidungsmatrizen in Bedingung 2 (in chronologischer Reihenfolge):

1

	Taubate	Aracaju
Universität	+	-
Flughafen	-	+

2

	Juazeiro de Norte	Mogi das Cruzes
Universität	-	+
Flughafen	+	-

3

	Parauabebas	Blumenau
Universität	-	+
Flughafen	+	-

4

	Santa Maria	Uruguaiana
Universität	+	-
Flughafen	-	+

5

	Marilia	Palmas
Universität	+	-
Flughafen	-	+

6

	Pircacicaba	Macapa
Universität	+	-
Flughafen	-	+

7

	Ilheus	Sao Carlos
Universität	-	+
Flughafen	+	-

8

	Corumba	Anapolis
Universität	-	+
Flughafen	+	-

9

	Foz do Iguacu	Juiz de Fora
Universität	-	+
Flughafen	+	-

10

	Sao Luis	Santos
Universität	-	+
Flughafen	+	-

11

	Maringa	Varzea Grande
Universität	+	-
Flughafen	-	+

12

	Passo Fundo	Petrolina
Universität	+	-
Flughafen	-	+

Anhang A-4: Weiteres Material für Experiment 3

ENTSCHEIDUNGSMATRIZEN

Bedingung 1:

	Ferienort A	Ferienort B
ZDF	-	+
www.wetteronline.de	+	-

Bedingung 2:

	Ferienort A	Ferienort B
www.wetteronline.de	-	+
SAT.1	+	-
Bild-Zeitung	-	+
ZDF	-	+

Bedingung 3:

	Ferienort A	Ferienort B
Süddeutsche Zeitung	+	-
ZDF	+	-
Bild-Zeitung	-	+
www.N24.de	-	+
SAT.1	-	+
www.wetteronline.de	+	-

NACHBEFRAGUNGSBOGEN FÜR EXPERIMENT 3 (EXEMPLARISCH FÜR BEDINGUNG 3)

Bitte schätzen Sie im Folgenden ein, wie häufig Sie den Wetterbericht von der Süddeutschen Zeitung, ZDF, Bild-Zeitung, www.N24.de, SAT.1 und www.wetteronline.de sehen bzw. lesen!

Wetterbericht der Süddeutschen Zeitung

○	○	○	○	○
nie	selten	manchmal	oft	sehr oft

Wetterbericht des ZDF am Ende des heute journals

○	○	○	○	○
nie	selten	manchmal	oft	sehr oft

Wetterbericht der Bild-Zeitung

○	○	○	○	○
nie	selten	manchmal	oft	sehr oft

Wetterbericht im Internet unter www.N24.de

○	○	○	○	○
nie	selten	manchmal	oft	sehr oft

Wetterbericht von SAT.1

○	○	○	○	○
nie	selten	manchmal	oft	sehr oft

Wetterbericht im Internet unter www.wetteronline.de

○	○	○	○	○
nie	selten	manchmal	oft	sehr oft

Anhang A-5: Material und Instruktionen für Experiment 4

INSTRUKTIONEN FÜR EXPERIMENT 4

Die Instruktionen werden gemeinsam mit dem entsprechenden Material für Bedingung 3 dargestellt. Die Instruktionen für die anderen beiden Bedingungen ergeben sich durch die Modifikation der Anzahl der Cues (zwei bzw. vier).

Einführung:

Herzlich Willkommen!

Im Folgenden möchten wir Sie zunächst um genaue Einschätzungen bitten.

So sollen Sie angeben, inwiefern eine bestimmte Information (im Folgenden Kennzeichen genannt) bezüglich einer Orange die Qualität der Orangensorte vorhersagt.

Die Vorhersagekraft eines Kennzeichens für die Qualität einer Orangensorte wird als dessen VALIDITÄT bezeichnet.

Die Validität eines Kennzeichens ist definiert als die Wahrscheinlichkeit, mit der eine Orangensorte, bei deren Orangen dieses Kennzeichen vorliegt, von höherer Qualität ist als eine Orangensorte, bei deren Orangen dieses Kennzeichen nicht vorliegt.

Ein solches Kennzeichen könnte z.B. sein, dass die Orangen einen Stiel haben oder nicht.

Da ein Kennzeichen entweder eine höhere Qualität oder eine geringere Qualität vorhersagt, gibt es nur zwei Arten von Vorhersagen. Folglich hat ein Kennzeichen mit einer Validität von 50% keine Vorhersagekraft.

Nun sollen Sie die Vorhersagekraft (Validität) von sechs Kennzeichen einschätzen.

Da das Vorliegen dieses Kennzeichens stets für eine höhere Qualität einer Orangensorte spricht, reicht die entsprechende Skala von 50% (keine Vorhersagekraft) bis 100% (perfekte Vorhersagekraft).

Bitte geben Sie möglichst gute und differenzierte Einschätzungen ab.

Pretest:

Bitte schätzen Sie die Vorhersagekraft (Validität) der folgenden Kennzeichen für die Qualität einer Orangensorte möglichst gut und differenziert ein.

Zur Erinnerung: Die Validität eines Kennzeichens ist definiert als die Wahrscheinlichkeit, mit der eine Orangensorte, bei deren Orangen dieses Kennzeichen vorliegt, von höherer Qualität ist als eine Orangensorte, bei deren Orangen dieses Kennzeichen nicht vorliegt.

Optimale Lagerdauer

50% 100%

Intensiver fruchtiger Geruch

50% 100%

Sehr lange Reifeperiode

50% 100%

Kräftige/ gleichmäßige Färbung

50% 100%

Optimales Lagerklima

50% 100%

Sorgfältiger Transport

50% 100%

Treatment:

Bitte stellen Sie sich vor, dass Sie heute Abend eine wichtige Dinnerparty geben und für die Zubereitung der Nachspeise Orangen benötigen.

Für diese Nachspeise möchten Sie Orangen einer qualitativ hochwertigen Sorte kaufen.

Im Supermarkt sind derzeit zwei Sorten Orangen vorrätig, welche sich im Preis nicht unterscheiden.

Die Qualität dieser beiden Orangensorten ist Ihnen nicht bekannt.

Sie erhalten aber Informationen darüber, ob bei der Orangensorte Kennzeichen, die für die Qualität einer Orangensorte sprechen, vorliegen (+) oder nicht vorliegen (-).

Sie können sich über die folgenden Kennzeichen informieren:

- Optimale Lagerdauer
- Intensiver fruchtiger Geruch
- Sehr lange Reifeperiode
- Kräftige/ gleichmäßige Färbung
- Optimales Lagerklima
- Sorgfältiger Transport

Da Sie noch die Meinung eines Supermarkt-Mitarbeiters einholen wollen, treffen Sie noch keine endgültige Entscheidung darüber, welche der beiden Sorten eine höhere Qualität aufweist.

Sehen Sie sich nun in der folgenden Matrix an, ob die jeweiligen Kennzeichen vorhanden sind (+) oder nicht (-), und überlegen Sie, von welcher Sorte Sie aufgrund der vorliegenden Informationen tendenziell eher eine höhere Qualität erwarten.

	Orangensorte A	Orangensorte B
Optimale Lagerdauer	+	-
Intensiver fruchtiger Geruch	+	-
Sehr lange Reifeperiode	-	+
Kräftige/ gleichmäßige Färbung	-	+
Optimales Lagerklima	-	+
Sorgfältiger Transport	+	-

Posttest:

Siehe Pretest

Entscheidung:

Sie finden keinen Supermarkt-Angestellten, der Ihnen Informationen über die Qualität der Orangensorten geben kann.

Da Sie aufgrund der bevorstehenden Dinnerparty unter Zeitdruck geraten, entscheiden Sie aufgrund der bisherigen Informationen, von welcher der beiden Orangensorten Sie eine höhere Qualität erwarten.

Bitte entscheiden Sie: Welche Orangensorte hat eine höhere Qualität?

	Orangensorte A	Orangensorte B
Optimale Lagerdauer	+	-
Intensiver fruchtiger Geruch	+	-
Sehr lange Reifeperiode	-	+
Kräftige/ gleichmäßige Färbung	-	+
Optimales Lagerklima	-	+
Sorgfältiger Transport	+	-

◯ Orangensorte A

◯ Orangensorte B

Wie sicher sind Sie sich bei dieser Entscheidung?

sehr unsicher sehr sicher

ENTSCHEIDUNGSMATRIZEN FÜR DIE BEDINGUNGEN VON EXPERIMENT 4

Bedingung 1:

	Orangensorte A	Orangensorte B
Optimale Lagerdauer	-	+
Sehr lange Reifeperiode	+	-

Bedingung 2:

	Orangensorte A	Orangensorte B
Kräftige/ gleichmäßige Färbung	-	+
Optimale Lagerdauer	+	-
Sehr lange Reifeperiode	-	+
Optimales Lagerklima	-	+

Bedingung 3:

	Orangensorte A	Orangensorte B
Optimale Lagerdauer	+	-
Intensiver fruchtiger Geruch	+	-
Sehr lange Reifeperiode	-	+
Kräftige/ gleichmäßige Färbung	-	+
Optimales Lagerklima	-	+
Sorgfältiger Transport	+	-

Anhang A-6: Instruktionen für die Entscheidung in Experiment 5

Sie können den Wetterbericht der ARD am Ende der Tagesthemen doch nicht mehr sehen. Da Sie nur bis heute die Möglichkeit haben, das Sonderangebot Ihres Reisebüros wahrzunehmen, entscheiden Sie nun aufgrund der bisherigen Informationen, an welchem Ferienort Sie eher sonniges Wetter erwarten.

Bitte entscheiden Sie:
An welchem Ferienort erwarten Sie eher sonniges Wetter?

	Ferienort A	Ferienort B
www.wetteronline.de	-	+
SAT.1	+	-
Bild-Zeitung	-	+
ZDF	-	+

◯ Ferienort A
◯ Ferienort B

Wie sicher sind Sie sich bei dieser Entscheidung?

sehr unsicher sehr sicher

Anhang A-7: Gemeinsames Material für Experiment 6 und 7

1. Blumengroßhandel „Blütenzauber“

Der Fall

Im Blumengroßhandel „Blütenzauber“ schließt der Assistent des Geschäftsführers die Einnahmen abends in einen Geldschrank, der sich im Nachbarraum befindet. Eines Abends möchte er das Geld um 18:45 Uhr einschließen, stellt aber fest, dass er den Schlüssel vergessen hat. Er lässt das Geld liegen und holt den Schlüssel aus seinem Büro. Als er etwa eine halbe Minute später wieder in den Raum mit dem Geldschrank geht, stellt er fest, dass die Einnahmen in Höhe von 2100€ fehlen. Im Verdacht steht Maria F., die als Verwaltungsangestellte arbeitet.

Die Fakten

Maria F. wurde gegen 19:05 Uhr von Nachbarn gesehen, als sie nach Hause kam. Man fährt mit dem Auto etwa 40 Minuten, um von dem Blumengroßhandel zum Haus von Maria F. zu gelangen.

Maria F. hat noch nie eine Straftat begangen.

Ein Kurier, der dem Geschäftsführer wichtige Post überbrachte, sah, wie gegen 18:45 Uhr eine Person aus dem Raum, in dem der Geldschrank steht, rannte. Er hat Maria F. als diese Person identifiziert.

Maria F. gilt allseits als eine Mitarbeiterin, die einen hohen Arbeitseinsatz zeigt, und wird für ihre guten Ergebnisse geschätzt.

2. Baufirma „Hövel und Partner“

Der Fall

Die leitende Büroangestellte der Baufirma „Hövel und Partner“ legt abends das Bargeld in den Safe, der in ihrem Büro steht. Dieser Safe ist mit einem Zeitmechanismus ausgestattet, der das Öffnen und Schließen des Safes dokumentiert. Als die Angestellte eines Tages wie jeden Morgen den Safe kontrolliert, bemerkt sie, dass 4700€ Bargeld fehlen. Laut Zeitmechanismus wurde der Safe zuletzt um 19:53 Uhr geöffnet. Im Laufe Ihrer Recherchen haben Sie sich zusehends auf den Angestellten Stefan K. konzentriert.

Die Fakten

Ein Techniker, der noch Geräte wartete, sah, wie eine Person gegen 19:55 Uhr aus dem Büro eilte. Er identifizierte Stefan K. als diese Person.

Um 20:15 Uhr erschien Stefan K. in seiner Stammkneipe, um gemeinsam mit Freunden ein Fußballspiel zu sehen. Man benötigt mit dem Auto etwa 40 Minuten, um von der Baufirma zu der Stammkneipe zu fahren.

Stefan K. wird bei Vorgesetzten und Kollegen für seinen Eifer und seine guten Leistungen geschätzt.

Stefan K. hat noch nie eine strafbare Tat begangen.

3. Bekleidungsgeschäft „Passform“

Der Fall

Das Bargeld wird im Bekleidungsgeschäft „Passform“ von der stellvertretenden Geschäftsführerin abends in einen Safe, der im Büro der Geschäftsführerin steht, eingeschlossen. Dieser Safe ist mit einer Zeituhr ausgestattet, wodurch nachvollzogen werden kann, wann der Safe geöffnet und geschlossen wurde. Eines Morgens bemerkt die stellvertretende Geschäftsführerin, dass 2300€ gestohlen wurden. Laut Zeituhr wurde der Safe zuletzt um 20:19 Uhr geöffnet. Im Laufe Ihrer Recherchen haben Sie sich zusehends auf die Mitarbeiterin Linda S. fokussiert. Linda S. arbeitet als Verkäuferin bei „Passform“.

Die Fakten

Ein Passant, der die Schaufenster des Bekleidungsgeschäftes betrachtete, sah, wie jemand gegen 20:20 Uhr eilig das Bekleidungsgeschäft verließ. Der Passant identifizierte Linda S. als diese Person.

Linda S. erschien um 20:30 Uhr auf einer Geburtstagsfeier, die bei einer Freundin stattfand. Mit dem Auto dauert es vom Bekleidungsgeschäft bis zur Wohnung der Freundin 30 - 35 Minuten.

Sowohl die Geschäftsführerin als auch die Kollegen von Linda S. halten diese für arbeitsam und schätzen ihre guten Arbeitsresultate.

Am Handgriff des Safes wurde ein einziger Fingerabdruck gefunden. Eine Analyse ergab, dass der Fingerabdruck nicht mit dem von Linda S. übereinstimmt.

4. Immobilienfirma „Hoch hinaus“

Der Fall

In der Immobilienfirma „Hoch hinaus“ legt die Buchhalterin abends das gesamte Bargeld in einen Safe, der im hinteren Teil ihres Büros steht. Dieser Safe ist mit einem Zeitmechanismus ausgestattet, der das Öffnen und Schließen des Safes dokumentiert. Eines Morgens stellt die Buchhalterin fest, dass aus dem Safe 6500€ entwendet wurden. Laut Zeitmechanismus wurde der Safe am Vorabend um 19:14 Uhr geöffnet. Ihre Recherchen konzentrieren sich auf Werner B. Werner B. arbeitet als Verwaltungsleiter bei „Hoch hinaus“.

Die Fakten

Die Sicherheitskamera, die der Überwachung des firmeneigenen Parkplatzes dient, filmte an dem Abend einen schwarzen Wagen, der gegen 19:20 Uhr eilig aus der Parklücke fuhr. Werner B. fährt einen roten Wagen.

Die Putzfrau der Firma berichtet, dass sie um etwa 19:15 Uhr eine Person aus dem Büro der Buchhalterin hat eilen sehen. Sie identifiziert Werner B. als diese Person.

Werner B. hat um 19:18 Uhr mit seinem Handy einen Anruf getätigt. Es konnte festgestellt werden, dass er zum Zeitpunkt des Anrufs etwa 30 Minuten von der Immobilienfirma entfernt war.

Werner B. wurde um 19:30 Uhr bei einer Theatervorstellung gesehen. Das Theater ist mit dem Auto 35 - 40 Minuten vom Tatort entfernt.

5. Maklerbüro „Kompetent“

Der Fall

Die leitende Büroangestellte des Maklerbüros „Kompetent“ ist dafür verantwortlich, das Bargeld abends in den Safe zu legen, der in ihrem Büro steht. Am Safe ist eine Zeituhr befestigt, die das Öffnen und Schließen des Safes aufzeichnet. Eines Morgens bemerkt die Büroangestellte, dass Bargeld in Höhe von 3000€ im Safe fehlt. Laut Zeituhr wurde der Safe zuletzt um 21:54 Uhr geöffnet. Im Verdacht steht Hajo S. Hajo S. arbeitet als Makler bei „Kompetent“.

Die Fakten

Hajo S. hat wenige Tage, nachdem das Geld gestohlen wurde, eine hochwertige Armbanduhr im Wert von 2500€ gekauft. Er kann belegen, dass er das Geld angespart hatte.

Die Videokamera, die den Parkplatz überwacht, filmte, wie ein roter Audi gegen 22:00 Uhr rasant den Parkplatz verließ. Hajo S. fährt einen weißen Audi.

Die Putzfrau der Firma erzählt Ihnen, sie habe gegen 21:55 Uhr eine Person vor dem Büro der Büroangestellten gesehen. Sie erkennt Hajo S. als diese Person wieder.

Gegen 22:00 Uhr traf Hajo S. zum Pokerabend bei einem Freund ein, dessen Wohnung 25 - 30 Minuten vom Maklerbüro entfernt liegt.

6. Versicherungsgesellschaft „Gut geschützt“

Der Fall

Das Bargeld der Versicherungsgesellschaft „Gut geschützt“ wird abends von der Buchhalterin in einen Safe geschlossen, der in einem benachbarten Raum steht. Eine Videokamera überwacht diesen Raum. Eines Morgens bemerkt die Buchhalterin, dass aus dem Safe 7000€ entwendet wurden. Die Videoaufzeichnungen sind aufgrund eines technischen Fehlers unscharf, aber es ist deutlich zu erkennen, dass der Safe zuletzt um 19:39 Uhr geöffnet wurde. Sonja T. ist des Diebstahls verdächtig. Sonja T. ist bei „Gut geschützt“ als Kundenberaterin tätig.

Die Fakten

Sonja T. wurde um 20:15 Uhr in ihrem Fitnessstudio gesehen, das etwa 55 Minuten von dem Tatort entfernt ist.

Der Revisor, der im Buchhaltungsbüro etwas vergessen hatte, hat gegen 19:40 Uhr eine Person aus dem Raum, in dem der Safe steht, eilen sehen. Er identifiziert Sonja T. als diese Person.

Sonja T. hat kurz nach dem Vorfall eine Möbelgarnitur für 6500€ gekauft. Sie kann belegen, dass sie das Geld durch den Verkauf von Aktien erworben hatte.

Am Eingang des Gebäudes ist eine Videokamera installiert, die den Firmenparkplatz filmt. Die Aufnahme zeigt, dass um 19:44 Uhr ein Auto der Marke Seat zügig vom Parkplatz gefahren ist. Sonja T. fährt ein Auto der Marke BMW.

7. Werbefirma „Kreativ“

Der Fall

Die leitende Büroangestellte der Werbefirma „Kreativ“ bringt abends das gesamte Bargeld zur Bank. Eines Abends lässt sie das Geld kurz auf ihrem Schreibtisch liegen, um ihre Tasche aus dem Nachbarraum zu holen. Sie kommt direkt zurück und stellt fest, dass 5900€ fehlen. Dies war um 20:45 Uhr. Im Laufe Ihrer Recherchen haben Sie sich auf Katja D. konzentriert. Katja D. arbeitet als Illustratorin bei „Kreativ“.

Die Fakten

Katja D. hat um 20:50 Uhr von ihrer Wohnung aus, die etwa 25 Minuten vom Tatort entfernt liegt, einen Anruf getätigt.

Die Überwachungskamera des Firmenparkplatzes zeigt, dass ein blaues Auto um 20:48 Uhr schnell den Parkplatz verließ. Das Auto von Katja D. ist weiß.

Um 21:00 Uhr erschien Katja D. pünktlich zum Elternabend in dem Kindergarten ihrer Kinder. Zu dieser Zeit benötigt man 35 – 40 Minuten, um vom Büro zum Kindergarten zu kommen.

Eine Kundin der Firma, die nach einem Termin das Gebäude verlassen wollte, berichtet, sie habe gegen 20:45 Uhr eine Person in unmittelbarer Nähe des Büros gesehen, die schnell Richtung Ausgang gegangen sei. Diese Kundin gibt an, sie habe Katja D. wiedererkannt.

8. Telekommunikationsunternehmen „Vernetzt“

Der Fall

Die Chefsekretärin des Telekommunikationsunternehmens „Vernetzt“ legt abends das gesamte Bargeld in einen Geldkoffer. Der Geldkoffer hat einen Sicherheitsalarmverschluss. Eines Tages ertönt um 05:41 Uhr das Alarmsignal. Es wird festgestellt, dass 5200€ fehlen. Im Laufe Ihrer Recherchen haben Sie sich auf Thomas W. als Tatverdächtigen konzentriert. Thomas W. ist als Programmierer bei „Vernetzt“ tätig.

Die Fakten

Eine Joggerin sah Thomas W. morgens gegen 6:00 Uhr, als er seinen Hund im Stadtwald ausführte. Der Stadtwald ist mit dem Auto 40 - 45 Minuten vom Tatort entfernt.

Der Wachbeamte der Firma, der wenige Minuten nach Auslösen des Alarms am Tatort war, sah dort eine Person, die weglief. Er identifiziert Thomas W. als diese Person.
Ein Hobbyfilmer nahm zufällig auf, wie ein Alfa Romeo gegen 05:45 Uhr zügig vom firmeneigenen Parkplatz fuhr. Thomas W. fährt einen Golf.
Thomas W. hat um 05:47 Uhr von seiner Wohnung aus, die ca. 35 Minuten entfernt vom Tatort liegt, einen Anruf getätigt.

9. Autohaus „Rasant“

Der Fall

Die Buchhalterin des Autohauses „Rasant“ legt abends das Bargeld in einen Geldkoffer, der mit einem Sicherheitsverschluss ausgestattet ist. Eines Morgens ertönt um 06:42 Uhr das Alarmsignal. Es fehlen 6300€ im Geldkoffer. Verdächtig ist Gertrud N., die bei „Rasant“ in der Produktentwicklung arbeitet.

Die Fakten

Gertrud N. gilt bei dem Geschäftsführer und unter den Kollegen als eine tüchtige und effiziente Mitarbeiterin.
Gertrud N. hat noch nie das Gesetz übertreten.
Gertrud N. nahm um 07:03 Uhr den Zug nach München, wo sie einen geschäftlichen Termin hatte. Um vom Autohaus zum Bahnhof zu fahren, benötigt man 40 - 45 Minuten.
Ein Angestellter der Stadtreinigung hat eine Person gesehen, die gegen 06:45 Uhr aus dem Autohaus lief. Er erkannte Gertrud N. als diese Person wieder.

10. Apotheke „Gesund bleiben“

Der Fall

In der Apotheke „Gesund bleiben“ schließt der Apotheker persönlich abends das Geld in einen Geldkoffer, der mit einem Sicherheitsalarmverschluss versehen ist. Eines Abends wird um 19:01 Uhr der Alarm ausgelöst. 5200€ fehlen im Geldkoffer. Im Verdacht steht der Mitarbeiter Fritz B.

Die Fakten

Eine Arzthelferin der Praxis, die über der Apotheke liegt, verließ um kurz nach 19:00 Uhr die Arztpraxis. Auf der Treppe sah sie eine Person, die Richtung Ausgang stürmte. Sie erkannte Fritz B. wieder.

Der einzige Fingerabdruck, der am Geldkoffer gefunden wurde, stimmt nicht mit dem von Fritz B. überein.

Um 19:30 Uhr erschien Fritz B. zu seiner Tennisstunde. Um von der Apotheke zum Tennisplatz zu fahren, benötigt man 50 - 55 Minuten.

Der Apotheker und Mitarbeiter sind sich einig, dass Fritz B. sehr eifrig ist. Seine guten Leistungen werden anerkannt.

11. Kreditbank „Sorgenfrei“

Der Fall

In der Kreditbank „Sorgenfrei“ werden das Bargeld und wichtige Dokumente in einem Safe-Raum verwahrt. Dieser Safe-Raum wird von einer Videokamera überwacht. Eines Morgens wird ein Diebstahl festgestellt: Es fehlen 8300€. Die Video-Aufzeichnungen belegen, dass der Diebstahl um 20:31 erfolgte. Verdächtigt wird Josef A., der bei „Sorgenfrei“ als Kreditberater arbeitet.

Die Fakten

Josef A. erschien um 20:40 Uhr am Kino, an dem er mit seiner Frau verabredet war. Das Kino liegt am anderen Ende der Stadt und man benötigt 30 - 35 Minuten, um von der Bank zum Kino zu kommen.

Kurz nach dem Vorfall hat Josef A. Schulden in Höhe von 7500 € bei seinem Kreditkarten-unternehmen beglichen. Josef A. kann nachweisen, dass er das Geld von seiner Schwester erhalten hat, die ihm ihrerseits Geld geschuldet hatte.

Ein Techniker, der den Kopierer reparieren wollte, hat gegen 20:35 Uhr eine Person gesehen, die eilig vom Safe-Raum weggegangen ist. Er hat Josef A. als diese Person identifiziert.

Die Tiefgarage der Kreditbank ist videoüberwacht. Es wurde aufgezeichnet, wie ein Auto der Marke BMW um 20:38 Uhr sehr schnell die Tiefgarage verlassen hat. Josef A. fährt einen Mercedes.

12. Supermarkt „Hansmann“

Der Fall

Der Betriebsleiter des Supermarktes „Hansmann“ schließt abends das Bargeld in seinem Büro in einem Schrank ein. Dieses Büro ist mit einer Alarmanlage ausgestattet. Eines Abends ertönt um 20:34 Uhr das Alarmsignal. 4500€ wurden gestohlen. Sie haben sich bei Ihren Ermittlungen auf die Mitarbeiterin Tina K. konzentriert.

Die Fakten

Tina K. gilt bei ihrem Vorgesetzten und bei ihren Kollegen als eine tüchtige Mitarbeiterin, die gute Arbeitsergebnisse erzielt.

Tina K. wurde um 21:00 Uhr bei einer Lesung in der Stadtbücherei gesehen. Um vom Supermarkt bis zur Stadtbücherei zu kommen, benötigt man etwa 45 Minuten.

Am Schrankgriff wurde ein einziger Fingerabdruck gefunden. Eine Analyse ergab, dass er nicht mit dem Fingerabdruck von Tina K. übereinstimmt.

Ein Spaziergänger, der am Supermarkt vorbeiging, sah gegen 20:35 Uhr eine Person aus dem Supermarkt rennen. Er gibt an, Tina K. wiedererkannt zu haben.

Anhang A-8: Instruktionen für Experiment 6

PRETEST/POSTTEST IN EXPERIMENT 6

Einführung:

Herzlich Willkommen!

Im Folgenden möchten wir Sie zunächst um genaue Einschätzungen bitten.

Ein Mitarbeiter eines Unternehmens ist des Diebstahls verdächtig. Sie sollen nun angeben, inwiefern ein bestimmter Fakt die Schuld bzw. Unschuld dieses Verdächtigen richtig vorhersagt.

Ein Fakt kann eher für die Schuld oder eher für die Unschuld eines Verdächtigen sprechen. Je eher ein Fakt für die Schuld (Unschuld) eines Verdächtigen spricht, desto eher spricht er gegen die Unschuld (Schuld) eines Verdächtigen.

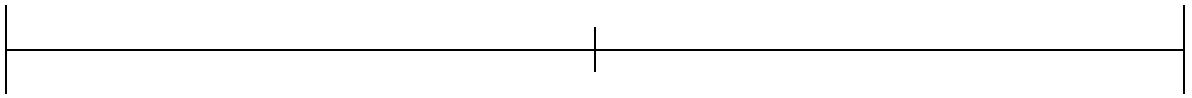
Die Vorhersagekraft eines Faktes für die Schuld bzw. Unschuld eines Verdächtigen wird als VALIDITÄT bezeichnet.

Die Validität eines Faktes ist definiert als die Wahrscheinlichkeit, dass bei Vorliegen des Faktes ein Verdächtiger eher schuldig (unschuldig) ist als unschuldig (schuldig).

Die Wahrscheinlichkeit, mit der ein Fakt für die Schuld des Verdächtigen spricht, und die Wahrscheinlichkeit, mit der ein Fakt für die Unschuld des Verdächtigen spricht, ergeben zusammen 100%, da „schuldig“ und „unschuldig“ die einzigen beiden Möglichkeiten sind.

Ein Fakt mit einer Validität von 50% hat keine Vorhersagekraft und entspricht dem Zufallsfehler. Das heißt, dass bei Vorliegen dieses Faktes 50% schuldig und 50% unschuldig sind.

Bitte benutzen Sie die folgende Skala, um einzuschätzen, inwiefern ein bestimmter Fakt die Schuld bzw. Unschuld eines Verdächtigen richtig vorhersagt. Machen Sie hierfür an der entsprechenden Stelle ein Kreuz.



Immer unschuldig	Keine Vorhersagekraft	Immer schuldig
0% schuldig	50% schuldig	100% schuldig
100% unschuldig	50% unschuldig	0% unschuldig

Bitte geben Sie möglichst gute und differenzierte Einschätzungen ab.

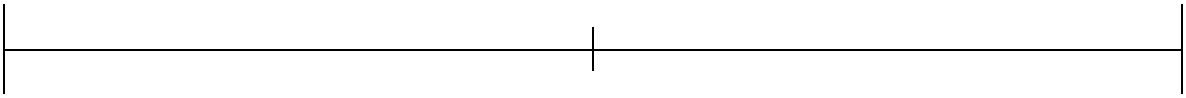
Falls Sie noch Fragen haben, wenden Sie sich bitte an den Versuchsleiter!

Pretest/Posttest:

Bitte schätzen Sie die Vorhersagekraft (Validität) der folgenden Fakten für die Schuld bzw. Unschuld eines Verdächtigen möglichst gut und differenziert ein.

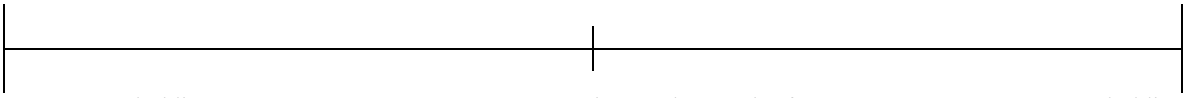
Zur Erinnerung: Die Validität eines Faktes ist definiert als die Wahrscheinlichkeit, dass bei Vorliegen des Faktes ein Verdächtiger eher schuldig (unschuldig) ist als unschuldig (schuldig).

1. Ein Augenzeuge hat den Verdächtigen als die Person identifiziert, die er vom Tatort hat wegeilen sehen.



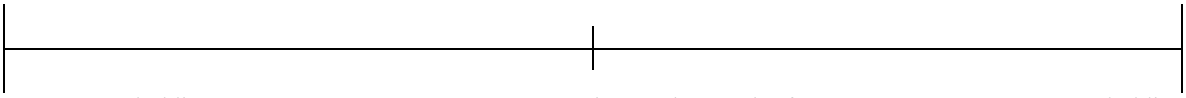
Immer unschuldig	Keine Vorhersagekraft	Immer schuldig
0% schuldig	50% schuldig	100% schuldig
100% unschuldig	50% unschuldig	0% unschuldig

2. Der Verdächtige gilt als fleißiger und guter Mitarbeiter.



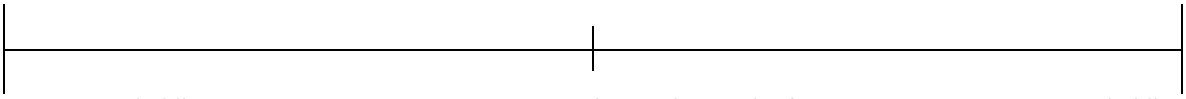
Immer unschuldig	Keine Vorhersagekraft	Immer schuldig
0% schuldig	50% schuldig	100% schuldig
100% unschuldig	50% unschuldig	0% unschuldig

3. Der Verdächtige hat finanzielle Ausgaben getätigt, die in etwa dem Betrag des gestohlenen Geldes entsprechen. Er kann aber belegen, woher er das Geld für diese Ausgaben hat.



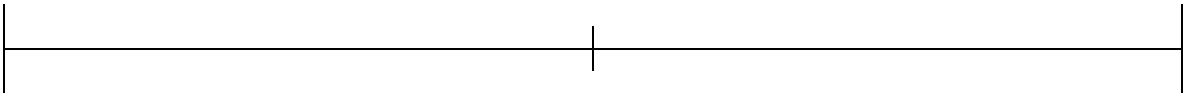
Immer unschuldig	Keine Vorhersagekraft	Immer schuldig
0% schuldig	50% schuldig	100% schuldig
100% unschuldig	50% unschuldig	0% unschuldig

4. Der Verdächtige ist noch nie mit dem Gesetz in Konflikt geraten.



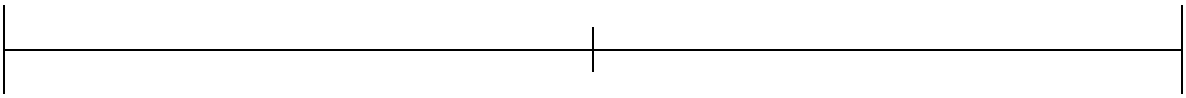
Immer unschuldig	Keine Vorhersagekraft	Immer schuldig
0% schuldig	50% schuldig	100% schuldig
100% unschuldig	50% unschuldig	0% unschuldig

5. Eine Analyse ergibt, dass der einzige Fingerabdruck, der am Tatort gefunden wurde, nicht mit dem Fingerabdruck des Verdächtigen übereinstimmt.



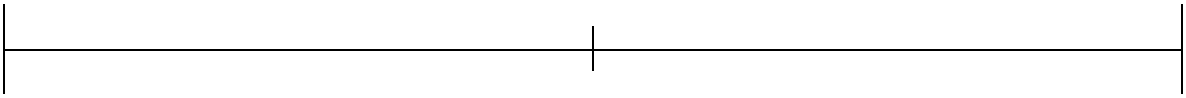
Immer unschuldig	Keine Vorhersagekraft	Immer schuldig
0% schuldig	50% schuldig	100% schuldig
100% unschuldig	50% unschuldig	0% unschuldig

6. Der Verdächtige hat etwa fünf Minuten nach der Tat einen Anruf von einem Ort aus gemacht, der ca. 30 Minuten weit entfernt vom Tatort ist.



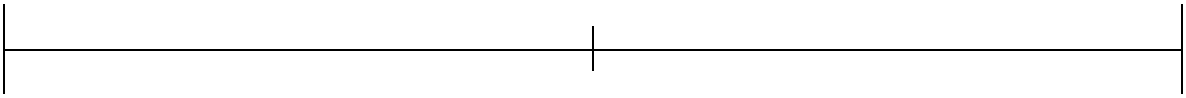
Immer unschuldig	Keine Vorhersagekraft	Immer schuldig
0% schuldig	50% schuldig	100% schuldig
100% unschuldig	50% unschuldig	0% unschuldig

7. Wenige Minuten nach dem Diebstahl wurde von der Überwachungskamera gefilmt, wie ein Auto rasant den Firmen-Parkplatz verlässt. Der Verdächtige fährt nicht das gleiche Auto, das von der Überwachungskamera gefilmt wurde.



Immer unschuldig	Keine Vorhersagekraft	Immer schuldig
0% schuldig	50% schuldig	100% schuldig
100% unschuldig	50% unschuldig	0% unschuldig

8. Der Verdächtige wurde etwa 20 Minuten nach der Tat an einem Ort gesehen, der etwa 40 Minuten vom Tatort entfernt liegt.



Immer unschuldig	Keine Vorhersagekraft	Immer schuldig
0% schuldig	50% schuldig	100% schuldig
100% unschuldig	50% unschuldig	0% unschuldig

TREATMENT FÜR BEIDE BEDINGUNGEN VON EXPERIMENT 6

Es werden die Instruktionen für die Variante, in der eine tendenzielle Entscheidung vor dem Posttest abzugeben war, vorgestellt.

Bedingung 1:

Bitte stellen Sie sich vor, dass Sie als Detektiv bei der Wirtschaftsdetektei „Diskreta“ arbeiten. Diese Detektei deckt Delikte von Wirtschaftskriminalität auf, wobei Sie selbst sich auf Fälle von Diebstahl spezialisiert haben.

Ihnen liegt der Auftrag eines Unternehmens vor, in dem es zu einem Diebstahl gekommen ist. Sie wurden damit beauftragt, den Täter zu finden. In den vergangenen Wochen und Monaten haben Sie in diesem Unternehmen recherchiert. Mit der Zeit haben sich Ihre Ermittlungen auf einen Mitarbeiter konzentriert. Ihnen liegen nun verschiedene Fakten vor, die für oder gegen die Schuld des Mitarbeiters sprechen.

Da Sie eine DNA-Analyse in Auftrag gegeben haben und das Ergebnis abwarten möchten, treffen Sie noch keine endgültige Entscheidung darüber, ob Sie den betreffenden Mitarbeiter für den Täter halten oder nicht.

Nach einer kurzen Fallschilderung werden Ihnen die Fakten, die Sie recherchiert haben, vorgelegt.

Lesen Sie sich aufmerksam durch, welche Fakten für oder gegen die Schuld des Mitarbeiters sprechen, und überlegen Sie, ob Sie den Mitarbeiter aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten.

Präsentation des Falls „Immobilienfirma ‚Hoch hinaus’“.

Bitte überlegen Sie nun, ob Sie Werner B. aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten. Nehmen Sie sich dafür so viel Zeit wie nötig.

Bitte treffen Sie noch keine endgültige Entscheidung!

Bitte kreuzen Sie nun an, ob Sie Werner B. aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten:

☐ schuldig ☐ unschuldig

Bedingung 2:

Bitte stellen Sie sich vor, dass Sie als Detektiv bei der Wirtschaftsdetektei „Diskreta“ arbeiten. Diese Detektei deckt Delikte von Wirtschaftskriminalität auf, wobei Sie selbst sich auf Fälle von Diebstahl spezialisiert haben.

Ihnen liegen die Aufträge verschiedener Unternehmen vor, in denen es zu Diebstählen gekommen ist.

Sie wurden damit beauftragt, den Täter zu finden. In den vergangenen Wochen und Monaten haben Sie in diesen verschiedenen Unternehmen recherchiert. Mit der Zeit haben sich in jedem Unternehmen Ihre Ermittlungen auf einen Mitarbeiter konzentriert. Ihnen liegen nun verschiedene Fakten vor, die für oder gegen die Schuld des Mitarbeiters sprechen.

Da Sie für jeden Fall von Diebstahl eine DNA-Analyse in Auftrag gegeben haben und das Ergebnis abwarten möchten, treffen Sie noch keine endgültige Entscheidung darüber, ob Sie den jeweiligen Mitarbeiter für den Täter halten oder nicht.

Nach einer kurzen Fallschilderung werden Ihnen die Fakten, die Sie recherchiert haben, vorgelegt.

Lesen Sie sich aufmerksam durch, welche Fakten in den einzelnen Fällen für oder gegen die Schuld des Mitarbeiters sprechen, und überlegen Sie, ob Sie den Mitarbeiter aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten.

Präsentation der zwölf Fälle in der Reihenfolge wie in Anhang A-7 dargestellt.

Nach jedem Fall:

Bitte überlegen Sie nun, ob Sie [Name Beschuldigter] aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten. Nehmen Sie sich dafür so viel Zeit wie nötig.

Bitte treffen Sie noch keine endgültige Entscheidung!

Bitte kreuzen Sie nun an, ob Sie [Name Beschuldigter] aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten:

☐ schuldig ☐ unschuldig

Instruktionen für die Entscheidung für beide Bedingungen von Experiment 6

Es werden die Instruktionen für die Variante, in der eine tendenzielle Entscheidung vor dem Posttest abzugeben war, vorgestellt.

Bedingung 1:

Die DNA-Analyse, die Sie in Auftrag gegeben haben, benötigt noch etwas Zeit.
Ihr Auftraggeber möchte aber bereits von Ihnen wissen, wie Sie sich jetzt entscheiden würden, wenn Sie müssten (d.h. ohne das Ergebnis der DNA-Analyse abzuwarten).
Um diese Angabe zu machen, werden Ihnen noch einmal die Fallschilderung sowie die Fakten, die Sie recherchiert haben, vorgelegt.
Entscheiden Sie sich, ob Sie den Mitarbeiter aufgrund der vorliegenden Fakten für schuldig oder für unschuldig halten, und kreuzen Sie das entsprechende Kästchen an.
Erneute Präsentation des Falls.
Treffen Sie jetzt eine Entscheidung: Halten Sie Werner B. aufgrund der vorliegenden Fakten für schuldig oder unschuldig?

☐ schuldig ☐ unschuldig

Bedingung 2:

Die DNA-Analysen, die Sie in Auftrag gegeben haben, benötigen noch etwas Zeit.
Ihre Auftraggeber möchten aber bereits von Ihnen wissen, wie Sie sich jetzt entscheiden würden, wenn Sie müssten (d.h. ohne die Ergebnisse der DNA-Analysen abzuwarten).
Um diese Angaben zu machen, werden Ihnen noch einmal die Fallschilderungen sowie die Fakten, die Sie recherchiert haben, vorgelegt.
Entscheiden Sie sich, ob Sie den Mitarbeiter aufgrund der vorliegenden Fakten für schuldig oder für unschuldig halten, und kreuzen Sie das entsprechende Kästchen an.
Erneute Präsentation der Fälle.
Nach jedem Fall:
Treffen Sie jetzt eine Entscheidung: Halten Sie [Name Beschuldigter] aufgrund der vorliegenden Fakten für schuldig oder unschuldig?

☐ schuldig ☐ unschuldig

Anhang A-9: Instruktionen für Experiment 7

PRETEST/POSTTEST IN EXPERIMENT 7

Einführung:

Herzlich Willkommen!

Im Folgenden möchten wir Sie zunächst um genaue Einschätzungen bitten.

Ein Mitarbeiter eines Unternehmens ist des Diebstahls verdächtig. Sie sollen nun angeben, inwiefern ein bestimmter Fakt die Schuld bzw. Unschuld dieses Verdächtigen richtig vorhersagt.

Ein Fakt kann eher für die Schuld oder eher für die Unschuld eines Verdächtigen sprechen.

Die Vorhersagekraft eines Faktes für die Schuld bzw. Unschuld eines Verdächtigen wird als VALIDITÄT bezeichnet.

Die Validität eines Faktes für die Schuld eines Verdächtigen ist definiert als die Wahrscheinlichkeit, dass bei Vorliegen des Faktes ein Verdächtiger eher schuldig ist als unschuldig.

Die Validität eines Faktes für die Unschuld eines Verdächtigen ist definiert als die Wahrscheinlichkeit, dass bei Vorliegen des Faktes ein Verdächtiger eher unschuldig ist als schuldig.

Ein Fakt mit einer Validität von 50% hat keine Vorhersagekraft und entspricht dem Zufallsfehler. Das heißt, dass bei Vorliegen dieses Faktes 50% schuldig bzw. 50% unschuldig sind.

Bitte benutzen Sie die folgenden Skalen, um einzuschätzen, inwiefern ein bestimmter Fakt die Schuld bzw. Unschuld eines Verdächtigen richtig vorhersagt. Bitte schieben Sie den Schieberegler an die entsprechende Stelle.

Der Fakt spricht für die SCHULD des Verdächtigen:



Keine Vorhersagekraft
50% schuldig

Immer schuldig
100% schuldig

Der Fakt spricht für die UNSCHULD des Verdächtigen:



Keine Vorhersagekraft
50% unschuldig

Immer unschuldig
100% unschuldig

Bitte geben Sie möglichst gute und differenzierte Einschätzungen ab.

Falls Sie noch Fragen haben, wenden Sie sich bitte an den Versuchsleiter!

Pretest/Posttest:

Bitte schätzen Sie die Vorhersagekraft (Validität) der folgenden Fakten für die Schuld bzw. Unschuld eines Verdächtigen möglichst gut und differenziert ein.

Zur Erinnerung: Die Validität eines Faktes ist definiert als die Wahrscheinlichkeit, dass bei Vorliegen des Faktes ein Verdächtiger eher schuldig (unschuldig) ist als unschuldig (schuldig).

Der folgende Fakt spricht für die SCHULD des Verdächtigen:

1. Ein Augenzeuge hat den Verdächtigen als die Person identifiziert, die er vom Tatort hat wegeilen sehen.

Keine Vorhersagekraft
50% schuldig

Immer schuldig
100% schuldig

Die folgenden Fakten sprechen für die UNSCHULD des Verdächtigen:

1. Der Verdächtige gilt als fleißiger und guter Mitarbeiter.

Keine Vorhersagekraft
50% unschuldig

Immer unschuldig
100% unschuldig

2. Der Verdächtige hat finanzielle Ausgaben getätigt, die in etwa dem Betrag des gestohlenen Geldes entsprechen. Er kann aber belegen, woher er das Geld für diese Ausgaben hat.

Keine Vorhersagekraft
50% unschuldig

Immer unschuldig
100% unschuldig

3. Der Verdächtige ist noch nie mit dem Gesetz in Konflikt geraten.

Keine Vorhersagekraft
50% unschuldig

Immer unschuldig
100% unschuldig

4. Eine Analyse ergibt, dass der einzige Fingerabdruck, der am Tatort gefunden wurde, nicht mit dem Fingerabdruck des Verdächtigen übereinstimmt.

Keine Vorhersagekraft
50% unschuldig

Immer unschuldig
100% unschuldig

5. Der Verdächtige hat etwa fünf Minuten nach der Tat einen Anruf von einem Ort aus gemacht, der ca. 30 Minuten weit entfernt vom Tatort ist.



Keine Vorhersagekraft
50% unschuldig

Immer unschuldig
100% unschuldig

6. Wenige Minuten nach dem Diebstahl wurde von der Überwachungskamera gefilmt, wie ein Auto rasant den Firmen-Parkplatz verlässt. Der Verdächtige fährt nicht das gleiche Auto, das von der Überwachungskamera gefilmt wurde.



Keine Vorhersagekraft
50% unschuldig

Immer unschuldig
100% unschuldig

7. Der Verdächtige wurde etwa 20 Minuten nach der Tat an einem Ort gesehen, der etwa 40 Minuten vom Tatort entfernt liegt.



Keine Vorhersagekraft
50% unschuldig

Immer unschuldig
100% unschuldig

TREATMENT FÜR BEIDE BEDINGUNGEN VON EXPERIMENT 7

Bedingung 1:

Bitte stellen Sie sich vor, dass Sie als Detektiv bei der Wirtschaftsdetektei „Diskreta“ arbeiten. Diese Detektei deckt Delikte von Wirtschaftskriminalität auf, wobei Sie selbst sich auf Fälle von Diebstahl spezialisiert haben.

Ihnen liegen die Aufträge verschiedener Unternehmen vor, in denen es zu Diebstählen gekommen ist.

Sie wurden damit beauftragt, den Täter zu finden. In den vergangenen Wochen und Monaten haben Sie in diesen verschiedenen Unternehmen recherchiert. Mit der Zeit haben sich in jedem Unternehmen Ihre Ermittlungen auf einen Mitarbeiter konzentriert.

Ihnen liegen nun verschiedene Fakten vor, die für oder gegen die Schuld des Mitarbeiters sprechen.

Da Sie für jeden Fall von Diebstahl noch die Ergebnisse einer DNA-Analyse abwarten wollen, treffen Sie noch keine endgültige Entscheidung darüber, ob Sie den jeweiligen Mitarbeiter für den Täter halten oder nicht.

Nach einer kurzen Fallschilderung werden Ihnen die Fakten, die Sie recherchiert haben, vorgelegt. Lesen Sie sich sorgfältig durch, welche Fakten in den einzelnen Fällen für oder gegen die Schuld des Mitarbeiters sprechen, und versuchen Sie die Informationslage zu erfassen.

Überlegen Sie dann, ob Sie den Mitarbeiter aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten.

Bitte geben Sie außerdem auf der entsprechenden Skala an, wie sicher Sie sich bei dieser vorläufigen Tendenz sind.

Es ist sehr wichtig, dass Sie sich die Informationen aufmerksam durchlesen! Nehmen Sie sich hierfür so viel Zeit wie nötig.

Wenn Sie noch Fragen haben, dann wenden Sie sich bitte an den Versuchsleiter!

Präsentation der zwölf Fälle in der Reihenfolge wie in Anhang A-6 dargestellt.

Nach jedem Fall:

Bitte überlegen Sie nun, ob Sie [Name Beschuldigter] aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten. Nehmen Sie sich dafür so viel Zeit wie nötig.

Bitte treffen Sie noch keine endgültige Entscheidung!

Bitte geben Sie an, ob Sie [Name Beschuldigter] aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten:

◯ schuldig ◯ unschuldig

Wie sicher sind Sie sich?

Sehr unsicher sehr sicher

Bedingung 2:

Bitte stellen Sie sich vor, dass Sie als Detektiv bei der Wirtschaftsdetektei „Diskreta“ arbeiten. Diese Detektei deckt Delikte von Wirtschaftskriminalität auf, wobei Sie selbst sich auf Fälle von Diebstahl spezialisiert haben.

Es liegen die Aufträge verschiedener Unternehmen vor, in denen es zu Diebstählen gekommen ist.

Sie wurden damit beauftragt, den Täter zu finden. In den vergangenen Wochen und Monaten haben Sie in diesen verschiedenen Unternehmen recherchiert. Mit der Zeit haben sich in jedem Unternehmen Ihre Ermittlungen auf einen Mitarbeiter konzentriert.

Ihnen liegen verschiedene Fakten vor, die für oder gegen die Schuld des Mitarbeiters sprechen.

Da Sie für jeden Fall von Diebstahl noch eine weitere wichtige Information erwarten, treffen Sie noch keine endgültige Entscheidung darüber, ob Sie den jeweiligen Mitarbeiter für den Täter halten oder nicht.

Nach einer kurzen Fallschilderung werden Ihnen die Fakten, die Sie recherchiert haben, vorgelegt. Lesen Sie sich sorgfältig durch, welche Fakten in den einzelnen Fällen für oder gegen die Schuld des Mitarbeiters sprechen, und versuchen Sie die Informationslage zu erfassen.

Überlegen Sie dann, ob Sie den Mitarbeiter aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten.

Bitte geben Sie außerdem auf der entsprechenden Skala an, wie sicher Sie sich bei dieser vorläufigen Tendenz sind.

Daraufhin wird Ihnen eine weitere wichtige Information gegeben, die Ihnen Aufschluss über die Schuld oder Unschuld des Verdächtigen gibt.

Es ist sehr wichtig, dass Sie sich die Informationen aufmerksam durchlesen! Nehmen Sie sich hierfür so viel Zeit wie nötig.

Wenn Sie noch Fragen haben, dann wenden Sie sich bitte an den Versuchsleiter!

Präsentation der zwölf Fälle in der Reihenfolge wie in Anhang A-6 dargestellt.

Nach jedem Fall:

Bitte überlegen Sie nun, ob Sie [Name Beschuldigter] aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten. Nehmen Sie sich dafür so viel Zeit wie nötig.

Bitte treffen Sie noch keine endgültige Entscheidung!

Bitte geben Sie an, ob Sie [Name Beschuldigter] aufgrund der vorliegenden Fakten tendenziell eher für schuldig oder tendenziell eher für unschuldig halten:

◯ schuldig ◯ unschuldig

Wie sicher sind Sie sich?



Sehr unsicher sehr sicher

Sie erhalten nun eine weitere wichtige Information, die Ihnen Aufschluss über die Schuld oder Unschuld von [Name Beschuldigter] gibt!

Nach den Fällen 1, 5, 7, 10:

[Name Beschuldigter] hat ein Alibi. Dieses Alibi wurde von Ihren Mitarbeitern überprüft und ist vollkommen wasserdicht.

Demnach ist [Name Beschuldigter] unschuldig!

Nach den Fällen 2, 6, 9, 11:

Ein anderer Mitarbeiter hat ein umfassendes Geständnis abgegeben, dessen Wahrheitsgehalt überprüft wurde und das vollkommen glaubwürdig ist.

Demnach ist [Name Beschuldigter] unschuldig!

Nach den Fällen 3, 4, 8,12:

Es ist ein Überwachungsvideo aufgetaucht, auf dem eindeutig zu sehen ist, wie eine andere Person als [Name Beschuldigter] das Geld aus dem Safe nimmt. Es konnte nachgewiesen werden, dass [Name Beschuldigter] nichts mit dem tatsächlichen Täter zu tun hat.

Demnach ist [Name Beschuldigter] unschuldig!

INSTRUKTION DER ENTSCHEIDUNG FÜR BEIDE BEDINGUNGEN VON EXPERIMENT 7

Bedingung 1:

Ihr Auftraggeber möchte nun von Ihnen wissen, ob Sie den jeweiligen Mitarbeiter für schuldig oder für unschuldig halten.

Die DNA-Analysen, die Sie in Auftrag gegeben haben, sind im Labor verloren worden.

Sie müssen sich also ohne die Ergebnisse der DNA-Analysen endgültig entscheiden, ob Sie den Mitarbeiter aufgrund der vorliegenden Fakten für schuldig oder für unschuldig halten!

Um diese Angaben zu machen, werden Ihnen noch einmal die Fallschilderungen sowie die Fakten, die Sie recherchiert haben, vorgelegt.

Entscheiden Sie sich bitte, ob Sie den Mitarbeiter aufgrund der vorliegenden Fakten für schuldig oder für unschuldig halten!

Bitte geben Sie außerdem auf der entsprechenden Skala an, wie sicher Sie sich bei Ihrer Entscheidung sind.

Präsentation der zwölf Fälle in der Reihenfolge wie in Anhang A-6 dargestellt.

Nach jedem Fall:

Treffen Sie jetzt eine Entscheidung: Halten Sie [Name Beschuldigter] aufgrund der vorliegenden Fakten für schuldig oder unschuldig?

◯ schuldig ◯ unschuldig

Wie sicher sind Sie sich?



Sehr unsicher sehr sicher

Bedingung 2:

Sie haben erfahren, dass sämtliche Mitarbeiter, die des Diebstahls verdächtigt wurden, unschuldig sind.

Nun sollen Sie angeben, wie sicher Sie sich sind, dass diese Mitarbeiter unschuldig sind.

Um diese Angabe zu machen, werden Ihnen noch einmal die Fallschilderungen sowie die Fakten, die Sie recherchiert haben, vorgelegt.

Präsentation der Fälle sowie der zusätzlichen Information.

Nach jedem Fall:



Sehr unsicher sehr sicher

ANHANG B: DESKRIPTIVER ÜBERBLICK ÜBER DIE VALIDITÄTSEINSCHÄTZUNGEN

Experiment 1

Tabelle B-1: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Validitätseinschätzungen in Experiment 1

		Bedingung 1		Bedingung 2		Bedingung 3		Bedingung 4	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Pretest	wetteronline.de	81.14	13.68	84.86	13.05				
	Sat.1	74.81	11.86	73.84	14.84				
	Bild	68.28	12.24	66.59	14.34				
	ZDF	82.78	11.96	84.78	14.19				
Posttest	wetteronline.de	83.65	11.24	83.78	13.75	83.06	10.06	82.93	13.49
	Sat.1	73.79	11.54	73.23	13.83	74.10	10.17	73.64	12.73
	Bild	71.06	11.88	66.58	14.47	65.48	12.04	67.20	14.29
	ZDF	85.06	10.75	84.38	14.68	84.22	8.63	81.67	12.21
n		50		44		52		50	

Anmerkungen: Die Validitätseinschätzungen sind getrennt für den Messzeitpunkt (Pretest und Posttest) und die Bedingung aufgeführt. Aufgrund der Verwendung eines Solomon-Viergruppenplans wurde der Pretest lediglich in den Bedingungen 1 und 2 bearbeitet. Bedingung 1 umfasste Pretest, Treatment und Posttest, Bedingung 2 Pretest und Posttest, Bedingung 3 Treatment und Posttest und Bedingung 4 umfasste nur den Posttest.

Experiment 2

Tabelle B-2: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Validitätseinschätzungen in Experiment 2

		Einmalige Darbietung		Wiederholte Darbietung	
		M	*SD*	*M*	SD
Pretest	Universität	74.04	13.35	74.94	13.31
	Flughafen	77.01	13.35	75.48	13.39
Posttest	Universität	77.23	12.17	75.77	12.02
	Flughafen	75.36	12.13	77.91	11.80
n		60		57	

Anmerkungen: Die Validitätseinschätzungen sind getrennt für den Messzeitpunkt (Pretest und Posttest) und die Bedingung (einmalige vs. wiederholte Darbietung einer Entscheidungsaufgabe) dargestellt.

Tabelle B-3: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Validitätseinschätzungen in Experiment 2 in Abhängigkeit von der zu erwartenden Entscheidung

		Einmalige Darbietung		Wiederholte Darbietung	
		M	*SD*	*M*	*SD*
Pretest	aufzuwertend	83.24	12.18	80.55	11.37
	abzuwertend	70.04	11.54	68.09	10.73
Posttest	aufzuwertend	82.92	10.28	82.07	10.19
	abzuwertend	68.55	10.70	70.36	10.20
n		46		46	

Anmerkungen: Die Validitätseinschätzungen sind in Abhängigkeit von der gemäß der Höhe der korrigierten gewichteten Summen im Pretest zu erwartenden Entscheidung aufgeführt. „Aufzuwertend“ entspricht dem Cue, der für die zu erwartende Entscheidung spricht, „abzuwertend“ repräsentiert den widersprechenden Cue. Alle Einschätzungen sind getrennt für den Messzeitpunkt (Pretest und Posttest) und die Bedingung (einmalige vs. wiederholte Darbietung einer Entscheidungsaufgabe) dargestellt.

Experiment 3

Tabelle B-4: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Validitätseinschätzungen in Experiment 3

		Bed. 1 (1:1)		Bed. 2 (1:3)		Bed. 3 (3:3)	
		M	*SD*	*M*	*SD*	*M*	*SD*
Pretest	ZDF	77.13	11.46	82.82	10.99	86.63	11.00
	Wetteronline.de	78.56	12.18	80.85	9.40	84.21	11.72
	Bild			64.67	10.96	63.03	12.59
	Sat.1			76.32	9.26	74.19	11.23
	Süddeutsche Zeitung					79.83	10.07
	N24.de					80.47	10.21
Posttest	ZDF	76.32	13.35	84.87	9.62	83.32	11.56
	Wetteronline.de	77.43	12.20	82.80	8.54	84.23	10.36
	Bild			68.50	11.35	62.76	11.55
	Sat.1			73.74	9.95	72.08	12.08
	Süddeutsche Zeitung					80.87	10.47
	N24.de					77.02	11.69
n		37		38		40	

Anmerkungen: Die Validitätseinschätzungen sind getrennt für den Messzeitpunkt (Pretest und Posttest) und die Bedingung aufgeführt. In Bedingung 1 wurde ein 1:1-Muster dargeboten, in Bedingung 2 ein 1:3-Muster und in Bedingung 3 ein 3:3-Muster. Welche und wie viele Cues hinsichtlich ihrer Validität einzuschätzen waren, hing von der Bedingung ab.

Tabelle B-5: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Validitätseinschätzungen in Experiment 3 in Abhängigkeit von der zu erwartenden Entscheidung

		Bed. 1 (1:1)		Bed. 2 (1:3)		Bed. 3 (3:3)	
		M	*SD*	*M*	*SD*	*M*	*SD*
Pretest	aufzuwertend$_1$	82.81	10.91	80.85	9.40	79.48	10.17
	aufzuwertend$_2$			64.67	10.96	84.24	11.29
	aufzuwertend$_3$			82.82	10.99	83.58	11.93
	abzuwertend$_1$	71.95	10.41	76.32	9.26	61.28	11.18
	abzuwertend$_2$					79.55	9.98
	abzuwertend$_3$					72.87	10.55
Posttest	aufzuwertend$_1$	82.59	11.22	82.80	8.54	80.61	10.54
	aufzuwertend$_2$			68.50	11.35	83.06	11.83
	aufzuwertend$_3$			84.87	9.62	83.57	10.44
	abzuwertend$_1$	69.51	11.08	73.74	9.95	61.95	10.46
	abzuwertend$_2$					76.24	11.50
	abzuwertend$_3$					71.50	11.84
n		33		38		37	

Anmerkungen: Die Validitätseinschätzungen sind getrennt für den Messzeitpunkt (Pretest und Posttest) und die Bedingung dargestellt. In Bedingung 1 wurde ein 1:1-Muster dargeboten, in Bedingung 2 ein 1:3-Muster und in Bedingung 3 ein 3:3-Muster. Die Cues wurden entsprechend der zu erwartenden Entscheidung in aufzuwertende (sprechen für die zu erwartende Entscheidung) und abzuwertende (sprechen gegen die zu erwartende Entscheidung) Cues eingeteilt. Da in den Bedingungen 2 und 3 mehr als ein auf- und ein abzuwertender Cue vorliegen, dienen die Indizes der Unterscheidung einzelner Cues.

Experiment 4

Tabelle B-6: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Validitätseinschätzungen in Experiment 4

		Bed. 1 (1:1)		Bed. 2 (1:3)		Bed. 3 (3:3)	
		M	*SD*	*M*	*SD*	*M*	*SD*
Pretest	Lagerdauer	79.05	13.84	77.19	13.81	79.85	13.02
	Reifeperiode	77.98	13.73	75.27	15.92	74.00	13.98
	Färbung			81.28	13.21	74.84	14.91
	Lagerklima			83.07	13.45	84.84	11.33
	Geruch					83.76	13.29
	Transport					80.50	12.61
Posttest	Lagerdauer	78.04	14.48	76.17	13.63	79.94	11.83
	Reifeperiode	77.61	14.87	76.80	14.71	77.15	11.12
	Färbung			80.96	13.67	73.91	14.79
	Lagerklima			83.16	12.97	83.16	10.89
	Geruch					84.87	13.24
	Transport					82.14	9.71
n		37		37		32	

Anmerkungen: Die Validitätseinschätzungen sind getrennt für den Messzeitpunkt (Pretest und Posttest) und die Bedingung aufgeführt. In Bedingung 1 wurde ein 1:1-Muster dargeboten, in Bedingung 2 ein 1:3-Muster und in Bedingung 3 ein 3:3-Muster. Welche und wie viele Cues hinsichtlich ihrer Validität einzuschätzen waren, hing von der Bedingung ab.

Tabelle B-7: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Validitätseinschätzungen in Experiment 4 in Abhängigkeit von der Entscheidung

		Bed. 1 (1:1)		Bed. 2 (1:3)		Bed. 3 (3:3)	
		M	*SD*	*M*	*SD*	*M*	*SD*
Pretest	aufzuwertend$_1$	87.24	9.47	83.56	11.71	81.49	14.38
	aufzuwertend$_2$			76.00	15.83	80.96	16.17
	aufzuwertend$_3$			83.16	12.77	85.14	10.26
	abzuwertend$_1$	70.06	11.58	77.44	14.19	73.15	13.75
	abzuwertend$_2$					75.78	15.19
	abzuwertend$_3$					82.39	13.03
Posttest	aufzuwertend$_1$	85.19	12.98	83.47	12.07	84.04	11.02
	aufzuwertend$_2$			77.56	15.03	82.36	15.33
	aufzuwertend$_3$			84.16	12.25	83.24	10.52
	abzuwertend$_1$	70.34	11.92	75.63	13.60	74.21	12.19
	abzuwertend$_2$					75.96	15.55
	abzuwertend$_3$					83.34	11.02
n		32		33		25	

Anmerkungen: Die Validitätseinschätzungen sind getrennt für den Messzeitpunkt (Pretest und Posttest) und die Bedingung dargestellt. In Bedingung 1 wurde ein 1:1-Muster dargeboten, in Bedingung 2 ein 1:3-Muster und in Bedingung 3 ein 3:3-Muster. Die Cues wurden entsprechend der Entscheidung in aufzuwertende (sprechen für die zu erwartende Entscheidung) und abzuwertende (sprechen gegen die zu erwartende Entscheidung) Cues eingeteilt. Da in den Bedingungen 2 und 3 mehr als ein auf- und ein abzuwertender Cue vorliegen, dienen die Indizes der Unterscheidung einzelner Cues.

Experiment 5

Tabelle B-8: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Validitätseinschätzungen in Experiment 5

		M	*SD*
Pretest	wetteronline.de	84.18	9.82
	Sat.1	75.29	9.47
	Bild	66.94	11.35
	ZDF	85.09	9.42
Posttest	wetteronline.de	84.84	8.45
	Sat.1	73.70	11.03
	Bild	70.45	12.16
	ZDF	85.29	9.14
n		50	

Anmerkungen: Die Validitätseinschätzungen sind getrennt für den Messzeitpunkt (Pretest und Posttest) aufgeführt.

Experiment 6

Tabelle B-9: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Validitätseinschätzungen in Experiment 6

		Einmalige Darbietung		Wiederholte Darbietung	
		M	*SD*	*M*	*SD*
Pretest	Augenzeuge	**70.29**	**10.63**	63.44	12.20
	Arbeitsweise	55.68	9.70	58.01	11.98
	Ausgaben	64.94	21.34	64.28	21.99
	Gesetz	56.34	9.25	55.17	11.27
	Fingerabdruck	65.09	18.83	69.12	17.37
	Anruf	**75.68**	**24.62**	79.19	22.41
	Auto	**61.84**	**19.83**	59.74	17.24
	Ort	**75.51**	**21.59**	82.87	17.56
Posttest	Augenzeuge	**62.90**	**15.98**	57.35	10.06
	Arbeitsweise	53.92	7.46	56.43	10.38
	Ausgaben	65.42	19.39	52.85	14.37
	Gesetz[a]	55.74	9.39	55.72	12.02
	Fingerabdruck[a]	62.65	15.58	63.53	15.17
	Anruf	**77.96**	**18.56**	78.27	18.05
	Auto	**60.97**	**17.66**	54.74	13.64
	Ort	**77.87**	**21.09**	72.72	16.78
n		34		34	

Anmerkungen: Die Validitätseinschätzungen sind getrennt für den Messzeitpunkt (Pretest und Posttest) und die Bedingung (einmalige vs. wiederholte Darbietung einer Entscheidungsaufgabe) dargestellt. Fett hervorgehoben sind die Werte für die vier Cues, die in der Entscheidungssituation der Bedingung „Einmalige Darbietung" auftraten. [a]Eine Person gab keine Validitätseinschätzung ab.

Experiment 7

Tabelle B-10: Mittelwerte (*M*) und Standardabweichungen (*SD*) der Validitätseinschätzungen in Experiment 7

		Einmalige Darbietung		Wiederholte Darbietung	
		M	*SD*	*M*	*SD*
Pretest	Augenzeuge	74.28	12.91	72.78	12.19
	Fingerabdruck	77.54	16.06	75.01	15.61
	Anruf	79.12	18.20	81.35	17.96
	Ort	80.71	18.11	81.33	16.79
	Auto	65.59	16.37	64.92	14.92
	Ausgaben	70.55	17.51	69.55	16.35
	Arbeitsweise	60.62	11.84	59.56	12.75
	Gesetz	65.14	13.61	64.84	14.87
Posttest	Augenzeuge	65.04	11.29	60.15	11.21
	Fingerabdruck	71.58	16.19	75.04	17.04
	Anruf	77.76	16.68	85.82	15.21
	Ort	80.17	15.35	87.92	11.37
	Auto	63.41	12.67	70.19	15.67
	Ausgaben	66.91	14.22	69.54	17.00
	Arbeitsweise	56.12	8.90	62.23	13.42
	Gesetz	61.96	13.19	63.52	14.59
n		62		64	

Anmerkungen: Die Validitätseinschätzungen sind getrennt für den Messzeitpunkt (Pretest und Posttest) und die Bedingung (einmalige vs. wiederholte Darbietung einer Entscheidungsaufgabe) dargestellt.

ANHANG C: WEITERE ERGEBNISSE

Experiment 6

Tabelle C-1: Klassifikation der Cues für Bedingung 1 von Experiment 6

Messzeitpunkt	Cue	Richtungsangabe			Klassifikation
		Weder noch	*Schuldig*	*Unschuldig*	
Pretest	Augenzeuge	3	31	0	Schuldig
	Arbeitsweise	20	1	13	Weder noch
	Ausgaben	11	7	16	/
	Gesetz	16	3	15	/
	Fingerabdruck	8	3	23	Unschuldig
	Anruf	1	9	24	Unschuldig
	Auto	10	5	19	Unschuldig
	Ort	4	3	27	Unschuldig
Posttest	Augenzeuge	3	28	3	Schuldig
	Arbeitsweise	23	1	10	Weder noch
	Ausgaben	8	5	21	Unschuldig
	Gesetz[a]	14	5	14	/
	Fingerabdruck[a]	7	4	22	Unschuldig
	Anruf	3	1	30	Unschuldig
	Auto	9	4	21	Unschuldig
	Ort	1	3	30	Unschuldig

Anmerkungen: Die Klassifikation wird auf Basis der Richtungsangaben (absolute Häufigkeiten) getrennt für den Pre- und den Posttest dargestellt. „Weder noch“ bezeichnet eine Angabe in der Mitte der Skala (50% schuldig/50% unschuldig), „schuldig“ bezeichnet eine Angabe auf der rechten Seite und „unschuldig“ auf der linken Seite der Skala. Es wurde keine Klassifikation vorgenommen, wenn in keiner Kategorie eine absolute Mehrheit vorlag. [a]Eine Versuchsperson gab keine Validitätseinschätzung ab.

Tabelle C-2: Klassifikation der Cues für Bedingung 2 von Experiment 6

Messzeitpunkt	Cue	Richtungsangabe			Klassifikation
		Weder noch	*Schuldig*	*Unschuldig*	
Pretest	Augenzeuge	3	29	2	Schuldig
	Arbeitsweise	18	0	16	Weder noch
	Ausgaben	14	4	16	/
	Gesetz	19	2	13	Weder noch
	Fingerabdruck	5	1	28	Unschuldig
	Anruf	2	6	26	Unschuldig
	Auto	13	2	19	Unschuldig
	Ort	2	2	30	Unschuldig
Posttest	Augenzeuge	8	24	2	Schuldig
	Arbeitsweise	17	1	16	Weder noch
	Ausgaben	14	8	12	/
	Gesetz	15	4	15	/
	Fingerabdruck	8	2	24	Unschuldig
	Anruf	2	2	30	Unschuldig
	Auto	13	3	18	Unschuldig
	Ort	1	3	30	Unschuldig

Anmerkungen: Die Klassifikation wird auf Basis der Richtungsangaben (absolute Häufigkeiten) getrennt für den Pre- und den Posttest dargestellt. „Weder noch“ bezeichnet eine Angabe in der Mitte der Skala (50% schuldig/50% unschuldig), „schuldig“ bezeichnet eine Angabe auf der rechten Seite und „unschuldig“ auf der linken Seite der Skala. Es wurde keine Klassifikation vorgenommen, wenn in keiner Kategorie eine absolute Mehrheit vorlag.

Tabelle C-3: Überblick über die zentralen Effekte für die Bedingung „einmalige Darbietung“ von Experiment 6

		alle Entscheidungen	$T_{schuldig} < 3$	$E_{schuldig} < 3$
Optionen-Ebene	Vergleich der Differenz zw. Optionen *(spreading apart)* [a]	$t(33) = 1.73$, $p < .05$	$t(32) = 1.59$, $p = .06$	$t(31) = 1.74$, $p < .05$
Cue-Ebene	MZP	$F(1, 33) < 1$, $p = .57$, $\eta_p^2 = .01$	$F(1, 32) < 1$, $p = .59$, $\eta_p^2 = .01$	$F(1, 31) < 1$, $p = .59$, $\eta_p^2 = .01$
	MZP x Cue	$F(2.65, 87.37) = 3.36$, $p < .05$, $\eta_p^2 = .09$ [b]	$F(3, 96) = 4.04$, $p < .01$, $\eta_p^2 = .11$	$F(3, 93) = 3.56$, $p < .05$, $\eta_p^2 = .10$
	Vergleich der Val. im Post- und Pretest [a]	$t(33) = -2.98$, $p < .01$	$t(32) = -3.27$, $p < .01$	$t(31) = -3.11$, $p < .01$

Anmerkungen: Es werden nur die Analysen mit 4 Cues und nur die interessierenden Effekte dargestellt. Alle Entscheidungen: In den Analysen wurden alle Personen unabhängig von der tendenziellen und finalen Entscheidung berücksichtigt. $T_{schuldig} < 3$: Es wurden nur die Versuchspersonen, die sich weniger als dreimal tendenziell für schuldig entschieden, berücksichtigt. $E_{schuldig} < 3$: Es wurden nur die Personen, die sich weniger als dreimal final für schuldig entschieden, berücksichtigt. MZP: Messzeitpunkt. [a]Einseitige Testung. [b]Korrektur der Freiheitsgrade nach Huynh-Feldt.

Tabelle C-4: Überblick über die zentralen Effekte für die Bedingung „wiederholte Darbietung“ von Experiment 6

		Alle Entscheidungen	$T_{schuldig} < 3$	$E_{schuldig} < 3$
Option-Ebene	MZP	$F(1, 32) < 1$, $p = .78$, $\eta_p^2 = .00$	$F(1, 20) < 1$, $p = .63$, $\eta_p^2 = .01$	$F(1, 21) < 1$, $p = .54$, $\eta_p^2 = .02$
	MZP x Fall	$F(3.72, 118.96) = 1.45$, $p = .23$, $\eta_p^2 = .04$[a]	$F(11,220) = 2.08$, $p < .05$, $\eta_p^2 = .10$	$F(11, 231) = 1.66$, $p = .19$, $\eta_p^2 = .07$
Cue-Ebene	MZP	$F(1, 33) = 12.02$, $p < .01$, $\eta_p^2 = .27$	$F(1, 21) = 3.80$, $p = .07$, $\eta_p^2 = .15$	$F(1, 22) = 4.89$, $p < .05$, $\eta_p^2 = .18$
	MZP x Cue	$F(5.73, 188.93) = 1.84$, $p = .10$, $\eta_p^2 = .05$[b]	$F(3.87, 81.26) = 1.92$, $p = .12$, $\eta_p^2 = .08$[a]	$F(3.96, 87.04) = 1.25$, $p = .29$, $\eta_p^2 = .05$[a]
	Vergleich der Val. im Post- und Pretest[c]	Augenzeuge: $t(33) = -2.99$, $p < .01$ Ausgaben: $t(33) = -3.28$, $p < .01$ Ort: $t(33) = -2.50$, $p < .05$	Augenzeuge: $t(21) = -3.02$, $p < .01$ Ausgaben: $t(21) = -2.98$, $p < .01$	Augenzeuge: $t(22) = -3.00$, $p < .01$ Ausgaben: $t(22) = -2.18$, $p < .05$

Anmerkungen: Es werden nur die interessierenden Effekte dargestellt. Alle Entscheidungen: In den Analysen wurden alle Personen unabhängig von der tendenziellen und finalen Entscheidung berücksichtigt. $T_{schuldig} < 3$: Es wurden nur die Versuchspersonen, die sich weniger als dreimal tendenziell für schuldig entschieden, berücksichtigt. $E_{schuldig} < 3$: Es wurden nur die Personen, die sich weniger als dreimal final für schuldig entschieden, berücksichtigt. MZP: Messzeitpunkt. [a]Korrektur der Freiheitsgrade nach Greenhouse-Geisser. [b]Korrektur der Freiheitsgrade nach Huynh-Feldt. [c]Einseitige Testung.

Tabelle C-5: Vergleich des *spreading aparts* für den in beiden Bedingungen von Experiment 6 präsentierten Fall

	Alle Entscheidungen	$T_{schuldig} < 3$	$E_{schuldig} < 3$
Spreading apart "einmalige Darbietung"[a]	$t(33) = 1.73$, $p < .05$	$t(32) = 1.59$, $p = .06$	$t(31) = 1.74$, $p < .05$
Spreading apart "wiederholte Darbietung"[a]	$t(33) = 0.91$, $p = .18$	$t(31) = 0.81$, $p = .21$	$t(29) = 0.06$, $p = .48$

Anmerkungen: Alle Entscheidungen: In den Analysen wurden alle Personen unabhängig von der tendenziellen und finalen Entscheidung berücksichtigt. $T_{schuldig} < 3$: Es wurden nur die Versuchspersonen, die sich weniger als dreimal tendenziell für schuldig entschieden, berücksichtigt. $E_{schuldig} < 3$: Es wurden nur die Personen, die sich weniger als dreimal final für schuldig entschieden, berücksichtigt. [a]Einseitige Testung.

Experiment 7

Tabelle C-6: Überblick über die zentralen Effekte für die Bedingung „reine Wiederholung“ von Experiment 7

		Alle Entscheidungen	$T_{schuldig} < 3$	$E_{schuldig} < 3$
Option-Ebene	MZP	$F(1, 61) < 1$, $p = .95$, $\eta_p^2 = .00$	$F(1, 43) < 1$, $p = .67$, $\eta_p^2 = .00$	$F(1, 41) < 1$, $p = .79$, $\eta_p^2 = .00$
	MZP x Fall	$F(5.43, 331.17) = 1.42$, $p = .21$, $\eta_p^2 = .02$[a]	$F(11, 473) = 1.50$, $p = .13$, $\eta_p^2 = .03$	$F(3.34, 136.78) = 1.35$, $p = .26$, $\eta_p^2 = .03$[a]
Cue-Ebene	MZP	$F(1, 61) = 23.71$, $p < .001$, $\eta_p^2 = .28$	$F(1, 43) = 15.62$, $p < .001$, $\eta_p^2 = .27$	$F(1, 41) = 15.06$, $p < .001$, $\eta_p^2 = .27$
	MZP x Cue	$F(5.92, 361.00) = 1.86$, $p = .09$, $\eta_p^2 = .03$[b]	$F(4.59, 197.51) = 2.18$, $p = .06$, $\eta_p^2 = .05$[a]	$F(5.57, 228.17) = 2.02$, $p = .07$, $\eta_p^2 = .05$[b]
	Vergleich der Val. im Post- und Pretest[c]	Augenzeuge: $t(61) = -4.88$, $p < .001$ Arbeitsweise: $t(61) = -3.05$, $p < .01$ Ausgaben: $t(61) = -1.70$, $p < .05$ Gesetz: $t(61) = -1.93$, $p < .05$ Fingerabdruck: $t(61) = -3.18$, $p < .01$	Augenzeuge: $t(43) = -4.34$, $p < .001$ Ausgaben: $t(43) = -2.81$, $p < .01$ Fingerabdruck: $t(43) = -2.67$, $p < .01$	Augenzeuge: $t(41) = -3.72$, $p < .001$ Arbeitsweise: $t(41) = -1.69$, $p < .05$ Ausgaben: $t(41) = -2.58$, $p < .01$ Fingerabdruck: $t(41) = -2.82$, $p < .01$

Anmerkungen: Es werden nur die interessierenden Effekte dargestellt. Alle Entscheidungen: In den Analysen wurden alle Personen unabhängig von der tendenziellen und finalen Entscheidung berücksichtigt. $T_{schuldig} < 3$: Es wurden nur die Versuchspersonen, die sich weniger als dreimal tendenziell für schuldig entschieden, berücksichtigt. $E_{schuldig} < 3$: Es wurden nur die Personen, die sich weniger als dreimal final für schuldig entschieden, berücksichtigt. MZP: Messzeitpunkt. [a]Korrektur der Freiheitsgrade nach Greenhouse-Geisser. [b]Korrektur der Freiheitsgrade nach Huynh-Feldt. [c]Einseitige Testung.

Tabelle C-7: Überblick über die zentralen Effekte für die Bedingung „Information“ von Experiment 7

		Alle Entscheidungen	$T_{schuldig} < 3$
Option-Ebene	MZP	$F(1, 63) = 25.69$, $p < .001$, $\eta_p^2 = .29$	$F(1, 58) = 28.34$, $p < .001$, $\eta_p^2 = .33$
	MZP x Fall	$F(11, 693) = 1.95$, $p < .05$, $\eta_p^2 = .03$	$F(11, 638) = 2.93$, $p < .01$, $\eta_p^2 = .05$
Cue-Ebene	MZP	$F(1, 63) < 1$, $p = .60$, $\eta_p^2 = .00$	$F(1, 58) < 1$, $p = .36$, $\eta_p^2 = .02$
	MZP x Cue	$F(7, 441) = 9.03$, $p < .001$, $\eta_p^2 = .13$	$F(7, 406) = 9.57$, $p < .001$, $\eta_p^2 = .14$
	Vergleich der Val. im Post- und Pretest[a]	Augenzeuge: $t(63) = -6.71$, $p < .001$ Auto: $t(63) = 2.39$, $p < .05$ Ort: $t(63) = 2.90$, $p < .01$	Augenzeuge: $t(58) = -6.60$, $p < .001$ Anruf: $t(58) = 1.99$, $p < .05$ Auto: $t(58) = 2.75$, $p < .01$ Ort: $t(58) = 3.01$, $p < .01$

Anmerkungen: Es werden nur die interessierenden Effekte dargestellt. Alle Entscheidungen: In den Analysen wurden alle Personen unabhängig von der tendenziellen Entscheidung berücksichtigt. $T_{schuldig} < 3$: Es wurden nur die Versuchspersonen, die sich weniger als dreimal tendenziell für schuldig entschieden, berücksichtigt. MZP: Messzeitpunkt. [a]Einseitige Testung.

Lebenslauf

Tanja Ostermann
Diplom-Psychologin

28. April 1982	Geboren in Haselünne/Emsland
06/2001	Abitur am Kreisgymnasium St. Ursula, Haselünne/Emsland
10/2001 – 09/2003	Grundstudium der Psychologie, Universität Trier
10/2003 – 03/2004	AIESEC Auslandsaufenthalt in Kolkata, Indien
10/2003 – 03/2007	Hauptstudium der Psychologie, Universität Trier Abschluss: Diplom-Psychologin
08/2007 – 07/2010	Promotionsstipendiatin am Max-Planck-Institut zur Erforschung von Gemeinschaftsgütern in Bonn, Mitglied der Forschungsgruppe *Intuitive Experts* (Leitung: Dr. Andreas Glöckner)

Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts habe ich keine Unterstützungsleistung erhalten.

Weitere Personen waren an der geistigen Herstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe eines Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit oder Teile davon wurden bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde als Dissertation vorgelegt. Ferner erkläre ich, dass ich nicht bereits eine gleichartige Doktorprüfung an einer Hochschule endgültig nicht bestanden habe.

Düsseldorf, 2010
Tanja Ostermann