

# **Informationsanordnung und kausale Urteilsbildung**

## **Einfluss unterschiedlicher externaler Repräsentationen auf die Beurteilung der Stärke monokausaler Zusammenhänge**

### **Dissertation**

der Fakultät für Informations- und Kognitionswissenschaften

der Eberhard-Karls-Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

**Dipl.-Psych. Ralf Decker**

aus Fulda

Tübingen 2007



## Danksagung

Auch ich möchte einigen Menschen danken, die mir die Erstellung dieser Arbeit ermöglicht und mich in den verschiedenen Phasen ihrer Entstehung unterstützt haben.

Erstellt wurde diese Arbeit im Rahmen des Virtuellen Graduiertenkollegs (VGK) *Wissenserwerb und Wissensaustausch mit neuen Medien*, welches von der DFG finanziert wurde. Insofern möchte ich mich bei der DFG bedanken, die mich und mein Projekt in den ersten drei Jahren durch ein Stipendium finanziell unterstützt hat, und bei den Dozenten des VGK, die sich für meine Aufnahme in das VGK entschieden haben. Insbesondere möchte ich Herrn Prof. Dr. Dr. Friedrich W. Hesse für die Erstbetreuung, hilfreiche Ratschläge und Weichenstellungen danken. Er hat es mir auch ermöglicht, dass mir am Institut für Wissensmedien (IWM) in Tübingen ein Arbeitsraum zur Verfügung gestellt wurde und dass ich die Infrastruktur am IWM nutzen konnte. Danken möchte ich auch meinem Zweitbetreuer Herrn Prof. Dr. Rolf Plötzner für konstruktive Rückmeldungen und Gespräche. Ein ganz besonderer Dank gebührt auch Herrn Dr. Uwe Oestermeier. Er war es, der mich auf das VGK aufmerksam gemacht hat und der mich in allen Phasen des Projekts intensiv betreut und tatkräftig unterstützt hat. Er hat mich immer wieder zum Weitermachen motiviert und die gewiss nicht immer angenehme Aufgabe übernommen, meine Entwürfe Korrektur zu lesen.

Danken möchte ich auch allen VGK-Dozenten und StipendiatInnen, mit denen ich gemeinsam an VGK-Treffen teilgenommen habe, für ihre hilfreichen Rückmeldungen zu meinen Präsentationen, für gemeinsame Unternehmungen und für angenehme Gespräche, die wir während der Treffen geführt haben. Vom IWM möchte ich vielen Menschen für vielfältigste Unterstützung und die angenehme Arbeitsatmosphäre danken, insbesondere Tanja Engelmann, Ingo Kollar, Carmen Zahn, Erik Ründal, Sebastian Groteloh, Jörn Töpfer, Daniel Bodemer, Susanne Frings, Frank Fischer, Stefan Schwan, Jürgen Buder, Daniela Straub, Markus Huff, Charlotte Haimerl, Rosi Croizier, Margot Stoll, Steffen-Peter Ballstaedt, Stefan Martin, Peter Zentel, Peter Gerjets, Ulrike Cress, Oliver Knabel und Herrn Alonso.

Schließlich möchte ich mich auch bei meiner Familie bedanken; bei meinen Eltern, Eva, Lea-Marie und Salome. Sie haben den intensivsten Anteil an der Entstehung dieser Arbeit genommen und mussten deswegen auf Vieles verzichten. Entsprechend freuen Sie sich mit mir am meisten über den Abschluss dieser Arbeit. Besonders möchte ich auch meiner Schwiegermutter Luise Reich danken, die mir gerade in den letzten Wochen der Entstehung dieser Arbeit immer wieder den Rücken frei gehalten hat.



# Inhaltsverzeichnis

<b>Überblick und Zusammenfassung .....</b>	<b>I</b>
--	----------

<b>Teil 1: Theoretischer Hintergrund .....</b>	<b>1</b>
--	----------

<b>1 Einführung .....</b>	<b>3</b>
---------------------------	----------

<b>2 Philosophische Sichtweisen zur Kausalität und ihr Einfluss auf die kognitionspsychologische Theorienbildung zum kausalen Lernen .....</b>	<b>5</b>
--	----------

2.1 Traditionelles Kausalitätsverständnis in der Philosophie: Aristoteles, Hume und Kant.....	5
--	---

2.2 Neuere Konzeptionen des Kausalitätsbegriffs .....	11
---	----

2.3 Psychologische Rezeption von neueren Konzeptionen des Kausalitätsbegriffs .....	15
---	----

<b>3 Der Erwerb von Kausalwissen .....</b>	<b>19</b>
--	-----------

3.1 Regel-basierte Ansätze .....	23
----------------------------------	----

3.1.1 Kontingenztheoretische Ansätze .....	23
--	----

3.1.2 Das Probabilistische Kontrastmodell .....	34
---	----

3.1.3 Die Power PC Theorie .....	38
----------------------------------	----

3.1.4 Bestimmung der kausalen Stärke nach der Theorie der empirischen Evidenzevaluation .....	46
---	----

3.1.5 Über den kognitiven Status regel-basierter Strategien. Oder: wie sind regel-basierte Strategien mental repräsentiert?.....	49
---	----

3.2 Assoziationalistische Ansätze .....	53
---	----

3.3 Weitere Ansätze: Die Rolle von Vorwissen .....	59
--	----

3.3.1 Der mechanistische Ansatz .....	59
---------------------------------------	----

3.3.2 Theorie der Kausalmodelle und Bayes-Netze .....	61
---	----

3.4 Kritik an monomechanistischen Theorieansätzen: Schließen assoziative und regel-basierte Mechanismen einander aus? .....	63
--	----

3.5 Ein möglicher Rahmen für einen integrativen Ansatz: Die Theorie von Shanks (1991) .....	65
--	----

<b>4</b>	<b>Der Einfluss von Darbietungsart, Informationsart und Präsentationsform auf die kausale Urteilsbildung.....</b>	<b>69</b>
4.1	Die Bedeutung situativer und aufgabenbezogener Faktoren.....	69
4.2	Empirische Befunde zum Einfluss von Darbietungsart, Informationsart und Präsentationsform .....	71
4.2.1	Einfluss der Darbietungsart .....	76
4.2.2	Einfluss der Informationsart .....	82
4.2.3	Einfluss der Präsentationsform .....	83
<b>5</b>	<b>Individuelle Unterschiede und Gedächtnisbeanspruchung bei der kausalen Urteilsbildung.....</b>	<b>87</b>
5.1	Einfluss individueller Unterschiede.....	87
5.2	Bedeutung von Gedächtnisbeanspruchung und Verarbeitungsaufwand .....	90
<b>6</b>	<b>Theorie und Empirie zum Einfluss von Medien auf die menschliche Informationsverarbeitung .....</b>	<b>97</b>
6.1	Die Kosten-Nutzen-Theorie von Kleinmuntz & Schkade.....	100
6.2	Auswirkungen externaler Repräsentationen in komplexen Informationsverarbeitungsaufgaben: Die Theorie von Zhang .....	105
6.3	Charakteristika und kognitive Verarbeitung von Diagrammen, Tabellen und Texten .....	109
6.3.1	Texte .....	110
6.3.2	Tabellen .....	111
6.3.3	Diagramme .....	112
6.4	Empirische Studien zum Vergleich zwischen Diagrammen, Tabellen und Texten.....	119
6.4.1	Die Theorie der kognitiven Passung von Vessey .....	121
<b>7</b>	<b>Fazit und Implikationen für die empirischen Untersuchungen.....</b>	<b>123</b>
7.1	Zusammenfassung der wesentlichen Ergebnisse .....	123
7.2	Die vier Ebenen der Analyse kognitiver Prozesse bei der Beurteilung der Stärke kausaler Zusammenhänge.....	129
7.3	Unter welchen Bedingungen kommen regel-basierte Strategien mit höherer Wahrscheinlichkeit zum Einsatz? .....	134
7.4	Ausblick auf Fragestellungen und Untersuchungsziele .....	138

## Teil 2: Empirische Untersuchungen .....139

<b>8 Experiment 1:</b>	
<b>Einfluss von Präsentationsform und Informationsart .....</b>	<b>141</b>
8.1 Untersuchungsdesign .....	143
8.2 Forschungshypothesen .....	150
8.2.1 Einfluss des Faktors Präsentationsform .....	150
8.2.2 Einfluss des Faktors Informationsart .....	153
8.2.3 Vergleich der Experimentalbedingungen mit einer Protokollliste .....	154
8.2.4 Einfluss des Faktors Studenausrichtung .....	156
8.3 Methodik und Vorgehensweise .....	158
8.4 Explorative Datenanalysen .....	161
8.5 Einfluss der Faktoren Präsentationsform und Informationsart .....	162
8.5.1 Ergebnisse .....	162
8.5.2 Diskussion .....	166
8.6 Häufigkeitsanalysen zum Einfluss der Versuchsbedingungen auf die Strategiewahl .....	174
8.6.1 Ergebnisse .....	174
8.6.2 Diskussion .....	181
8.7 Einfluss der Studenausrichtung und Analysen zu Teilstichprobeneffekten .....	185
8.7.1 Einfluss der Studenausrichtung .....	185
8.7.1.1 Ergebnisse .....	185
8.7.1.2 Diskussion .....	188
8.7.2 Teilstichprobeneffekte.....	188
8.7.2.1 Ergebnisse .....	188
8.7.2.2 Diskussion .....	191
8.8 Ergebnisse der Nachbefragung .....	193
8.9 Abschließende Diskussion von Experiment 1 .....	197

<b>9</b>	<b>Experiment 2: Einfluss von Darbietungsart und visueller Strukturierung des darstellenden Mediums .....</b>	<b>203</b>
9.1	Untersuchungsdesign .....	205
9.2	Forschungshypothesen .....	209
9.2.1	Einfluss des Faktors Darbietungsart.....	209
9.2.2	Einfluss des Faktors visuelle Strukturierung des darstellenden Mediums .....	211
9.2.4	Einfluss der Kovariate numerisches Denken .....	213
9.2.5	Einfluss des Faktors Studenausrichtung .....	214
9.3	Methodik und Vorgehensweise .....	216
9.4	Explorative Datenanalysen .....	221
9.5	Einfluss der Faktoren Darbietungsart und visuelle Strukturierung .....	222
9.5.1	Ergebnisse .....	222
9.5.2	Diskussion .....	232
9.6	Häufigkeitsanalysen zum Einfluss der Versuchsbedingungen auf die Strategiewahl .....	237
9.6.1	Ergebnisse .....	237
9.6.2	Diskussion .....	243
9.7	Einfluss von Studenausrichtung und Unterschieden im numerischem Denken .....	247
9.7.1	Einfluss der Studenausrichtung .....	247
9.7.1.1	<i>Ergebnisse</i> .....	247
9.7.1.2	<i>Diskussion</i> .....	251
9.7.2	Einfluss von Unterschieden im numerischen Denken .....	254
9.7.2.1	<i>Ergebnisse</i> .....	254
9.7.2.2	<i>Diskussion</i> .....	257
9.8	Abschließende Diskussion von Experiment 2 .....	260



<b>10 Experiment 3: Einfluss von Notizen bei Einzelfalldarbietung .....</b>	<b>265</b>
10.1 Untersuchungsdesign .....	268
10.2 Forschungshypothesen .....	270
10.2.1 Einfluss des Faktors Notizen anfertigen .....	270
10.2.2 Einfluss der Kovariate numerisches Denken .....	273
10.2.3 Einfluss des Faktors Studenausrichtung .....	273
10.3 Methodik und Vorgehensweise .....	275
10.4 Explorative Datenanalysen.....	276
10.5 Einfluss des Faktors Notizen anfertigen .....	277
10.5.1 Anzahl regel-konformer Lösungen .....	277
10.5.1.1 Ergebnisse .....	277
10.5.1.2 Diskussion .....	281
10.5.2 Bearbeitungszeit .....	288
10.5.2.1 Ergebnisse .....	288
10.5.2.2 Diskussion .....	291
10.5.3 Verarbeitungseffizienz .....	293
10.5.3.1 Ergebnisse .....	293
10.5.3.2 Diskussion .....	294
10.6 Häufigkeitsanalysen zum Einfluss von Notizen auf die Strategiewahl .....	296
10.6.1 Ergebnisse .....	296
10.6.2 Diskussion .....	300
10.7 Einfluss von Studenausrichtung und Unterschieden im numerischem Denken .....	303
10.7.1 Studenausrichtung .....	303
10.7.1.1 Ergebnisse .....	303
10.7.1.2 Diskussion .....	307
10.7.2 Einfluss von Unterschieden im numerischen Denken .....	311
10.7.2.1 Ergebnisse .....	311
10.7.2.2 Diskussion .....	313
10.8 Abschließende Diskussion von Experiment 3 .....	318

<b>Teil 3: Schlussfolgerungen .....</b>	<b>321</b>
<b>11 Zusammenfassende Diskussion .....</b>	<b>323</b>
<b>12 Ausblick .....</b>	<b>335</b>
<b>Literaturverzeichnis .....</b>	<b>341</b>
<b>Abbildungsverzeichnis .....</b>	<b>373</b>
<b>Tabellenverzeichnis .....</b>	<b>377</b>
<b>Anhang .....</b>	<b>A 1</b>

# Überblick und Zusammenfassung

## Teil 1: Theoretischer Hintergrund

**Kapitel 1** soll eine Einordnung der vorliegenden Arbeit in das weite Feld der Forschung zur Kausalität innerhalb der kognitiven Psychologie ermöglichen. Des Weiteren werden Ausgangspunkt und Zielsetzung dieser Arbeit skizziert.

In **Kapitel 2** werden philosophische Sichtweisen zum Kausalitätsbegriff vorgestellt. Neben den traditionellen Sichtweisen von Aristoteles, Hume und Kant werden neuere Ansätze, insbesondere der bedingungsanalytische Ansatz von Mackie (1974) und probabilistische Theorien zur Kausalität, zusammengefasst. Darüber hinaus werden die Einflüsse dieser verschiedenen philosophischen Sichtweisen auf die psychologische Theorienbildung zum Erwerb von Kausalwissen diskutiert.

In **Kapitel 3** wird der kognitionspsychologische Forschungsstand über den Erwerb von Kausalwissen referiert. Den größten Raum nehmen darin die im Fokus der vorliegenden Arbeit stehenden regel-basierten Ansätze ein. Sie werden in Kapitel 3.1 beschrieben. Besonders einflussreiche regel-basierte Ansätze sind die kontingenztheoretischen Ansätze und die Theorie der kausalen Power nach Cheng (1997). Kontingenztheoretischen Ansätzen zufolge, kann die Stärke eines monokausalen Zusammenhangs mit nur einer Wirkung und einer in Frage kommenden Ursache durch das Kontingenzmaß  $\Delta P$  (Allan, 1993; McKenzie, 1994; Ward & Jenkins, 1965) bestimmt werden. Bei der Theorie der kausalen Power handelt es sich um eine Weiterentwicklung kontingenztheoretischer Ansätze. Sie wurde unter anderem deshalb von Cheng (1997) entwickelt, um einige der Probleme zu überwinden, die mit rein kontingenztheoretischen Ansätzen verbunden sind. Weniger ausführlich werden in Kapitel 3.2 assoziationalistische Ansätze, sowie in Kapitel 3.3 der mechanistische Ansatz und die Theorie der Kausalmodelle vorgestellt. Der Überblick zu unterschiedlichen Theorien über den Erwerb von Kausalwissen mündet in einer Kritik der Annahme, dass der komplexe Prozess der kausalen Urteilsbildung nur durch einen einzigen Mechanismus allumfassend erklärt werden könne (Kapitel 3.4). Nach der Theorie von Shanks (1991), welche einen Rahmen zur Integration von assoziationalistischen und regel-basierten Ansätzen bereit stellt, hängt es stattdessen vom Situationstypus ab, ob der Bildung von Kausalurteilen assoziative Mechanismen oder kausale Schemata zugrunde liegen (Kapitel 3.5). Allerdings wird in

Verbindung mit der Theorie von Shanks (1991) auch darauf verwiesen, dass dieser Ansatz kaum dazu in der Lage ist, die unterschiedlichen Einflüsse der vielfältigen situativen Rahmenbedingungen, die bei der kausalen Urteilsbildung eine Rolle spielen, zu erklären und ihnen gerecht zu werden.

In **Kapitel 4** werden empirische Belege für mögliche Einflüsse der nachfolgend im empirischen Teil dieser Arbeit untersuchten medialen Faktoren auf die Beurteilung der Stärke von kausalen Zusammenhängen diskutiert. Kapitel 4.1 liefert zunächst ein Plädoyer für die Bedeutung von situativen und aufgabenspezifischen Faktoren für die bei der Beurteilung der Stärke von kausalen Zusammenhängen ablaufenden kognitiven Prozesse und Mechanismen. Dieses Plädoyer endet mit der Annahme, dass unterschiedliche Situations- und Aufgabenkontexte unterschiedliche Prozesse und Mechanismen bei der kausalen Urteilsbildung aktivieren oder auslösen können. In Kapitel 4.2 werden der innerhalb der Literatur zur Kausalkognition selten verwendete Medien- und der häufiger verwendete Aufgabenbegriff voneinander abgegrenzt und empirische Belege für einen möglichen Einfluss der ebenfalls definierten Faktoren Darbietungsart, Präsentationsform und Informationsart auf die Beurteilung der Stärke von kausalen Zusammenhängen präsentiert. Während im Zusammenhang mit den Einflüssen von Präsentationsform und Informationsart auf die kausale Urteilsbildung kaum Befunde existieren, liegen in Verbindung mit dem Einfluss der Darbietungsart einige Ergebnisse vor. Dabei ergab sich das Bild, dass regel-basierte Strategien häufiger bei summarischer Informationsdarbietung zum Einsatz kommen. Voraussetzung dafür ist allerdings, dass die Komplexität der Aufgabenstellung nicht zu groß ist. Bei einzelfallweiser Darbietung steigen die Komplexität der Aufgabenstellung und die Gedächtnisbeanspruchung in der Regel an. Das beinhaltet, dass im Vergleich zur summarischen Darbietung der Einsatz regel-basierter Strategien (wie z.B. der  $\Delta P$ -Regel) unwahrscheinlich, aber nicht ganz ausgeschlossen ist.

**Kapitel 5** kann als Argument für die Bedeutsamkeit des Einflusses von individuellen Unterschieden (Kapitel 5.1) und der Gedächtnisbeanspruchung (Kapitel 5.2) auf die kognitiven Prozesse während der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen verstanden werden. In Verbindung mit dem Einfluss individueller Unterschiede auf die Strategiewahl bei der Beurteilung kausaler Zusammenhänge vermuten Shaklee & Hall (1983), dass die von Ihnen berichteten Geschlechterunterschiede auf unterschiedlichen mathematischen Fertigkeiten beruhen könnten. Eine zu hohe Gedächtnisbeanspruchung kann entweder dazu führen, dass (i) fehlerhafte und/oder verzerrte

Kausalurteile gebildet werden (Fiedler, 2000; Kao & Wasserman, 1993; Shaklee & Mims, 1982), oder dazu, dass (ii) die Probanden einfachere und ungenauere Strategien einsetzen (Kleinmuntz & Schkade, 1993; Shaklee & Mims, 1982).

**Kapitel 6** befasst sich mit unterschiedlichen Theorien und Befunden zum Einfluss unterschiedlicher Medien auf die menschliche Informationsverarbeitung. In Kapitel 6.1 wird die ursprünglich im Bereich der Forschung zur Entscheidungsbildung entwickelte *Kosten-Nutzen-Theorie* von Kleinmuntz & Schkade (1993) vorgestellt. Nach der Kosten-Nutzen-Theorie können bei komplexen Entscheidungsprozessen und Aufgabenstellungen unterschiedliche Formen, Organisationen und Abfolgen von Information die Auswahl von Strategien beeinflussen. Es wird argumentiert, dass die Kosten-Nutzen-Theorie auch innerhalb der Forschung zur kausalen Urteilsbildung Anwendung finden könnte, um dort Erklärungen für einen möglichen Einfluss von Informationsanordnungen, insbesondere für mögliche Einflüsse der medienbezogenen Faktoren Darbietungsart, Präsentationsform und Informationsart, auf die Strategiewahl bei der Beurteilung der Stärke kausaler Zusammenhänge liefern zu können. In Kapitel 6.2 wird zu einer besseren Konzeptualisierung des differenziellen Einflusses unterschiedlicher Medien die Theorie von Zhang (Zhang, 1997, 2000; Zhang, Johnson, Malin & Smith, 2002; Zhang & Norman, 1994) herangezogen, in welcher die Auswirkungen externaler Repräsentationen auf die menschliche Informationsverarbeitung beschrieben werden. Danach werden in Kapitel 6.3 und 6.4 die Eigenschaften und die kognitive Verarbeitung derjenigen Medien beschrieben und miteinander verglichen, die in den empirischen Untersuchungen dieser Arbeit eingesetzt wurden. Bisherige Studien, in denen der Einfluss unterschiedlicher Medien in unterschiedlichen Aufgabenbereichen untersucht wurde, zeigen, dass Diagramme Tabellen vorzuziehen sind, wenn durch den Einsatz des Mediums Gruppierungen sichtbar gemacht, Vergleiche nahe gelegt und quantitative Beziehungen hergestellt werden sollen (s. Jacobs, 1994, 1999). Andere Studien belegen aber auch, dass die Wirksamkeit einer bestimmten Präsentationsform entscheidend von der auszuführenden Aufgabe abhängt (Wickens & Carswell, 1995; Vessey, 1991, 1994).

In **Kapitel 7** werden die wesentlichen Erkenntnisse aus den Kapiteln 3 bis 6 zusammengefasst und es wird daraus die zentrale These abgeleitet, dass die kognitiven Prozesse, die bei der Beurteilung kausaler Zusammenhänge ablaufen, von einem Zusammenspiel von aufgaben- und medienbezogenen Faktoren, sowie von individuellen Unterschieden und Voraussetzungen beeinflusst werden (Kapitel 7.1). Entsprechend ergibt sich daraus die

Forderung, dass es nicht genügt, die bei der Beurteilung der Stärke von Kausalzusammenhängen ablaufenden Prozesse nur auf der Ebene der Kognition zu analysieren, sondern auch auf einer Aufgaben-, einer Medien- und einer Individuenebene (Kapitel 7.2). In Kapitel 7.3 werden auf der Basis vorliegender Forschungsergebnisse für jede dieser Ebenen Bedingungen spezifiziert, unter denen regel-basierte Strategien mit einer höheren Wahrscheinlichkeit zum Einsatz kommen. Kapitel 7.4 beschließt den Theorieteil mit einer Auflistung der Fragestellungen, die in drei Experimenten eingehender untersucht wurden.

## **Teil 2: Empirische Untersuchungen**

In den Kapiteln 8, 9 und 10 werden die drei durchgeführten Experimente detailliert beschrieben und die gefundenen Ergebnisse diskutiert. Die abhängigen Variablen waren in allen drei Experimenten gleich: Gemessen wurden die *Anzahl regel-konformer Lösungen pro Proband*, die *Bearbeitungszeit*, die *Verarbeitungseffizienz* und die *Strategiewahl*. Mit Hilfe der *Anzahl regel-konformer Lösungen pro Proband* und der *Strategiewahl* sollte erfasst werden, ob die Probanden regel-basiert vorgegangen sind oder nicht. Weitere Analysen konnten darüber hinaus zeigen, dass eine hohe Anzahl regel-konformer Lösungen auf eine geringe mentale Beanspruchung während der Aufgabenbearbeitung hinweist. Durch die *Bearbeitungszeit* konnte der zeitliche Aufwand bei der Aufgabenbearbeitung erfasst werden. Die *Verarbeitungseffizienz* pro Person wurde aus dem Quotienten *Anzahl regel-konformer Lösungen / Bearbeitungszeit* ermittelt, aus einer hohen Anzahl regel-konformer Lösungen und einer geringen Bearbeitungszeit ergab sich also eine hohe Verarbeitungseffizienz. Mit Hilfe der *Verarbeitungseffizienz* sollte ein Maß bereitgestellt werden, welches Aufschluss darüber geben sollte, wie gut die jeweils untersuchte Bedingungskonstellation dazu geeignet ist, regel-konforme Lösungen zu ermöglichen. Lösungen wurden in den drei Untersuchungen dann als regel-konform eingestuft, wenn sie exakt mit den Vorhersagen der  $\Delta P$ -Regel (Allan, 1993; Ward & Jenkins, 1965) oder einer Berechnung der kausalen Stärke  $p$  nach der Theorie der kausalen Power von Cheng (1997) übereinstimmten.

**Kapitel 8** beginnt den Empirieteil der vorliegenden Arbeit mit einer Beschreibung von Experiment 1. In Experiment 1 wurde der Einfluss der beiden Präsentationsformen Stapeldiagramm und tabellarische Vierfeldertafel, sowie der beiden Inhaltsarten Häufigkeiten und Wahrscheinlichkeiten untersucht, darüber hinaus wurden die sich aus einer Kreuzung der Faktoren *Präsentationsform* und *Informationsart* ergebenden vier Experimentalbedingungen

mit einer Protokollliste (Kontrollbedingung) verglichen. Es zeigte sich, dass unter den beiden Wahrscheinlichkeitsbedingungen tendenziell mehr regel-konforme Lösungen gebildet wurden und mehr Probanden regel-basierte Strategien einsetzten. Allerdings hatte der Faktor *Informationsart* keine Auswirkungen auf die Bearbeitungszeiten und die Verarbeitungseffizienz. Demgegenüber ergab sich ein Trend, dass Stapeldiagramme zwar schneller verarbeitet werden konnten als tabellarische Vierfeldertafeln, doch hatte die Unterscheidung zwischen Stapeldiagrammen und tabellarischer Vierfeldertafel keinerlei Einfluss auf die Anzahl regel-konformer Lösungen und die Strategiewahl. Diese Ergebnisse entsprachen zwar teilweise den formulierten Hypothesen, doch fielen sie schwächer aus als erwartet. Die schwachen Effekte lassen sich möglicherweise durch Deckeneffekte erklären (knapp 61 % aller Probanden setzten durchgängig regel-basierte Strategien ein, knapp 82 % zumindest teilweise), denkbar ist aber auch, dass sich die vier Experimentalbedingungen nur unwesentlich hinsichtlich ihres Potenzials unterscheiden, die mentale Beanspruchung zu reduzieren und die Probanden zum Gebrauch regel-basierter Strategien anzuregen. Gemeinsam ist allen Bedingungen in Experiment 1, dass die Informationen in summarischer Form dargeboten wurden. Immerhin wurden die vier Experimentalbedingungen im Vergleich zur Kontrollbedingung Protokollliste schneller und effizienter verarbeitet, außerdem wurden mit Ausnahme der Tabelle mit Häufigkeiten mehr regel-konforme Lösungen gebildet und mehr regel-basierte Strategien eingesetzt. Diese Ergebnisse entsprachen im Wesentlichen den Erwartungen und wurden von uns auf eine geringere mentale Beanspruchung und eine größere Übersichtlichkeit der Informationsanordnung in den beiden Diagrammbedingungen und der Bedingung Tabelle mit Wahrscheinlichkeiten zurückgeführt. Experiment 1 untersuchte darüber hinaus den Einfluss des Faktors *Studienausrichtung* auf die abhängigen Variablen, allerdings unterschieden sich naturwissenschaftlich orientierte Probanden nicht signifikant von geisteswissenschaftlich orientierten Probanden. Möglicherweise wurden auch hier die vermuteten Deckeneffekte bedeutsam. Da eine Nachbefragung der Probanden keine Hinweise darauf ergab, worin sich Probanden, die regel-basiert vorgehen, von solchen unterscheiden, die nicht regel-basiert vorgehen, wurde in den Experimenten 2 und 3 auf die Nachbefragung verzichtet. Weitere Häufigkeits- und Teilstichprobenanalysen zeigten aber, dass in Experiment 1 ausgeprägte individuelle Unterschiede zu beobachten waren. Deshalb haben wir uns auf der Basis eigener Annahmen und der Vermutung von Shaklee & Hall (1983), dass unterschiedliche mathematische Fertigkeiten bei der Beurteilung der Stärke von Kausalzusammenhängen bedeutsam sein könnten, dazu entschlossen, in den Experimenten 2 und 3 den Einfluss individueller Unterschiede im numerischen Denken mit zu untersuchen.

In **Kapitel 9** wird über Experiment 2 berichtet. Darin wurde der Einfluss der beiden Darbietungsarten einzelfallweise und summarische Informationsdarbietung, sowie einem hohen und einem geringen Ausmaß an visueller Strukturierung des darstellenden Mediums untersucht. Visuell hoch strukturierte Grafiken unterscheiden sich von visuell gering strukturierten einfachen Fliesstexten darin, dass in ihnen die Informationen über Ursache-Wirkungs-Zusammenhänge übersichtlicher angeordnet sind. Zudem machen visuell hoch strukturierte Grafiken im Gegensatz zu einfachen Fliesstexten Vergleichs- und Inferenzprozesse salient, die die Probanden beim Einsatz regel-basierter Strategien und bei der Bildung regel-konformer Lösungen berücksichtigen müssen. Aus der Kreuzung der beiden Faktoren *Darbietungsart* und *Visuelle Strukturierung* ergaben sich die vier Bedingungen *Stapeldiagramm mit summarischer Darbietung*, *Fliesstext mit summarischer Darbietung*, *Vierfeldertafel mit Positionskreuz und einzelfallweiser Darbietung*, sowie *Fliesstext mit einzelfallweiser Darbietung*. Außerdem wurden noch der aus Experiment 1 bekannte Faktor *Studienausrichtung* und der Einfluss der Kovariate *numerisches Denken* untersucht. Erwartungsgemäß bildeten die Probanden bei summarischer Darbietungsart und bei visuell hoch strukturierten Grafiken eine höhere Anzahl regel-konformer Lösungen und setzten häufiger regel-basierte Strategien ein. Während für den Einfluss des Faktors *visuelle Strukturierung* nur ein Trend erzielt wurde, war der Einfluss des Faktors *Darbietungsart* hoch signifikant und überlagerte den Einfluss des Faktors *visuelle Strukturierung*. Entsprechend wirkte sich der Faktor *visuelle Strukturierung* nur bei summarischer Informationsdarbietung auf die Anzahl regel-konformer Lösungen und die Strategiewahl aus. Des Weiteren arbeiteten die Probanden bei summarischer Informationsdarbietung signifikant schneller und effizienter als bei einzelfallweiser Darbietung. In Verbindung mit dem Einfluss des Faktors *visuelle Strukturierung* ergab sich für die Bearbeitungszeiten eine signifikante Interaktion mit dem Faktor *Darbietungsart*, da Informationen in einer visuell hoch strukturierten und übersichtlichen Grafik nur bei einzelfallweiser und nicht bei summarischer Darbietung schneller verarbeitet werden konnten. Vermutet wurde, dass sich die kürzeren Bearbeitungszeiten der Grafik im Vergleich zum Fliesstext bei einzelfallweiser Darbietung vor allem auf eine höhere Geschwindigkeit von Prozessen während der Informationsnahme zurückführen lassen, die bei summarischer Darbietung deutlich weniger ins Gewicht fallen, da hierbei das zeitaufwändige Sammeln und Ordnen von Einzelfallinformationen entfällt. Die unterschiedlichen Auswirkungen des Faktors *visuelle Strukturierung* in Abhängigkeit von der *Darbietungsart* veranschaulichen auch, wie sich durch eine Variation der Darbietungsart die Beschaffenheit der Aufgabe, die die Probanden bei der Beurteilung der Stärke von Ursache-



Wirkungs-Zusammenhängen auszuführen haben, verändert. Die Ergebnisse aus Experiment 2 zeigen, dass sich mediale Rahmenbedingungen dann sehr stark auf die kognitiven Prozesse bei der Beurteilung der Stärke von kausalen Zusammenhängen auswirken können, wenn durch sie Aufgabenbeschaffenheit und Komplexität der Aufgabe entscheidend verändert werden. Dabei scheinen sich mediale Faktoren nicht nur auf die Gedächtnisbeanspruchung auszuwirken, sondern auch auf Wahrnehmungsorganisation, Handlungsabläufe und Prozesse des Schlussfolgerns. In Verbindung mit dem Einfluss des Faktors *Studienausrichtung* und der Kovariate *numerisches Denken* lassen die Ergebnisse aus Experiment 2 den Schluss zu, dass regel-basierte Strategien vor allem von Probanden eingesetzt werden, die ein mathematisch-naturwissenschaftlich ausgerichtetes Studienfach absolvieren und/oder über eine hohe numerische Verarbeitungskapazität verfügen. Damit wird auch die Annahme gestützt, dass individuelle Unterschiede im numerischen Denken von hoher Bedeutung für die Prozesse bei der Beurteilung der Stärke kausaler Zusammenhänge sein können.

In **Kapitel 10** wird zum Abschluss des empirischen Teils dieser Arbeit Experiment 3 vorgestellt. Darin wurden die vier Bedingungen *summarische Darbietung ohne Notizen*, *einzelfallweise Darbietung mit strukturierten Notizen*, *einzelfallweise Darbietung mit freien Notizen* und *einzelfallweise Darbietung ohne Notizen* hinsichtlich ihrer Auswirkungen auf die abhängigen Variablen, insbesondere die Anzahl regel-konformer Lösungen und die Strategiewahl, miteinander verglichen. Durch das Anfertigen von Notizen sollte die mentale Beanspruchung der Probanden reduziert und das Arbeitsgedächtnis der Probanden entlastet werden. Zudem sollte eine Antwort auf die Frage gefunden werden, ob sich die in Experiment 2 gefundene höhere Anzahl regel-konformer Lösungen bei summarischer Darbietung ausschließlich auf eine reduzierte Beanspruchung des Arbeitsgedächtnisses zurückführen lässt, oder ob dabei auch andere Faktoren, beispielsweise die Salienz oder Übersichtlichkeit der Informationsanordnung, eine Rolle spielen. Vermutet wurde, dass eine durch das Anfertigen von Notizen reduzierte Beanspruchung des Arbeitsgedächtnisses bei einzelfallweiser Darbietung insbesondere dann zu einer ähnlich hohen mittleren Anzahl regel-konformer Lösungen führt wie bei summarischer Darbietung, wenn die Probanden dazu angeleitet werden, ihre Notizen strukturiert in eine vorgegebene und übersichtlich gestaltete Vierfeldertafel einzutragen. Die frei oder strukturiert angefertigten Notizen sollten so als externale Repräsentationen dienen, die es den Probanden ermöglichen, interne Gedächtnisprozesse zu externalisieren und sie für Denkopoperationen verfügbar zu halten, die zur Beurteilung der Stärke kausaler Zusammenhänge erforderlich sind. Die Ergebnisse zeigten, dass die Anzahl regel-konformer Lösungen in den beiden Bedingungen mit Notizen

höher ausgefallen sind als in der Einzelfallbedingung ohne Notizen und geringer als in der summarischen Bedingung ohne Notizen. Obwohl die entsprechenden Einzelkontraste nicht signifikant wurden, haben wir die Unterschiede dennoch interpretiert, da sie von uns teilweise vorhergesagt worden waren. Der Unterschied zwischen den Einzelfallbedingungen mit und ohne Notizen, lässt sich nach unserer Auffassung auf eine durch das Anfertigen von Notizen reduzierte Beanspruchung des Arbeitsgedächtnisses zurückführen. Unklar ist allerdings, warum sich die beiden Bedingungen mit Notizen nicht oder kaum unterscheiden und warum in den beiden Bedingungen mit Notizen weniger regel-konforme Lösungen gebildet wurden als in der summarischen Bedingung ohne Notizen.

Der fehlende Unterschied zwischen den beiden Bedingungen mit Notizen kann post hoc damit erklärt werden, dass die frei erstellten Notizen für die Probanden nicht weniger übersichtlich waren als das strukturierte Eintragen von Strichen in eine vorgegebene Vierfeldertafel. Immerhin haben knapp 2/3 der Probanden in der Bedingung *Notizen frei* eigenständig eine Vierfeldertafel erstellt. Möglicherweise haben die Probanden in der Bedingung Notizen frei aber auch davon profitiert, dass das eigenständige Anfertigen von Notizen zu einer intensiveren, eingehenden Beschäftigung mit den selbst erstellten Notizen geführt hat und die Auswirkungen der möglicherweise fehlenden Übersichtlichkeit der freien im Vergleich zu den strukturierten Notizen neutralisiert hat.

Der Unterschied zwischen den beiden Notizenbedingungen und der summarischen Bedingung ohne Notizen scheint aus unserer Sicht darin begründet zu sein, dass die mentale Beanspruchung in den beiden Einzelfallbedingungen mit Notizen immer noch deutlich erhöht ist, weil die Probanden die zu verarbeitenden Informationen erst im Anschluss an eine einzelfallweise Informationsdarbietung und nicht schon von Anfang an als Häufigkeiten vorliegen hatten. So ist bei einer einzelfallweisen Darbietung mit Notizen die Aufgabenbeschaffenheit immer noch eine andere als in der summarischen Bedingung ohne Notizen, da in den beiden Einzelfallbedingungen mit Notizen viel Aufmerksamkeit für einfache, aber zeitaufwendige Handlungsrouninen (Ordnen und Sammeln von Daten, Notizen anfertigen) benötigt wird. Möglicherweise fehlt den Probanden dadurch der Überblick, um die zur Bildung regel-konformer Lösungen notwendigen Vergleichs- und Inferenzprozesse durchführen zu können. Denkbar ist auch, dass die Probanden sich aus Kosten-Nutzen-Erwägungen dafür entschieden haben, auf weniger komplexe Strategien zur Beurteilung der Stärke eines kausalen Zusammenhangs zurückzugreifen. Letzteres spräche dafür, dass bei der Beurteilung der Stärke von monokausalen Zusammenhängen Entscheidungsbildungsprozesse mitbeteiligt sind und die Probanden im Sinne der Kosten-Nutzen-Theorie von Kleinmuntz &

Schkade (1993) dazu angeregt werden, bei höherer Komplexität von Informationsanordnung und/oder Aufgabenbeschaffenheit von einfacheren Heuristiken und Strategien Gebrauch zu machen und entsprechend weniger regel-konforme Lösungen zu bilden.

Ferner wurde in Experiment 3 der Einfluss von Studienausrichtung und numerischem Denken erfasst. Während die Kovariate *numerisches Denken* erneut hochsignifikant war, hatte der Faktor *Studienausrichtung* dieses Mal nur einen trendnahen Einfluss auf die Anzahl regel-konformer Lösungen, da die Probanden mit naturwissenschaftlicher Ausrichtung in der Bedingung *Notizen strukturiert* weniger regel-konforme Lösungen produzierten als die geisteswissenschaftlich ausgerichteten Studenten. Weitere Analysen konnten zum einen zeigen, dass sich die Probanden unterschiedlicher Studienausrichtung im Gegensatz zu den anderen drei Bedingungen in der Bedingung *Notizen strukturiert* nur geringfügig im numerischen Denken unterschieden. Zum anderen konnte aber auch nachgewiesen werden, dass in den beiden Notizenbedingungen der Einfluss des numerischen Denkens schwächer ausgeprägt war als in den beiden Bedingungen ohne Notizen. Möglicherweise führt die Entlastung des Arbeitsgedächtnisses in den beiden Bedingungen mit Notizen dazu, dass andere individuelle Voraussetzungen, etwa die Motivation der Probanden, ein im Vergleich zu den Fertigkeiten im numerischen Denken höheres Gewicht erhalten.

### **Teil 3: Schlussfolgerungen**

In **Kapitel 11** werden die wesentlichen Ergebnisse des Theorieteils und der empirischen Untersuchungen noch einmal zusammengefasst, insbesondere werden die Ergebnisse explizit auf die am Ende des Theorieteils formulierten Forschungsfragen bezogen. Im Wesentlichen stützen die gefundenen Ergebnisse die Behauptung, dass die Häufigkeit des Einsatzes regel-basierter Strategien bei der Beurteilung der Stärke von monokausalen Ursache-Wirkungs-Zusammenhängen entscheidend von der jeweiligen Zusammensetzung von aufgaben-, medien- und personenbezogenen Faktoren in einer Untersuchung abhängt.

**Kapitel 12** beschließt die Arbeit mit einem Ausblick auf mögliche Folgeuntersuchungen und weitere Forschungsfelder im Zusammenhang mit dem Erwerb von Kausalwissen und der kausalen Urteilsbildung.



Teil 1

# **Theoretischer Hintergrund**



# 1 Einführung

In dieser Arbeit geht es um die Untersuchung des Einflusses unterschiedlicher Informationsanordnungen auf die Beurteilung der Stärke monokausaler Zusammenhänge. In Verbindung mit Kausalität interessiert man sich innerhalb der kognitiven Psychologie unter anderem dafür, wie Menschen Informationen über kausale Zusammenhänge verarbeiten. Zentral ist dabei die Fragestellung, wie Menschen Wissen über Ursache-Wirkungs-Zusammenhänge erwerben oder erlernen bzw. wie sie kausale Urteile bilden. In der Literatur finden sich dazu die Schlagwörter *Erwerb von Kausalwissen* (Hagmayer & Waldmann, 2006), *kausales Lernen* (Hagmayer & Waldmann, 2006; englisch: *causal learning*, s. Dickinson, 2001; Shanks, Holyoak & Medin, 1996), *kausale Urteilsbildung* (englisch: *causal judgment*, s. Perales & Shanks, 2003, in press; Shanks, 1991; White, 2004); *Kontingenzlernen* (englisch: *contingency learning*, s. De Houwer & Beckers, 2002; Pineño & Miller, 2007; Shanks, 2007), *Kontingenzurteile* (englisch: *contingency judgment*, s. Allan, 1993; Kao & Wasserman, 1993; Price & Yates, 1995, Shanks, 1987; Shanks, López, Darby & Dickinson, 1996); *Kausalinduktion* (englisch: *causal induction*, s. Baker, Murphy & Vallée-Tourangeau, 1996; Buehner & Cheng, 1997; Cheng & Novick, 1990, 1992; Collins & Shanks, 2006; Griffiths & Tenenbaum, 2005; Lober & Shanks, 2000; Waldmann, 1996) und *Kovariationsurteil* (englisch: *covariation judgment*, s. Shaklee, 1983).

In Verbindung mit diesen Begriffen wurde vor allem die Frage kontrovers diskutiert, ob sich die Prozesse bei der kausalen Urteilsbildung besser durch regel-basierte Ansätze oder durch assoziative Lernmechanismen erklären oder vorhersagen lassen (Allan, 1993; Baker, Murphy & Vallée-Tourangeau, 1996; Perales & Shanks, in press). Regel-basierte Ansätze gehen von der Annahme aus, dass Menschen Wissen über Kausalzusammenhänge erwerben, indem sie kognitive Operationen anwenden, die rationalen Vorgehensweisen ähneln. Eine schwächere Variante der Grundannahme regel-basierter Ansätze geht lediglich davon aus, dass das Resultat der beim Erwerb von Kausalwissen ablaufenden psychologischen Prozesse mit dem Ergebnis einer rationalen oder normativen Vorgehensweise übereinstimmt, unabhängig von den tatsächlich ablaufenden Operationen, mit deren Hilfe der Lernende zu dieser Übereinstimmung gelangt (Perales & Shanks, in press). In Experimenten zur Untersuchung der Frage, ob der Erwerb von Kausalwissen regel-basiert oder assoziativ abläuft werden die Probanden sehr häufig vor die Aufgabe gestellt, die Stärke eines Ursache-Wirkungs-Zusammenhangs quantitativ einzuschätzen.

Ausgangspunkt der vorliegenden Arbeit war die Fragestellung, unter welchen medialen Rahmenbedingungen Probanden bei der Beurteilung der Stärke monokausaler Zusammenhänge regel-basiert vorgehen. Dabei ist die Annahme zentral, dass es in Verbindung mit der Beurteilung der Stärke monokausaler Zusammenhänge nicht genügt, die dabei ablaufenden kognitiven Prozesse und Mechanismen zu untersuchen, sondern dass die Einflüsse der Informationsumgebung mit berücksichtigt werden müssen. Innerhalb der Literatur zum kausalen Lernen und zur kausalen Urteilsbildung wurden der Einfluss der Informationsumgebung eher vernachlässigt, vor allem wurden nur vereinzelt und am Rande Theorien (z.B. Oestermeier & Hesse, 2000; Shanks, 1991) formuliert, in denen die Prozesse bei der kausalen Urteilsbildung mehr (Oestermeier & Hesse, 2000) oder weniger explizit (Shanks, 1991) als eine dynamische Interaktion von Individuum und Informationsumgebung konzipiert wurden. Bei näherer Betrachtung der Literatur zeigt es sich zwar, dass unterschiedliche *Aufgabendimensionen* (Allan, 1993, S. 435) die Prozesse bei der kausalen Urteilsbildung beeinflussen können (s. Allan, 1993), doch erst in jüngerer Zeit wurde beispielsweise der Einfluss von Fragestellung und dabei verwendeter Wortwahl intensiver untersucht (Buehner, Cheng & Clifford, 2003; Collins & Shanks, 2006; Shanks, 2004; White, 2003b). Die Auswahl der in den unterschiedlichen Untersuchungen zur kausalen Urteilsbildung eingesetzten Anordnung von Informationen über Ursache-Wirkungs-Zusammenhänge erfolgte eher willkürlich und nicht theoriegeleitet.

Ziel dieser Arbeit war es, den Einfluss unterschiedlicher Informationsanordnungen auf die Beurteilung der Stärke monokausaler Zusammenhänge zu untersuchen. Im Besonderen sollte dabei - unter Konstanthaltung des Einflusses aufgabenbezogener Faktoren wie z.B. der Fragestellung - der Einfluss von unterschiedlichen Präsentationsformen, Informations- und Darbietungsarten auf die Beurteilung der Stärke von monokausalen Zusammenhängen untersucht werden. Darüber hinaus sollte herausgefunden werden, ob sich Individuen hinsichtlich Urteilsverhalten und Strategiewahl unter verschiedenen medialen Bedingungen voneinander unterscheiden und ob unterschiedliche individuelle Voraussetzungen, insbesondere die numerische Verarbeitungskapazität der Probanden, einen Einfluss darauf haben, ob Probanden regel-basiert vorgehen oder nicht.



## **2 Philosophische Sichtweisen zur Kausalität und ihr Einfluss auf die kognitionspsychologische Theorienbildung zum kausalen Lernen**

Im Allgemeinen versteht man unter Kausalität einen Zusammenhang zwischen Objekten, Ereignissen, Prozessen, Sachverhalten usw. bei dem eine Erscheinung A (Ursache genannt) unter bestimmten Bedingungen eine bestimmte andere Erscheinung B (Wirkung genannt) mit Notwendigkeit hervorruft. Die Begriffe Kausalität, Kausalzusammenhang oder Ursache-Wirkungs-Zusammenhang sollen dabei synonym verwendet werden. In der philosophischen Literatur zur Kausalität nimmt die Bestimmung von Begrifflichkeiten eine zentrale Bedeutung ein, sie ist insofern vorwiegend normativ ausgerichtet (Van der Meer, 2003; Oestermeier, 1997). Psychologische Ansätze interessieren sich dagegen eher deskriptiv dafür, wie kausales Denken im Alltag abläuft und wie Menschen Informationen über Kausalzusammenhänge verarbeiten. Während in der Psychologie wichtige philosophische Theorien zur Kausalität rezipiert wurden, wurden psychologische Beiträge, etwa zum Verstehen kausaler Zusammenhänge bei Kindern (Bullock et al., 1982; Leslie, 1982, 1995; Leslie & Keeble, 1987; Sodian, 2002; Spelke, Phillips & Woodward, 1995), von der Philosophie weitgehend ignoriert (Oestermeier, 1997). Nachfolgend sollen aus philosophischer Sicht traditionelle (Aristoteles, Hume, Kant) und moderne (bedingungsanalytische und probabilistische) Theorien zur Kausalität vorgestellt werden. Die Rezeption dieser philosophischen Ansätze von psychologischer Seite beschließt die jeweiligen Abschnitte (für einen Überblick s. Oestermeier, 1997, sowie Hagmayer & Waldmann, 2006).

### **2.1 Traditionelles Kausalitätsverständnis in der Philosophie: Aristoteles, Hume und Kant**

Philosophiegeschichtlich sind neben dem aristotelischen Kausalitätsverständnis vor allem die Arbeiten von David Hume und Immanuel Kant von herausragender Bedeutung. Zwar haben sich Philosophen sowohl vor als auch nach dem Erscheinen von Humes *Treatise of Human Nature* (1739) und Kants *Kritik der reinen Vernunft* (1781) mit dem Kausalitätsbegriff auseinandergesetzt, doch haben sie keinen vergleichbaren Einfluss auf die psychologische Theorienbildung im Bereich der Kausalkognition entfaltet.

Aristoteles (1970) formulierte in seiner *Metaphysik* ein Kausalitätsprinzip ("Alles Entstehende wird durch etwas und aus etwas ein Was" *Metaphysik*, Buch VII, 1032, 10-15) und unterschied vier Arten von Ursachen: Die *causa materialis* (materielle oder stoffliche Ursache), beinhaltet die Materie (z.B. Holz), aus der ein Ding oder ein Gegenstand (z.B. Tisch) zusammengesetzt ist. In der *causa formalis* (formale Ursache) ist die Form, das Wesen oder auch der Plan eines Gegenstandes oder Ereignisses fest gelegt (z.B. setzt sich ein Tisch aus einer runden oder eckigen Tischplatte mit vier gleichlangen, geraden Tischbeinen zusammen). Die *causa efficiens* (Antriebs- oder Wirkursache) beinhaltet die Kraft, den Antrieb, das Motiv, die Urheberschaft oder die Dynamik, wodurch ein Gegenstand entsteht oder ein Ereignis seinen Anfang nimmt (z.B. entsteht der Tisch durch die Arbeit eines Tischlers oder Schreiners). Die *causa finalis* (Zweckursache) schließlich umfasst den Zweck oder das Ziel, dem ein Gegenstand oder Ereignis dienen soll (z.B. dient der Tisch dem gemeinsamen Essen oder allgemeiner der Ablage und Anordnung von Gegenständen). Bunge (1987) hebt hervor, dass das neuzeitliche Denken nur eine der vier Ursachenarten des Aristoteles als Kausalität im eigentlichen Sinne behandelt, nämlich die *causa efficiens*.

### *David Hume*

Nach der so genannten Regularitätstheorie von David Hume (1739/1778 in *A Treatise of Human Nature*, s. vor allem Buch I, Teil III, und 1748/1775 *An Enquiry of Human Understanding*) beinhaltet der Begriff der Ursache dreierlei: (i) eine Aufeinanderfolge (Sukzession) von Ereignissen der Gestalt, dass die Ursache der Wirkung zeitlich vorausgeht, (ii) dass eine raum-zeitliche Nähe (englisch: *contiguity*) zwischen zwei Phänomenen besteht und (iii) dass das, was geschieht, notwendig geschieht (Notwendigkeit). Die Erkenntnis von Sukzession und raum-zeitlicher Nähe basiert jeweils auf sinnlicher Erfahrung, die Erkenntnis von Notwendigkeit nicht. Selbst wenn ein Ereignis A in unserer bisherigen Erfahrung konstant und ständig wiederholend ein Ereignis B hervorruft, können wir nicht wissen, dass diese konstante Verknüpfung (englisch: *constant conjunction*) notwendig ist, da wir keine hundertprozentige Gewissheit darüber haben, ob diese Verknüpfung gesetzmäßig ist und auch in Zukunft immer Bestand haben wird. Wenn die Erkenntnis von Notwendigkeit nicht aus der Erfahrung stammen kann, so wirft das die Frage auf, wie dann dennoch die Vorstellung von der notwendigen Kausalverknüpfung zweier Ereignisse in uns entstehen kann. Die Antwort des Skeptikers Hume darauf ist "halb psychologisch" (Skirbekk & Gilje, 1993): Wenn Geschehnisse oder Ereignisse immer wieder auf die gleiche Weise assoziativ miteinander verknüpft sind, gewöhnen wir uns an diese Verknüpfungen und bilden Erwartungen in uns, dass sie auch künftig fortbestehen werden. Und diese Erwartungen sind es, die in uns die

Vorstellung einer (subjektiv) notwendigen Kausalverknüpfung hervorrufen. Hume (1975) betont dabei, dass der Schluss - oder besser Übergang - von Ursache auf Wirkung nicht rational, d.h. nicht durch die Tätigkeit des Verstandes (englisch: *reason*), erklärbar sei, sondern bloß das Ergebnis angeborener und Instinkt gesteuerter Assoziationsmechanismen in unserer Einbildungskraft (englisch: *imagination*): Habe ich in der Vergangenheit wiederholt beobachtet, dass auf Ereignisse von der Art des Ereignisses A (oder von einer ähnlichen Art) ein Ereignis von der Art B (oder von einer ähnlichen Art) folgte, dann sind die Vorstellungen von der Folge dieser beiden Ereignisse in mir so fest assoziiert, dass ich mit meiner Einbildungskraft die Erwartung ausbilde, dass die beiden Ereignisarten auch in Zukunft regelmäßig aufeinander folgen werden. Ursache und Wirkung werden also von unserer Einbildungskraft durch ihre Aufeinanderfolge und ihre raumzeitliche Nähe, sowie durch ihre konstante Verknüpfung über einen basalen Assoziationsmechanismus zu einer Vorstellungseinheit geformt (Hume, 1978).

Hume hat wohl das moderne Verständnis von Kausalität wie kein anderer Philosoph geprägt. Innerhalb der kognitiven Psychologie wurde Humes Kausalitätsverständnis vor allem von Vertretern assoziationalistischer Ansätze aufgegriffen. Beispielsweise wurden Humes Prinzipien der Sukzession und der raum-zeitlichen Kontiguität von Ursache und Wirkung mehr oder weniger unverändert übernommen (Baker et al., 1996; Rescorla & Wagner, 1972). Seine Regularitätsannahme hat neben assoziationalistischen vor allem auch regel-basierte Ansätze beeinflusst.

### *Immanuel Kant*

Kant stimmt in seiner *Kritik der reinen Vernunft* (1781/1794) oder auch in den *Prolegomena zu einer jeden künftigen Metaphysik* (1783/1796) mit Hume darin überein, dass unmittelbare Erfahrung und Induktion keine Einsicht in etwas Notwendiges und Allgemeingültiges vermitteln. Und er leugnet auch weder die Existenz einer Einbildungskraft, die Kant als eine blinde und begriffsleere, aber unentbehrliche "Funktion der Seele" (*Kritik der reinen Vernunft* B 104) bezeichnet, noch die Tatsache, dass jede Erkenntnis mit der Erfahrung beginnt (Einleitung zur *Kritik der reinen Vernunft* B 1). Aber er zweifelt die These Humes an, dass Kausalität das Resultat von auf Gewöhnung beruhenden (psychischen) Assoziierungsprozessen sei und dadurch das Kausalitätsprinzip seine Allgemeingültigkeit verliert. Um die von Hume herabgesetzte Bedeutung der Vernunft zu rehabilitieren und um die Allgemeingültigkeit des Kausalitätsprinzips ("alles, was geschieht, hat seine Ursache") wiederherzustellen, bestreitet Kant die empiristische Grundannahme, dass alle Erkenntnis

letztlich von sinnlichen Wahrnehmungen herrührt und sich nach den Gegenständen der Außenwelt richtet. Stattdessen dreht Kant mit seiner so genannten "kopernikanischen Wende" in der Metaphysik den Spieß herum und fragt sich, welche subjektiven Voraussetzungen gegeben sein müssen, damit Erfahrung überhaupt möglich wird. In den Worten Kants bedeutet die kopernikanische Wende, dass sich die Anschauung nicht nach der Beschaffenheit der erfahrbaren Gegenstände richtet, sondern der Gegenstand als Objekt der Sinne richtet sich nach der Beschaffenheit unseres Anschauungsvermögens (*Vorrede zur Kritik der reinen Vernunft* B XVII). Kausalität ist für Kant keine Vorstellung, die aus der Erfahrung wiederholter Ereignisverknüpfungen abgeleitet wird, sondern eine von zwölf allgemeingültigen Verstandeskategorien oder Denkformen, welche die Erkenntnis von Ursache-Wirkungs-Zusammenhängen in unserer Erfahrungswelt überhaupt erst ermöglicht, indem begriffsleere, sinnliche Wahrnehmungen unter diese in unserem Verstand vorgegebene Begriffskategorie subsumiert und damit erkannt werden können. Präziser formuliert versteht Kant das Kausalprinzip "alles, was geschieht, hat seine Ursache" als eine synthetische Erkenntnis a priori. Synthetische Erkenntnisse a priori sind unabhängig von Erfahrung (daher a priori), doch ist in ihnen das logische Prädikat (z.B. "hat seine Ursache") nicht im logischen Subjekt (z.B. "alles, was geschieht") enthalten. Verstanden als eine synthetische Erkenntnis a priori ist das Kausalprinzip eine notwendige Voraussetzung in den Naturwissenschaften, ohne die naturwissenschaftliche Erkenntnis unmöglich ist (Skirbekk & Gilje, 1993).

Versucht man den Prozess der Erkenntnis kausaler Zusammenhänge im Sinne Kants funktionalistisch zu rekonstruieren, so scheint sich dieser in zwei Stufen zu vollziehen: In einem ersten Schritt werden Erscheinungen (der für uns nicht erkennbaren Dinge an sich) in der blinden Einbildungskraft begriffslos miteinander verknüpft oder assoziiert. Das vorläufige Produkt dieser assoziativen Prozesse sind Anschauungen, die der Sinnlichkeit angehören und als solche Gegenstand der empirischen Psychologie sind (*Prolegomena*, § 21). Anschauungen können grob als unstrukturierte Sinnesempfindungen konzipiert werden (Höffe, 1996). In einem zweiten Schritt werden die begriffsleeren sinnlichen Wahrnehmungen unter die a priori gegebenen Begriffskategorien des Verstandes subsumiert oder auf sie bezogen. Erst durch die Tätigkeit des Verstandes können die unstrukturierten Sinnesempfindungen oder Anschauungen als Gegenstand gedacht und zu Erfahrungsurteilen verknüpft werden. Anders als bei Hume, wo der Verstand durch Gewohnheit zu dem Trugschluss der Notwendigkeit von Kausalurteilen verleitet wird, sind Kausalurteile bei Kant allgemeingültig, weil die Erfahrung von Kausalität erst durch eine notwendig in unserem Verstand angelegte und a priori gegebene Kategorie der Kausalität möglich gemacht werden

kann. Kant behauptet damit nicht, dass die durch unseren Verstand gebildeten Kausalurteile etwas über Verhältnisse in der Welt der Dinge an sich aussagen, weil sich diese jenseits unserer Erfahrungsgrenzen befinden und der Verstand nur die für uns wahrnehmbaren Erscheinungen der Dinge an sich unter seine Kategorien fassen kann. Die von Kant postulierte Allgemeingültigkeit von synthetischen Urteilen a priori wie z.B. dem Kausalprinzip liegt also nicht in den Dingen an sich, sondern in der allgemeinen Beschaffenheit des Verstandes begründet.

Kants Kausalitätsverständnis hat kognitionspsychologische Überlegungen zum Erlernen von Ursache-Wirkungs-Zusammenhängen und zum Erwerb von Kausalwissen ebenfalls stark beeinflusst, allerdings oft weniger explizit als dasjenige David Humes. Ein Grund hierfür mag seine idealistische empirieferne Konzeption sein, ein anderer, dass seine Begriffe erst in psychologische Termini und Theorien übersetzt werden müssen (Hagmayer & Waldmann, 2006). Seine Vorstellung, dass der Begriff der Kausalität notwendig für das Erkennen von Kausalität ist, ist grundlegend für die Annahme, dass Vorwissen über Kausalität die Wahrnehmung und die Beurteilung von Zusammenhängen als kausal beeinflusst.

Ein Beispiel für einen durch Kant inspirierten psychologischen Ansatz zum Erlernen von Ursache-Wirkungs-Zusammenhängen ist die von Waldmann und Mitarbeitern entwickelte *Theorie der Kausalmodelle* (Waldmann, 1994, 1996, 2000; Waldmann & Holyoak, 1992; Waldmann et al. 1995). Trotz vieler Kritikpunkte an der Kantischen Epistemologie wird Kants Grundannahme einer Interaktion zwischen gesammelten Erfahrungen und abstrakt gespeicherten Wissenskategorien (den Kausalmodellen), die es erst ermöglichen, den sensorischen Input zu interpretieren, von der Theorie der Kausalmodelle übernommen.

Der so genannte *Power-Ansatz* (Ahn et al. 1995; Michotte, 1963; White, 1989) zum kausalen Lernen wurde ebenfalls mit dem Kantischen Kausalitätsverständnis in Verbindung gebracht und von Cheng in die von ihr entwickelte Power PC Theorie (Cheng, 1997) integriert. Die kausale Kraft (power) wird dabei als ein intuitiver Begriff dafür verstanden, dass ein Ereignis durch irgendeine Kraft oder Energie ein anderes verursacht. Entsprechend dem Power-Ansatz treten Ursachen nicht einfach zeitlich vor ihren Effekten auf, sondern sie bringen diese Effekte irgendwie hervor oder erzeugen sie.

#### *Zur Ausschließlichkeit der unterschiedlichen Konzeptionen des Kausalitätsbegriffs nach Hume und Kant*

Oft wird der Gegensatz der Auffassungen von Hume und Kant in den Vordergrund gestellt. Nach Stegmüller (1983) sollte dabei aber nicht außer Acht gelassen werden, dass beide von

unterschiedlichen Problemstellungen ausgingen. Während Hume den Begriff der Ursache und den singulärer Kausalbehauptungen klären wollte, ging es Kant um das allgemeine Kausalprinzip und dessen Begründung. Wissenschaftstheoretisch betrachtet, blieb dabei der Einfluss Kants geringer als der Humes, was hauptsächlich daran liegt, dass die meisten Wissenschaftstheoretiker die kantische Herleitung der Verstandeskategorien für "vollkommen unzulänglich" (Stegmüller, 1983, S. 518) halten.

Auch aus kognitionspsychologischer Perspektive müssen Hume und Kant nicht als wechselseitig ausschließend betrachtet werden, sondern können – bei aller gerechtfertigten inhaltlichen Kritik – als wechselseitig ergänzend angesehen werden (siehe auch Cheng et al., 1996). Aus Sicht der Forschung zur Kausalkognition können Humes Ausführungen so verstanden werden, dass er einen Versuch gestartet hat, zu erklären, wie aus der wiederholten sinnlichen Wahrnehmung der konstanten Verknüpfung zweier aufeinander folgender raumzeitlich benachbarter Ereignisse eine mentale Repräsentation von einer Ursache-Wirkungs-Einheit oder, anders formuliert, ein kausales Schema entsteht. Kant hingegen setzt die Existenz einer allgemeinen mentalen Repräsentation von Kausalität oder eines allgemeinen kausalen Schemas bereits voraus, welches unsere sinnlichen Empfindungen strukturiert und dadurch die Erfahrung von Ursache-Wirkungs-Zusammenhängen möglich macht. Die beiden Sichtweisen Kants und Humes lassen sich kognitionspsychologisch miteinander in Einklang bringen, wenn sie in Analogie zu Marrs (1982) Unterscheidung zwischen einer algorithmischen und einer computationalen Ebene als unterschiedliche Ebenen der Analyse kognitiver Prozesse konzipiert werden. Während Hume unter Vernachlässigung der Rolle höherwertiger kognitiver Prozesse (in Humes Worten des Verstandes) auf das *Wie* der Entstehung von abstrakten Vorstellungen (englisch: *ideas*) im Allgemeinen und von kognitiven kausalen Schemata im Besonderen Bezug zu nehmen scheint, könnten die Ausführungen Kants kognitionspsychologisch als eine grobe funktionalistische Beschreibung (*was geschieht warum in welchen Teilschritten*) darüber verstanden werden, wie existierende kausale Schemata (ob angeboren oder erlernt) unstrukturierte Wahrnehmungen erfahrbar machen. Auf der Seite Kants müsste dabei die Prämisse fallen gelassen werden, dass es nur ein a priori gegebenes allgemeines Kausalitätsschema gibt. Stattdessen müsste von Anhängern Kants anerkannt werden, dass es viele assoziativ erworbene kausale Schemata gibt (Zugeständnis von erfahrungsabhängigen bottom-up Prozessen). Auf der Seite Humes müsste zugestanden werden, dass - unabhängig von der Frage, wie kausale Schemata entstehen - kausale Schemata existieren, die neue Erfahrungen ordnen, strukturieren und in

präexistierende neuronale Wissensnetze einbauen (Zugeständnis von top-down agierenden kausalen Schemata).

## 2.2 Neuere Konzeptionen des Kausalitätsbegriffs

In der Nachfolge Humes erlangten noch die Ausführungen von John Stewart Mill (1979) erkenntnistheoretische Bedeutung. Mill zufolge beruht unser Wissen über Ursache-Wirkungs-Zusammenhänge nicht auf der wiederholten Beobachtung der Ereignis- oder Objektfolge A B, sondern kann viel besser durch das Kovarianzprinzip verstanden werden. Nach dem Kovarianzprinzip sind vor allem solche Faktoren als kausal zu bezeichnen, die sowohl *hinreichend* (wenn A, dann B) als auch *notwendig* (wenn nicht A, dann nicht B) für die Wirkung sind. Zudem hat Mill darauf hingewiesen, dass in der Regel nicht nur ein, sondern mehrere Faktoren bestimmte Wirkungen zur Folge haben.

Im 20. Jahrhundert führten modernere philosophische und wissenschaftstheoretische Ansätze (z.B. Suppes, 1970; Mackie, 1974; Cartwright, 1989; Salmon, 1984; Bunge, 1987; Eells, 1991; Pearl, 2000; Meixner, 2001), sowie der Einfluss neuer wissenschaftlicher Erkenntnisse (z.B. physikalische Quantentheorie, Unbestimmtheitsprinzip von Heisenberg) zu einer zunehmenden Ausdifferenzierung und Facettenerweiterung des Kausalitätsbegriffs. Stegmüller (1983) zufolge lassen sich die neueren Arbeiten zur Kausalität dabei in zwei Arten von Abhandlungen unterteilen. Während es in den Arbeiten der ersten Art darum geht, die Humesche Theorie wesentlich zu *verbessern*, wird in der zweiten Art versucht, Humes Theorie wesentlich zu *verallgemeinern*. Exemplarisch für den Versuch einer Verbesserung von Humes Kausalitätsbegriffs ist der *bedingungsanalytische Ansatz* von John Leslie Mackie (1974). *Probabilistische Theorien zur Kausalität* (Suppes, 1970; Salmon, 1984; Eells, 1991) werden von Stegmüller als herausragender Versuch gesehen, Humes Theorie zu erweitern.

### *Der bedingungsanalytische Ansatz von Mackie*

Aufbauend auf Mills Kovarianzprinzip hat Mackie (1974) darauf hingewiesen, dass nicht alle alltagssprachlich als Ursachen bezeichneten Faktoren im strikten Sinne *hinreichend* und *notwendig* sind. So kann z.B. die singuläre Kausalaussage "Der Auto-Unfall wurde durch eine eisglatte Fahrbahn verursacht" wahr sein, obwohl die Ursache "eisglatte Fahrbahn" weder *hinreichend* (denn das Fahrzeug wäre bei langsamerer Fahrgeschwindigkeit, höheren Temperaturen oder gerader Streckenführung nicht ins Schleudern gekommen) noch *notwendig* (denn der Unfall hätte sich auch ohne Eisglätte durch ein entgegenkommendes Fahrzeug oder ausgelaufenes Öl auf der Fahrbahn ereignen können) für die Wirkung "Auto-

Unfall" ist. Aber wenn die eisglatte Fahrbahn weder hinreichend noch notwendig für den Unfall ist, in welchem Sinn hat sie denn dann den Unfall *verursacht*? Die Antwort liegt darin, dass ein komplexes Bedingungsgefüge K, welches sich z.B. aus zu schneller Fahrgeschwindigkeit, Kurve in einem Waldstück, Unaufmerksamkeit eines übermüdeten Fahrers, sinkenden Temperaturen unter den Gefrierpunkt und eisglatter Fahrbahnstelle infolge überfrierender Nässe in der Nacht zusammensetzt, den Unfall verursacht hat. *Alle* diese Bedingungen *zusammen* mögen in unserem Beispiel als *hinreichend* gelten und innerhalb dieses Bedingungskomplexes K ist die eisglatte Fahrbahn ein *unabdingbarer* Bestandteil, da die anderen Bedingungen allein nicht zu dem Unfall geführt hätten. Dabei ist aber zusätzlich zu berücksichtigen, dass der Bedingungskomplex K zwar hinreichend, aber nicht notwendig ist, da noch unzählige andere hinreichende Bedingungsgefüge vorstellbar sind, die den Unfall verursacht haben könnten. Die in unserem Beispiel als Ursache bezeichnete eisglatte Fahrbahn wird von Mackie als INUS-Bedingung charakterisiert, als ein *nicht hinreichender* (englisch: **insufficient**), aber *notwendiger* (englisch: **necessary**) Teil einer komplexen Bedingung K<sub>i</sub>, die selbst als ganze *nicht notwendig* (englisch: **unnecessary**), aber *hinreichend* (englisch: **sufficient**) für das fragliche Ereignis ist. Die eisglatte Fahrbahn ist aber nur dann als Ursache des zu erklärenden Autounfalls anzusehen, wenn auch die anderen Komponenten (zu schnelle Fahrgeschwindigkeit, Kurve in einem Waldstück, Unaufmerksamkeit eines übermüdeten Fahrers, Temperaturen unter dem Gefrierpunkt) der komplexen Bedingung K<sub>i</sub> gegeben sind; die Wirkung der Ursache muss also auf die jeweils gegebenen Umstände relativiert werden.

Zu beachten ist dabei auch die Abgrenzung der Ursachen von den dauerhaften Bedingungen. Dass wir in unserem Beispiel eine zu schnelle Fahrgeschwindigkeit oder die eisglatte Fahrbahn, nicht aber die Kurve in einem Waldstück oder die Temperaturen unter dem Gefrierpunkt als Ursachen ansehen, folgt nach Mackie zwar nicht strengen Auswahlregeln, unterliegt aber doch systematischen Tendenzen.

Das bisher Gesagte bezog sich auf singuläre Ereignisse, lässt sich aber auf die Analyse kausaler Regularitäten übertragen, indem von individuellen Ereignissen auf entsprechende Ereignistypen übergegangen wird. Wie bei singulären Kausalaussagen analysiert Mackie also kausale Regularitäten als komplexe Bedingungssätze und relativiert diese auf ein kausales Feld oder einen kausalen Kontext, wobei in der Alltagssprache meist nur ein Faktor aus dem ganzen und oft nicht vollständig erfassten Bedingungsgefüge genannt wird (z.B. "eisglatte Fahrbahn").



Obwohl Mackie selbst keinen psychologischen Erklärungsanspruch erhoben hat, wurde seine Analyse im Bereich der menschlichen Kausalkognition mehrfach als Beschreibung aufgegriffen, wie Menschen im Alltag etwas kausal erklären oder beurteilen (z.B. Einhorn & Hogarth, 1986).

### *Probabilistische Theorien der Kausalität*

Die zweite Art der Auseinandersetzung mit Hume, nämlich diejenige sein Werk zu verallgemeinern, kann mit dem Namen Paul Suppes (1970; s.a. Stegmüller, 1983, und Eells, 1991) und seiner *probabilistischen Theorie der Kausalität* verbunden werden. Suppes Kritik an Hume lässt sich besonders prägnant in dem Zitat zusammenfassen, dass "the omission of probability considerations [...] is perhaps the single greatest weakness in Hume's famous analysis of causality." (Suppes, 1970, S. 9; zit. n. Stegmüller, 1983, S.600). Die empirische Basis für diese Feststellung liegt in der Beobachtung, dass der Begriff der Kausalität alltagssprachlich meist nicht deterministisch, sondern probabilistisch verwendet wird. Humes Vorstellung einer konstanten, d.h. deterministischen Verknüpfung ist demnach durch den Gedanken an eine probabilistische Aufeinanderfolge zu ersetzen. Suppes zufolge ist ein Ereignis A dann die Ursache eines anderen Ereignisses B, wenn A die Wahrscheinlichkeit des Auftretens von B erhöht und wenn es kein drittes Ereignis C gibt, das es gestattet, die Wahrscheinlichkeitsbeziehung zwischen A und B auszuklammern. Probabilistische Theorien der Kausalität verstehen Ursache-Wirkungs-Zusammenhänge statistisch, indem sie der Kausalanalyse bedingte Wahrscheinlichkeiten zu Grunde legen. Ähnlich wie bei Mackie (1974) müssen also Ursachen weder hinreichend noch notwendig sein, um als kausal relevant angesehen zu werden. Notwendige und hinreichende Bedingungen werden einfach als statistische Grenzfälle  $P(E|C) = 0$  bzw.  $P(E|C) = 1$  definiert. Probabilistische Theorien der Kausalität haben den Vorzug, dass sie mathematisch präzise formuliert, sehr allgemein und breit anwendbar sind (s. Oestermeier, 1997). Allerdings läuft man bei einer formalen Analyse der Struktur von Kausalrelationen aber auch Gefahr, inhaltsleer zu werden (fehlende Validität), wenn keine weiteren Angaben gemacht werden, welche Eigenschaften mental repräsentiert und als kausal relevant angesehen werden.

Gegen die probabilistischen Ansätze (Cartwright, 1989; Eells, 1991, Salmon, 1984; Suppes, 1970, 1984) aber auch gegen Humes Regularitätstheorie, wurde häufig eingewendet, dass nicht alle regulären Abfolgen (z.B. "der Tag folgt auf die Nacht und umgekehrt") und nicht alle statistischen Zusammenhänge kausal sein müssen. Bezogen auf Letzteres, kann beispielsweise eine höhere Anzahl der Geburten statistisch gesehen mit einer höheren Anzahl Störche einhergehen oder positiv korrelieren, ohne dass Störche Geburten verursachen.

Dementsprechend kann also nicht zwischen Kausalzusammenhängen und zufälligen statistischen Zusammenhängen, wie z.B. einer positiven Korrelation zwischen dem Auftreten von Störchen und der Geburtenrate differenziert werden. Mögliche Antworten auf dieses Problem der *akzidentellen Regularitäten* bieten *bedingte Kontraste* ("Die Störche wären nur dann kausal relevant, wenn andere potentielle erklärende Ursachen, wie z.B. der Einsatz von Verhütungsmitteln oder die Nahrungs- und Nistmöglichkeiten für Störche konstant gehalten wurden") und *kontrafaktische Konditionale* ("die Störche wären nur dann kausal relevant, wenn bei ausbleibender Erhöhung der Anzahl die Störche die Geburtenzahlen nicht gestiegen wären"). Das Problem, dass nicht alle regulären Abfolgen kausal sein müssen, kann mit Hilfe von Mackies bedingungsanalytischem Ansatz gelöst werden: Tag und Nacht verursachen einander deshalb nicht, weil in einer vollständigen Bedingungsanalyse von Tag und Nacht Sonneneinstrahlung und Erdrotation (sowohl um die eigene Achse als auch um die Sonne herum) miteinbezogen werden müssen. Für sich genommen ist weder der Tag für die Nacht, noch die Nacht für den Tag eine INUS-Bedingung.

Das Problem der akzidentellen Regularitäten und seine Lösungsversuche zeigen, dass der Versuch, Kausalität allein auf probabilistische Beziehungen zu reduzieren, als gescheitert angesehen werden muss (vgl. Hagmayer & Waldmann, 2006; Pearl, 2000). Der Kern der Unfähigkeit probabilistischer Ansätze, kausale Zusammenhänge oder Abfolgen einerseits von zufälligen Zusammenhängen oder nicht kausalen regulären Abfolgen andererseits zu unterscheiden, liegt im Wesentlichen darin, dass kausale Beziehungen gerichtet sind (von Ursachen zu Effekten und nicht umgekehrt) und sich diese Richtung nur in bestimmten Fällen und nur durch weitere Zusatzannahmen aus probabilistischen Daten herleiten lässt. Ein weiteres Problem bei der Beurteilung des Kausalstatus von statistischen Zusammenhängen sind die potentiell konkurrierenden Ko-Faktoren. Statistische Zusammenhänge sind nur dann kausal, wenn der Einfluss aller konfundierenden Faktoren konstant gehalten wird. In der Realität kann man aber nicht wissen, ob man alle Faktoren kennt und ob es sich dabei tatsächlich um bedeutsame Kausalfaktoren und nicht um irrelevante Ereignisse handelt (vgl. Hagmayer & Waldmann, 2006; Cartwright, 1989). Dadurch entsteht ein Zirkularitätsproblem, weil die Kontrolle konfundierender Faktoren voraussetzt, dass bekannt ist, welche Faktoren kausal bedeutsam sind und welche nicht.

Bei der Theorie der kausalen Bayesnetze (s. Glymour & Cooper, 1999; Hagmayer, 2001; Hagmayer & Waldmann, 2006; Pearl, 1988, 2000; Spirtes, Glymour & Scheines, 1993, 2000 für Überblicke) werden die probabilistischen Ansätze erweitert. Nach dieser Theorie wird der Erwerb von adäquatem Kausalwissen aus einem Zusammenspiel von probabilistischen

Informationen, abstraktem Vorwissen über Kausalmodelle (wie z.B. Hinweise über die zeitliche Reihenfolge), sowie einschränkenden Zusatzannahmen erklärt. Vertreter dieser Theorie (wie z.B. Pearl, 1988, 2000; Glymour & Cooper, 1999; Woodward, 2003) haben sich mit der Frage beschäftigt, wie und unter welchen Bedingungen sich aus vorliegenden probabilistischen Daten über die Zusammenhänge von Ereignissen die Struktur der zugrunde liegenden Kausalzusammenhänge rekonstruieren lässt und haben hierfür auch Computermodelle, so genannte "kausale Bayes-Netzwerke" entwickelt.

Ein weiterer Vorschlag, die Humeschen Regularitätstheorie zu ergänzen oder gar zu ersetzen, besteht in der so genannten *mechanistischen Sichtweise*, dass die Ursache nicht nur regelmäßig die Wirkung zur Folge hat, sondern durch einen kausalen Prozess, einen generativen Mechanismus oder die Übertragung von Energie mit der Wirkung verbunden ist (Bunge, 1987; Cartwright, 1989). So kritisiert beispielsweise Mario Bunge (1987), dass Humes Konzept der "konstanten Verknüpfung" den Unterschied zwischen einfacher Korrelation und aktivem Hervorbringen verwischt. In der Psychologie wurden ähnliche Annahmen (*causal power view*) zum Verständnis von Ursache-Wirkungsrelationen von White (1993), Ahn et al. (1995) oder Cheng (1993, 1997) vertreten.

### **2.3 Psychologische Rezeption von neueren Konzeptionen des Kausalitätsbegriffs**

Der mechanistische Power-Ansatz, Mackies bedingungsanalytischer Ansatz, die Berücksichtigung bedingter Kontraste und kontrafaktischer Konditionale, sowie die probabilistischen Theorien zur Kausalität und die Theorie der kausalen Bayes-Netzwerke haben die psychologische Theorienbildung zum kausalen Lernen vielfältig und stark beeinflusst.

Beispielsweise haben Kovariationsansätze (Cheng, 1993, 1997; Cheng et al., 1996), in denen auf irgendeine Art die Kovariation zwischen einer Ursache und einem Effekt eingeschätzt wird, ihren Ursprung in Humes Regularitätstheorie und ihren probabilistischen Weiterentwicklungen. Ein einflussreiches Modell zur Berechnung der Stärke des Zusammenhangs zwischen Ursache und Wirkung beinhaltet die Berechnung des Kontingenzmaßes  $\Delta P$  (Jenkins & Ward, 1965; Allan, 1980; Allan & Jenkins, 1983; s.a. Kapitel 3.1.1). Probabilistische Theorien liegen auch den Weiterentwicklungen von Kontingenztheorien wie dem Probabilistischen Kontrastmodell nach Cheng und Novick (1990, 1992), der Theorie der kausalen Power nach Cheng (1997) und der Theorie der Evidenzevaluation nach White (2003a, 2003b, 2003c, 2004, 2005a) maßgeblich zugrunde.

Das Probabilistische Kontrastmodell hat darüber hinaus eine weitere philosophische Wurzel in Mackies (1965, 1974; s.a. Hart & Honoré, 1959) Unterscheidung zwischen Ursachen und ermöglichenden Bedingungen (englisch: *enabling conditions*). Spezifisch für Chengs (1997) Ansatz ist die theoretische Fundierung durch philosophische Überlegungen zum Konzept der kausalen Power (Cartwright, 1989). Für Vertreter des Power-Ansatzes genügt es nicht, Ursachen als Ereignisse zu definieren, die zeitlich vor ihren Effekten auftreten. Vielmehr sind Ursachen dadurch charakterisiert, dass sie diese Effekte irgendwie hervorbringen oder sie erzeugen.

Die Theorie der kausalen Bayesnetze wurde vor allem seit der Jahrtausendwende von psychologischer Seite in immer größerem Umfang zur Erstellung kognitiver Modellierungen rezipiert (Danks, 2005; Gopnik, Glymour, Sobel, Schulz, Kushnir, & Danks, 2004; Gopnik, Sobel, Schulz, & Glymour, 2001; Glymour, 1998, 2001, 2003; Griffiths & Tenenbaum, 2005; Hagmayer, 2001; Hagmayer & Waldmann, 2006; Tenenbaum & Griffiths, 2001, 2003; Waldmann, 1996; Waldmann & Hagmayer, 2001; Waldmann & Holyoak, 1992).

Aus dem Facettenreichtum und der Bedeutungsvielfalt des Kausalitätsbegriffs ergeben sich für die empirisch psychologische Forschung sprachliche Probleme. So gehen verschiedene Forschungszweige im Bereich der Kausalkognition von unterschiedlichen Vorannahmen bezüglich des Kausalitätsbegriffs aus und versuchen dementsprechend, den Kausalitätsbegriff über unterschiedliche Grundbegriffe oder Basiskriterien zu definieren. Die psychologische Rezeption der vielfältigen philosophischen Kausalitätskonzeptionen erfolgte also uneinheitlich. Vertreter des Power-Ansatzes gehen beispielsweise von einer mechanistischen Begriffskonzeption aus und versuchen, Kausalität auf Begriffe wie "Kraft", "Energie", oder "Erzeugung" zurückzuführen. Anhänger assoziationalistischer Ansätze halten dagegen Begriffe wie "konstante Verknüpfung", "raum-zeitliche Nähe" oder "Sukzession" für grundlegender, während Kontingenztheoretiker den statistischen Zusammenhang zwischen zwei Ereignissen als Basiselement für ihr Kausalitätsverständnis begreifen. Die vielfältigen Verwendungsmöglichkeiten des Kausalitätsbegriffs entsprechen unserem alltagssprachlichen Verständnis von Kausalität, weil wir Kausalität sowohl über mechanistische Verursachungsketten (C ist Ursache für E, weil C über den Mechanismus D zu E führt) und kontrafaktische Formulierungen (C ist Ursache für E, weil E nicht gewesen wäre, wenn C nicht gewesen wäre), als auch über konstante Verknüpfungen und das Aufzeigen von statistischen Zusammenhängen (C ist Ursache für E, weil C die Wahrscheinlichkeit von E erhöht) begreifen können. Oestermeier (1997, S.81) hält es deshalb für sinnvoller, kausales Denken nicht auf die Anwendung möglichst einfacher und einheitlicher Kriterien zu

reduzieren, sondern durch den Versuch zu ersetzen, *"die Komplexität des alltäglichen kausalen Denkens durch multikriteriale Theorieentwürfe zu approximieren"*. Das bedeutet, dass sich die verschiedenen Theorieansätze zum kausalen Lernen nicht widersprechen müssen, sondern auf unterschiedliche Aspekte von Kausalität Bezug nehmen. Cheng (1997) versucht in ihrer in Kapitel 3.1.3 näher beschriebenen Power PC Theorie beispielsweise, mechanistische Power- und Kovariationstheorien miteinander zu vereinen und hebt dabei hervor, dass beide Ansätze auf unterschiedliche Probleme Bezug nehmen: Während Kovariationsansätze darauf Bezug nehmen, wie aus beobachtbaren Ereignissen Ursache-Wirkungs-Zusammenhänge erschlossen oder berechnet werden können (Problem der Kausalinduktion), gehen mechanistische Power-Ansätze von der Annahme aus, dass irgendeine Art von bereichsspezifischem oder bereichsübergreifendem Vorwissen existiert, welches die Interpretation von Beobachtungsdaten im Zusammenhang mit dem kausalen Induktionsprozess einbettet und strukturiert.

In Kapitel 3 sollen verschiedene Theorieansätze aus dem Bereich der psychologischen Kausalitätsforschung dargestellt werden. Zu Beginn wird kurz auf Befunde zur visuellen Wahrnehmung von Kausalität und zur Kausalattributionsforschung (Einleitung zu Kapitel 3) eingegangen. Der Schwerpunkt von Kapitel 3 liegt aber auf der Darstellung regel-basierter (Kapitel 3.1) und assoziationalistischer Theorien (Kapitel 3.2). Diese Theorien beschäftigen sich mit dem Problem der Kausalinduktion, mit dem Erwerb von Wissen über kausale Zusammenhänge aus Beobachtungsdaten. Am Rande soll in Kapitel 3 noch auf Theorien eingegangen werden, die sich mit dem Einfluss von abstraktem und bereichsspezifischem Vorwissen über Kausalmechanismen auf den Erwerb von Kausalwissen befassen (Kapitel 3.3). Eine Kritik an monomechanistischen Ansätzen (Kapitel 3.4) zum Erwerb von Kausalwissen und die Vorstellung eines integrativen Rahmens für Theorien zum Erwerb von Kausalwissen (Kapitel 3.5) beenden Kapitel 3.



### 3 Der Erwerb von Kausalwissen

Kapitel 2 hat gezeigt, dass in der Psychologie philosophische Theorien zur Kausalität rezipiert und zum Teil aufgegriffen worden sind (Oestermeier, 1997; Van der Meer, 2003). Innerhalb der Psychologie selbst hat die Forschung zur Kausalkognition viele verschiedene Richtungen eingeschlagen, einen Eindruck von der Vielfalt der Forschungslandschaft liefern Sperber, Premack & Premack (1995); May & Oestermeier (2001) und Van der Meer (2003). Nach Van der Meer (2003) haben zwei inhaltliche Schwerpunkte die traditionelle psychologische Kausalitätsforschung geprägt: Untersuchungen zur visuellen Wahrnehmung und zum Wissen über kausale Zusammenhänge.

*Visuelle Wahrnehmung von Kausalität:* Untersuchungen zur visuellen Wahrnehmung von Kausalität thematisieren die Frage, ob kausale Zusammenhänge unmittelbar beobachtbar sind oder nicht. Hier wurde ausgehend von Hume die These vertreten, dass Ursache-Wirkungs-Zusammenhänge als solche nicht unmittelbar beobachtbar sind, sondern nur aus der wiederholten Beobachtung zweier aufeinander folgender und raum-zeitlich nahe stehender Ereignisse oder Objekte erschlossen werden können (Cheng, 1997; Ross & Fletcher, 1985). Demgegenüber kam Michotte (1963) in Experimenten zur Stützung seiner Theorie der phänomenalen Kausalität zu dem Schluss, dass zumindest typische Kausalbeziehungen zwischen einzelnen Ereignissen wie "stoßen", "ziehen" und "schieben" direkt im Einzelfall wahrnehmbar sind und nicht erst rational erschlossen werden müssen. Die psychologische Kontroverse, ob kausale Zusammenhänge beobachtbar sind oder nicht, lässt sich allerdings auflösen, wenn davon ausgegangen wird, dass innerhalb dieser Debatte unterschiedliche Konzepte von Beobachtung zugrunde gelegt werden (s. Oestermeier, 1997, 2001).

Ein weiterer Forschungsstrang im Zusammenhang mit der visuellen Wahrnehmung von Kausalität beschäftigt sich mit der Frage, ob Menschen über angeborene Module zur Wahrnehmung physikalischer Kausalität verfügen. Leslie & Keeble (1987, s.a. Leslie, 1995, S. 124 ff.) konnten bereits bei 6-Monate alten Kindern, die kausale und non-kausale Bewegungsabfolgen beobachteten, unterschiedliche Aufmerksamkeitsmuster nachweisen. Ob aus solchen Hinweisen zwingend geschlossen werden kann, dass Menschen über angeborene Module zur Wahrnehmung physikalischer Kausalität verfügen, wird kritisch diskutiert. Weitere Befunde sprechen dafür, dass auch Lernfaktoren bei der Wahrnehmung physikalischer Kausalität eine Rolle spielen. Dafür spricht zum einen die Beobachtung, dass sich die Wahrnehmung physikalischer Kausalität aus Bewegungsabfolgen von der Kindheit

bis ins Erwachsenenalter verbessert (Kaiser & Proffitt, 1984), zum anderen ergaben sich bei Replikationen der Experimente von Michotte (1963) Hinweise auf beträchtliche interindividuelle Unterschiede (s. Morris, Nisbett & Peng, 1995). Kerzel & Hecht (2001) ergänzen die Arbeiten zur phänomenalen Kausalität um die pragmatische Annahme, dass die Wahrnehmung von Kausalbeziehungen stark von Intentionen, Zielen und Zwecken geprägt ist.

*Wissen über kausale Zusammenhänge:* Innerhalb der Psychologie wurden Fragen, die in Verbindung mit dem Wissen über kausale Zusammenhänge stehen, vor allem von der kognitiven und der Sozialpsychologie aufgegriffen. Innerhalb der Sozialpsychologie wurden vor allem die Untersuchungsansätze von Heider (1958) und Kelley (1967, 1972, 1973) prominent. Dabei wurde der Frage nachgegangen, welche Attributionsmuster menschlichen Kausalurteilen zu Grunde liegen. Dabei geht es weniger um die Frage, wie Wissen über kausale Zusammenhänge erworben wird, sondern darum, wie Menschen bereits vorhandenes Wissen zur Erklärung menschlicher Verhaltensweisen nutzen und einsetzen. Nach dem *varianzanalytischen Ansatz von Kelley* (1967, 1972, 1973) versucht ein Beobachter menschliche Verhaltensweisen (die Effekte) zu erklären, indem er die Ursachen derselben nach dem Kovariationsprinzip zu bestimmen versucht. In den Worten Kelleys (1967, S. 194) lautet das Kovariationsprinzip: *"The effect is attributed to that condition which is present, when the effect is present, and absent, when the effect is absent"*. Wenn für die Erklärung einer menschlichen Verhaltensweise mehrere Ursachen in Frage kommen, dann führen die urteilenden Personen bei der Ursachenzuschreibung implizit eine naive Varianzanalyse durch, und bestimmen so Haupteffekte und Interaktionen der Ursachen in Bezug auf die Effekte. Für die Erklärung menschlicher Verhaltensweisen nimmt Kelley an, dass die Ursachen für eine bestimmte menschliche Verhaltensweise in der Person, in den gegebenen Umweltreizen oder am spezifischen Zeitpunkt liegen können.

Das varianzanalytische Modell von Kelley (1967, 1972, 1973) hat zahlreiche Weiterentwicklungen erfahren. Nach dem Ansatz zum *kontrafaktischen Schließen* von Lipe (1991) lassen sich Kausalzuschreibungen grundlegend auf kontrafaktische Informationen zurückführen. Das Finden einer Kausalerklärung beruht auf einem dreistufigen Prozess, der sich aus den Teilprozessen Hypothesenbildung, Abschätzung der Stringenz der Hypothese und Prüfung alternativer Erklärungen zusammensetzt, wobei Lipe (1991) nur die beiden letzten Prozessstufen interessieren. Die Abschätzung der Hypothesenstringenz erfolgt nach dem Kovarianzprinzip, wobei die Kovarianzinformation nicht direkt zur Kausalerklärung heran gezogen, sondern zur Erzeugung der kontrafaktischen Information benutzt wird. Die



gewonnene kontrafaktische Information wird dann zur Prüfung alternativer Erklärungen verwendet.

Die Ansätze von Kelley (1967, 1972, 1973) und Lipe (1991) beschäftigen sich mit der Nutzung vorhandener Informationen bei Erklärungsfindung und Ursachenzuschreibung. Schwerpunkt der vorliegenden Arbeit ist nicht die Nutzung vorhandenen Wissens, sondern die Prozesse, die dem Erwerb von Wissen über Ursache-Wirkungs-Zusammenhänge zu Grunde liegen.

#### *Der Erwerb von Wissen über kausale Zusammenhänge*

Bis in die frühen 90er Jahre des vergangenen Jahrhunderts wurde die Forschung zum Erwerb von Kausalwissen von assoziationalistischen und regel-basierten Ansätzen dominiert (Allan, 1993; Baker et al., 1996; Waldmann, 1994). *Assoziationalistische Ansätze* (z.B. Allan, 1993; Dickinson, 2001; Dickinson, Shanks & Evenden, 1984; Shanks, 1995, 1993; Shanks & Dickinson, 1987; Shanks, Lopez, Darby & Dickinson, 1996; Wasserman et al., 1993; Wasserman et al., 1996) haben ihre philosophischen Wurzeln in einem Humeschen Kausalitätsverständnis. In der Psychologie wurde die Entwicklung assoziationalistischer Ansätze im Bereich der Kausalkognition durch die Idee inspiriert, dass sich die in Untersuchungen mit Tieren gefundenen Prinzipien des assoziativen Lernens auch zur Erklärung höherer Lernprozesse wie zum Beispiel dem Erwerb von Kausalwissen anwenden lassen. Das bekannteste assoziationalistische Modell ist das *Rescorla-Wagner-Modell*, kurz RWM (Rescorla & Wagner, 1972; s.a. Kapitel 3.2). *Regel-basierte Ansätze* (z.B. Allan, 1980; Cheng, 1997; Cheng & Novick, 1990, 1991, 1992; Novick & Cheng, 2004; Ward & Jenkins, 1965; White, 2003a, 2003b, 2003c, 2004, 2005a) sind maßgeblich durch ein statistisch-probabilistisches Kausalitätsverständnis geprägt und konzipieren den Lernenden als einen *intuitiven Statistiker*, der intuitiv den statistischen Zusammenhang zwischen Ursachen und Effekten abschätzt oder berechnet. Die dabei zum Einsatz kommenden Denk- und Lernprozesse sind komplexer als beim Assoziationslernen. Gemeinsam ist beiden Ansätzen, dass der Erwerb von Kausalwissen primär als eine Abfolge datengesteuerter (*bottom up*) Informationsverarbeitungsprozesse gesehen wird (Waldmann, 1994). Die Möglichkeit von Wissenseinflüssen wird zwar nicht ausgeschlossen, doch wird der Erwerb von Kausalwissen meist in Domänen untersucht, für die kein spezifisches Vorwissen verfügbar ist.

In der Mitte der 90er Jahre des vergangenen Jahrhunderts wurden verstärkt Theorien entwickelt, die von der Annahme ausgehen, dass die datengesteuerte Informationsaufnahme durch wissensgeleitete Prozesse (*top down*) gesteuert und strukturiert wird. *Mechanistische Ansätze* (Ahn, Kalish, Medin & Gelman, 1995; Koslowski, 1996) gehen von der Annahme

aus, dass der Erwerb von Kausalwissen hauptsächlich durch unser Vorwissen über kausale Mechanismen gesteuert wird. Eine grundlegende Schwäche mechanistischer Ansätze liegt darin, dass die Existenz von spezifischem Vorwissen über kausale Mechanismen einfach vorausgesetzt wird und nicht erklärt wird, wie dieses Wissen über kausale Mechanismen ursprünglich erworben wird (Cheng, 1993, 1997), d.h. das Problem der Kausalinduktion wird nicht adressiert. Die *Theorie der Kausalmodelle* (Waldmann, 1994, 1996, 2000; Waldmann & Hagmayer, 2001; Waldmann & Holyoak, 1992; Waldmann, Holyoak & Fratianne, 1995) postuliert eine Interaktion zwischen abstrakten Vorannahmen über hypothetische Kausalstrukturen (den Kausalmodellen) und dem Lern-Input. Zentral ist dabei die Vorstellung, dass die Struktur der Kausalmodelle die Interpretation beobachteter Zusammenhänge steuert, umgekehrt kann aber auch das ursprüngliche Kausalmodell durch den Lern-Input verändert werden.

Im Folgenden sollen verschiedene Theorieansätze zum Erwerb von Kausalwissen näher dargestellt werden. Nach Cheng (1997) kann die Frage, wie wir Wissen über kausale Zusammenhänge erwerben, in zwei Komponenten aufgeteilt werden: Zum einen, wie erwerben wir ursprünglich Wissen darüber, dass überhaupt ein kausaler Zusammenhang besteht (Problem der Kausalinduktion), und zum anderen, wie beeinflusst (angeborenes oder erworbenes) bereichsspezifisches Vorwissen die nachfolgende Beurteilung kausaler Zusammenhänge. Das Problem der Kausalinduktion veranschaulicht Cheng (1997) an einem Kind, das aus Erfahrung lernt, dass Keramikteller kaputt gehen, wenn sie fallen gelassen werden, während Plastikteller ganz bleiben. Den Einfluss von bereichsspezifischem Vorwissen veranschaulicht Cheng (1997) an dem Beispiel, dass der Sonnenaufgang nicht durch das Krähen eines Hahnes auf einer Farm verursacht wird, obwohl beide Ereignisse konstant miteinander verknüpft sind und das Krähen des Hahnes dem Sonnenaufgang zeitlich regelmäßig vorausgeht. Cheng (1997) zufolge wissen wir intuitiv aus Analogieschlüssen, dass dem Krähen des Hahnes keine kausale Kraft oder Mechanismus zugrunde liegen kann, die das Aufgehen der Sonne hervor bringt.

Cheng (1997) zufolge können Prozesse der Kausalinduktion und Prozesse der wissensgestützten Beeinflussung von Kausalurteilen getrennt voneinander betrachtet und untersucht werden. Der Schwerpunkt in der vorliegenden Arbeit liegt dabei auf Theorien, die sich mit dem Problem der Kausalinduktion, mit dem Erwerb von Wissen über kausale Zusammenhänge aus Beobachtungsdaten, beschäftigen. Auf Theorien, die sich mit dem Einfluss von bereichsspezifischem Vorwissen über kausale Mechanismen (mechanistische Ansätze) und von abstrakten, bereichsunspezifischen Kausalmodellen (Theorie der

Kausalmodelle) auf den Erwerb von Kausalwissen befassen, wird hier nur kurz eingegangen. Die Bedeutung und der Einfluss von sowohl bereichsspezifischem Vorwissen über kausale Mechanismen als auch von abstrakten, bereichsunspezifischen Kausalmodellen wurde mittlerweile in zahlreichen Untersuchungen nachgewiesen (s. Hagmayer & Waldmann, 2006, für einen Überblick).

Im Bereich der Kausalkognition wurde die Frage danach, wie Menschen Wissen über kausale Zusammenhänge aus Beobachtungsdaten erlangen und beurteilen, intensiv untersucht (s. z.B. Allan, 1993; Hagmayer & Waldmann, 2006; Shanks, 2007; Shanks, Holyoak & Medin, 1996, für Überblicke). Sehr häufig wurden die Probanden dabei dazu aufgefordert, ein quantitatives Urteil über die Stärke eines einfachen Ursache-Wirkungs-Zusammenhangs abzugeben. Die Stärke eines Ursache-Wirkungs-Zusammenhangs gibt dabei an, wie wirksam die Ursache darin ist, den Effekt hervorzurufen.

### 3.1 Regel-basierte Ansätze

Regel-basierte Ansätze gehen davon aus, dass sich Menschen wie intuitive Statistiker verhalten, wenn sie Wissen über Ursache-Wirkungs-Zusammenhänge erwerben. Mit Hilfe der Regeln werden Beobachtungen gebündelt und strukturiert, daran anschließend werden sie dazu genutzt, um auf die Stärke kausaler Zusammenhänge zu schließen. Aktuell werden verschiedene Regeln diskutiert (s. Hagmayer & Waldmann, 2006 für einen Überblick), drei davon, das *Probabilistische Kontrastmodell* nach Cheng & Novick (1990, 1991, 1992), die *Theorie der kausalen Power* nach Cheng (1997) und die *Theorie der Evidenzevaluation* (White, 2003a, 2003b, 2003c, 2004, 2005a) sollen hier näher beschrieben werden. Einen mehr oder weniger gemeinsamen Ausgangspunkt für alle diese drei Modelle bilden kontingenztheoretische Annahmen, mit denen deshalb die Darstellung regel-basierter Ansätze beginnt.

#### 3.1.1 Kontingenztheoretische Ansätze

Kontingenztheoretische Ansätze gehen von der Annahme aus, dass Kausalzusammenhänge probabilistisch sind. Ein Ereignis wird dann als Ursache eines anderen Ereignisses, dem Effekt, definiert, wenn die Wahrscheinlichkeit des Effekts bei Vorliegen der fraglichen Ursache ungleich der Wahrscheinlichkeit des Effekts bei Abwesenheit der fraglichen Ursache ist. Der Begriff der Kontingenz wird verwendet, weil er den Zusammenhang zwischen zwei diskreten Variablen oder zwei mehrklassigen, nominalskalierten Merkmalen quantifiziert.

### Bestimmung einfacher Kontingenzen: Die $\Delta P$ -Regel

Einfache Kontingenztheorien beziehen sich auf Situationen, in denen eine Ursache mit einem Effekt verknüpft ist. Vertreter von einfachen Kontingenztheorien nehmen an, dass Menschen die Kontingenz zwischen der vermeintlichen Ursache und dem fraglichen Effekt ermitteln, wenn sie die Stärke von Ursache-Wirkungs-Zusammenhängen beurteilen. Demnach verhalten sich Menschen bei der Beurteilung kausaler Zusammenhänge wie intuitive Statistiker (Allan, 1993; Peterson & Beach, 1967), die implizit Kontingenzen berechnen. Für monokausale Zusammenhänge mit nur einer Wirkung und einer in Frage kommenden Ursache kann die Stärke eines Ursache-Wirkungs-Zusammenhanges über das Kontingenzmaß  $\Delta P$  (Ward & Jenkins, 1965; McKenzie, 1994) bestimmt werden.  $\Delta P$  ist definiert als:

$$\Delta P = P(E|C) - P(E|\neg C) \quad (1)$$

wobei  $P(E|C)$  der bedingten Wahrscheinlichkeit des Effektes E entspricht, wenn die Ursache C gegeben ist, und  $P(E|\neg C)$  für die bedingte Wahrscheinlichkeit des Effektes E steht, wenn die Ursache C nicht gegeben ist.  $P(E|C)$  wird auch als Erfolgsrate bezeichnet und  $P(E|\neg C)$  als Basisrate. Das Kontingenzmaß  $\Delta P$  ergibt sich also aus der Differenz (dem Kontrast) zwischen Erfolgs- und Basisrate. Eine Ursache wird als *generativ* bezeichnet, wenn der Effekt bei vorhandener Ursache häufiger auftritt als bei fehlender Ursache. In diesem Fall ist die Kontingenz positiv. Beispiele für eine generative oder erzeugende Ursache wären ein Düngemittel, welches das Wachstum von Pflanzen fördert, oder Viren (Bakterien), die zu einer Infektion führen. Bei so genannten *präventiven* Ursachen wird das Auftreten des Effekts durch die Ursache seltener, d.h. unwahrscheinlicher gemacht. Die Differenz zwischen  $P(E|C)$  und  $P(E|\neg C)$  ist dann negativ. Beispiele für präventive Ursache wären eine vorbeugende Impfung, die das Ausbrechen einer Krankheit verhindert, oder Pflanzenschutzmittel, welche Ungeziefer vernichten.

Die beiden bedingten Wahrscheinlichkeiten  $P(E|C)$  und  $P(E|\neg C)$  lassen sich aus den Zellhäufigkeiten der 2 x 2 Kontingenztabelle in Abbildung 1 berechnen.

	Wirkung tritt ein	Wirkung tritt nicht ein
Ursache gegeben	a	b
Ursache nicht gegeben	c	d

**Abbildung 1:** Kombinationen in einer 2 (Wirkung tritt ein oder tritt nicht ein) x 2 (Ursache ist gegeben oder nicht gegeben) Kontingenztabelle. Die Kleinbuchstaben a, b, c und d stehen für Ereignishäufigkeiten.

Die Erfolgsrate  $P(E|C)$  kann berechnet werden, indem  $a$  zur Summe  $a + b$  ins Verhältnis gesetzt wird, die Basisrate  $P(E|\neg C)$  kann berechnet werden indem  $c$  zur Summe  $c + d$  ins Verhältnis gesetzt wird. Entsprechend kann Gleichung (1) durch Gleichung (2) ersetzt werden:

$$\Delta P = a / (a + b) - c / (c + d) \quad (2)$$

Von vielen Forschern (z.B. Ward & Jenkins, 1965; Allan, 1993) wurde die Berechnung der Kontingenz mit der  $\Delta P$ -Regel als statistisch angemessen und richtig angesehen. Deshalb wurde die  $\Delta P$ -Regel ebenso wie später die Berechnung der kausalen Stärke  $p$  nach Cheng (1997) als "normativ" bezeichnet, während einfachere Strategien wie die *Zelle a-Strategie*, die *a versus b Strategie* oder die  *$\Delta D$ -Regel* (s. Shanks, 1995 für einen Überblick) als nicht normative Heuristiken eingestuft werden, weil sie als statistisch unangemessen gelten und zu Fehleinschätzungen der tatsächlichen Kontingenz führen. In weiteren Verlauf dieses Abschnittes werden auf Seite 26 ff. diese nicht normativen Heuristiken unter der Überschrift "*Empirische Befunde und Kritik: Die Suche nach der Regel*" näher beschrieben und bewertet.

#### *Empirische Befunde und Kritik*

*Überprüfungen der  $\Delta P$ -Regel:* Vor dem Aufkommen der Power PC Theorie galt die  $\Delta P$ -Regel lange Zeit als die angemessene Regel um den statistischen Zusammenhang zwischen zwei möglicherweise kausal verknüpften Ereignissen zu ermitteln (Allan, 1980, 1993; Cheng & Holyoak, 1995; Ward & Jenkins, 1965). Zahlreiche Studien haben sich deshalb darauf beschränkt festzulegen, ob Menschen das Ausmaß oder zumindest die Gerichtetheit der Kontingenz zwischen zwei Ereignissen genau beurteilen können. In solchen Studien wurde den Probanden eine Reihe von Problemen vorgelegt, wobei die Kontingenzen zwischen der vermeintlichen Ursache und ihrer Wirkung jeweils variieren. Nach jedem Problem sollten die Probanden die Kontingenz zwischen den beiden Ereignissen einschätzen. Dabei muss aber berücksichtigt werden, dass in den verschiedenen Untersuchungen sehr unterschiedliche Paradigmen eingesetzt wurden (s. Allan, 1993; Alloy & Tabachnik, 1984; Hagmayer & Waldmann, 2006, für Überblicke). Beispielsweise wurde variiert, in welcher Form die Informationen dargeboten wurden, entweder summarisch in Tabellen (Ward & Jenkins, 1965; Anderson & Sheu, 1995), Tortendiagrammen (Lober & Shanks, 2000), Falllisten (Van Hamme & Wasserman, 1993), Bildertafeln (Buehner, Cheng & Clifford, 2003) oder als eine Abfolge von Einzelfällen in Form eines "trial-by-trial" Lernens (Smedslund, 1963; Shaklee & Tucker, 1980).

Auch hinsichtlich der Vorgehensweise bei der Analyse von Einschätzungen des Kausalzusammenhangs unterschieden sich die Studien, meist wurden entweder Regressionen bzw. Korrelationen zwischen den Probandenurteilen zu verschiedenen Regeln oder Heuristiken berechnet (z.B. Ward & Jenkins, 1965; Anderson & Sheu, 1995), oder es wurden die von den Probanden eingesetzten Strategien bei jeder Person über die Einschätzungen bei einer Vielzahl von Aufgaben rekonstruiert (z.B. Arkes & Harkness, 1983; Shaklee & Hall, 1983; Shaklee & Mims, 1981, 1982; Shaklee & Tucker, 1980; Shaklee & Wasserman, 1986). Auf diese unterschiedlichen Vorgehensweisen soll im Anschluss an den unmittelbar nachfolgenden Ergebnisüberblick kritisch eingegangen werden.

Die Ergebnisse von Studien mit dem Ziel der Überprüfung der  $\Delta P$ -Regel sind uneinheitlich (Allan, 1993; Hagmayer & Waldmann, 2006). Eine Reihe von Studien (z.B. Allan & Jenkins, 1980, 1983; Alloy & Abramson, 1979; Vallée-Tourangeau, Hollingsworth & Murphy, 1998; Wasserman et al., 1993) berichten von hohen Korrelationen zwischen den Kontingenzzurteilen der Probanden und der nach der  $\Delta P$ -Regel ermittelten Kontingenz. In anderen Studien (Allan, 1993; Allan & Jenkins, 1980; Anderson & Sheu, 1995; Arkes & Harkness, 1983; Baker, Barbier & Vallée-Tourangeau, 1989; Buehner & Cheng, 1997; Jenkins & Ward, 1965; Kao & Wasserman, 1993; Nisbett & Ross, 1980; Schustack & Sternberg, 1981; Shaklee & Tucker, 1980; Shanks, 1987, 1993; Shanks et al. 1996; Smedslund, 1963; Wasserman, Dorner & Kao, 1990, Wasserman, Elek, Chatlosh & Baker, 1993) konnten aber auch systematische Abweichungen von der  $\Delta P$ -Regel entdeckt werden. So konnte beispielsweise gefunden werden, dass Kontingenzzurteile bei konstantem  $\Delta P$  nicht gleich ausfallen, sondern mit der Häufigkeit der Wirkung ansteigen (Allan & Jenkins, 1980, 1983; Alloy & Abramson, 1979; Baker, Barbier, Vallée-Tourangeau, 1989; Shanks & Dickinson, 1991). Dieser so genannte "Dichte-Bias" (englisch: *density bias*, s.a. Allan, 1993) kann von einfachen Kontingenztheorien wie der  $\Delta P$ -Regel nicht erklärt werden.

Da sich die immer wieder gefundenen Abweichungen von der  $\Delta P$ -Regel mindernd auf die Korrelationen zwischen  $\Delta P$  und den tatsächlichen Kontingenzzurteilen der Probanden auswirken, wurde von einer Reihe von Autoren vorgeschlagen, die vier Zellohäufigkeiten, mit deren Hilfe das Kontingenzmaß  $\Delta P$  berechnet werden kann, unterschiedlich zu gewichten. So können gewichtete  $\Delta P$ -Regeln, in denen die Zellohäufigkeiten der Bedeutung nach in der Reihenfolge Zelle a > Zelle b > Zelle c > Zelle d gewichtet werden, die Kontingenzzurteile von Probanden besser vorhersagen als eine ungewichtete  $\Delta P$ -Regel, in der die vier Zellohäufigkeiten die gleiche Bedeutung haben (Anderson & Sheu, 1995; Kao & Wasserman, 1993; Mandel & Lehmann, 1998; Wasserman, Dorner & Kao, 1990; Wassermann, Kao, Van

Hamme, Katagiri & Young, 1996). Dieser Befund bedeutet aber nicht, dass die Probanden in ihrem Vorgehen die Zellhäufigkeiten a, b, c und d tatsächlich unterschiedlich gewichten, sondern eher, dass einige der Probanden entsprechend der  $\Delta P$ -Regel die vier Zellhäufigkeiten gleichermaßen berücksichtigen, während andere Probanden nur diejenigen Informationen berücksichtigen, in denen die Ursache gegeben ist (Zellen a und b). Anderson & Sheu (1995) stützen sich dabei auf Studien von Allan & Jenkins (1980), Shaklee & Tucker (1980) und Wasserman et al. (1990), die allesamt zu dem Ergebnis kamen, dass eine Untergruppe von Probanden sich in ihrem Urteil entsprechend der  $\Delta P$ -Regel verhält.

Weitere Möglichkeiten im Umgang mit Abweichungen von der  $\Delta P$ -Regel wurden zum einen darin gesehen, die Theorie weiterzuentwickeln und Zusatzannahmen zu bilden, wie es zum Beispiel in der Focal Set-Theorie der  $\Delta P$ -Regel (Cheng & Novick, 1990, 1991, 1992; Cheng & Holyoak, 1995) oder der Power PC Theorie (Cheng, 1997) geschehen ist.

*Die Suche nach der Regel:* Die vielen gefundenen Abweichungen von der  $\Delta P$ -Regel warfen natürlich auch die Frage auf, ob Probanden bei der Beurteilung von Kontingenzen nicht einfach von ganz anderen Regeln oder Strategien als der statistisch angemessenen  $\Delta P$ -Regel Gebrauch machen. Zudem besagen Studien, in denen hohe, aber nie perfekte Korrelationen zwischen den Probandenurteilen und der nach der  $\Delta P$ -Regel ermittelten Kontingenz gefunden wurden, nicht notwendigerweise, dass die Kausalurteile der Probanden tatsächlich auf den Gebrauch der  $\Delta P$ -Regel zurückzuführen sind (Allan, 1993).

Zahlreiche Untersuchungen haben sich deshalb mit der Frage auseinander gesetzt, ob Menschen beim Erlernen von Ursache-Wirkungs-Zusammenhängen überhaupt von der  $\Delta P$ -Regel Gebrauch machen und ob sie neben oder an Stelle der  $\Delta P$ -Regel andere Strategien oder einfachere Heuristiken einsetzen. Dabei kam eine Reihe von Studien zu dem Schluss, dass Menschen von komplexen Regeln wie der  $\Delta P$ -Regel keinen oder nur einen sehr eingeschränkten Gebrauch machten. (Inhelder & Piaget, 1958; Jenkins & Ward, 1965; Kao & Wasserman, 1993; Nisbett & Ross, 1980; Schustack & Sternberg, 1981; Shaklee & Tucker, 1980; Smedslund, 1963). Mit anderen Worten ausgedrückt, bedeutet das, dass Menschen keine bedingten Wahrscheinlichkeiten berechnen und dann die Differenz aus ihnen bilden. Stattdessen konnte gefunden werden, dass viele Probanden häufiger andere, weit einfachere Strategien verwenden als die  $\Delta P$ -Regel (Inhelder & Piaget, 1958; Shaklee & Tucker, 1980; Shaklee & Mims, 1981, 1982; Shaklee & Hall, 1983). Bei all diesen einfachen Strategien oder Heuristiken werden keine bedingten Wahrscheinlichkeiten miteinander verglichen, sondern Häufigkeitsvergleiche angestellt.

Die einfachste dieser Strategien, die so genannte *Zelle a-Strategie*, wurde von Smedslund (1963; s.a. Inhelder & Piaget, 1958; Nisbett & Ross, 1980) identifiziert. Er kam in seinen Untersuchungen mit Schwesternschülerinnen zu dem Ergebnis, dass normale Erwachsene ohne eine Statistik-Ausbildung über kein adäquates Verständnis von Korrelation bzw. Kontingenz im Sinne der  $\Delta P$ -Regel verfügen. Die Schwesternschülerinnen in der Untersuchung von Smedslund machten stattdessen ihr Urteil über den Zusammenhang von zwei Ereignissen exklusiv von der Anzahl der Fälle abhängig, in denen Ursache und Wirkung gemeinsam auftreten. Fälle, in denen entweder nur die Ursache oder nur die Wirkung oder keins von beiden vorkommen, blieben weitestgehend unberücksichtigt, d.h. aus der  $2 \times 2$  Kontingenztafel in Abbildung 1 (S. 22) wurden bei der Urteilsbildung nur Fälle aus Zelle a, nicht aber solche aus den Zellen b, c und d herangezogen. Nach einer etwas anderen Variante dieser *Zelle a-Strategie* (s. Inhelder & Piaget, 1958, Shaklee & Tucker, 1980) wird ein Zusammenhang zwischen zwei Ereignissen dann als positiv gesehen, wenn die Anzahl der Fälle in Zelle a größer ist als in den Zellen b, c und d.

Anders als Smedslund (1963) fanden Shaklee und Mitarbeiter (Shaklee & Mims, 1982; Shaklee & Hall, 1983; Shaklee & Wassermann, 1986) heraus, dass die meisten Ihrer Probanden (die Zahlen schwanken zwischen 35 und 38,1 %) von einer anderen Strategie, nämlich der *a versus b Strategie* Gebrauch machten. Nach der *a versus b Strategie* (Inhelder & Piaget, 1958) basiert die kausale Urteilsbindung der Probanden auf einem Vergleich der Häufigkeiten in Zelle a mit denen in Zelle b, es wird also geprüft, ob der interessierende Effekt bei Anwesenheit der Ursache häufiger vorkommt als nicht vorkommt.

Probanden die nach der *a versus c Strategie* (Inhelder & Piaget, 1958) vorgehen, vergleichen, ob der Effekt häufiger auftritt, wenn die Ursache vorliegt, als wenn sie nicht vorliegt.

Eine im Vergleich zur *Zelle a-Strategie*, zur *a versus b Strategie* und zur *a versus c Strategie* relativ komplexe Strategie, die so genannte  $\Delta D$ -Regel, berücksichtigt wie die  $\Delta P$ -Regel alle vier Zellen einer  $2 \times 2$ -Kontingenztafel. Der  $\Delta D$ -Regel (Shanks, 1995; Allan, 1993; Inhelder & Piaget, 1958) zu Folge sind Urteile über Ursache-Wirkungs-Zusammenhänge eine Funktion der Differenz der Evidenz, die für einen Kausalzusammenhang spricht, und der Evidenz, die dagegen spricht. Für einen Kausalzusammenhang sprechen solche Fälle, in denen die potentielle Ursache und die Wirkung gemeinsam auftreten (Zelle a) oder gemeinsam ausbleiben (Zelle d). Gegen einen Kausalzusammenhang sprechen Fälle, in denen entweder nur die potentielle Ursache (Zelle b) oder nur die Wirkung (Zelle c) in Erscheinung treten. Das Differenzmaß  $\Delta D$  bildet sich also aus der Differenz der Summe der Häufigkeiten aus den Zellen a und d und der Summe der Häufigkeiten aus den Zellen b und c:



$$\Delta D = (a + d) - (b + c) \quad (3)$$

Eine Reihe von Untersuchungen kam zu dem Ergebnis, dass unter bestimmten Umständen mehr Probanden von der  $\Delta D$ -Regel als von einer anderen Strategie Gebrauch machten (Allan & Jenkins, 1983; Arkes & Harkness, 1983, Experiment 4; Shaklee & Tucker, 1980; Shaklee & Mims, 1981). Die Zahlen der Probanden, die von der  $\Delta D$ -Regel Gebrauch machten, schwankten bei Shaklee und Mitarbeitern zwischen 35 und 41 %, bei Arkes & Harkness (1983) war die  $\Delta D$ -Regel sogar bei 51 % vorherrschend.

Es gab aber auch Studien, die belegen, dass die meisten Probanden von der  $\Delta P$ -Regel Gebrauch machten. In der Studie von Ward & Jenkins (1965) waren es ca. 43 %, in der von Allan & Jenkins (1980) knapp über 44 %. Alloy & Abramson (1979) kamen ebenfalls zu dem Ergebnis, dass ihre Probandenurteile insgesamt recht genau mit den Vorhersagen der  $\Delta P$ -Regel übereinstimmten.

Um die Frage zu beantworten, welche dieser Regeln spontan eingesetzt wird, wurde eine Reihe von Forschungsmethoden entwickelt (s. Waldmann, 1994, für einen Überblick). Eine bereits früh verwendete Methode bestand darin, Versuchsteilnehmer die Stärke des Ursache-Wirkungs-Zusammenhangs einschätzen zu lassen und dann diese Schätzungen mit den von den unterschiedlichen Regeln vorhergesagten Werten zu korrelieren (s. z.B. Allan & Jenkins, 1980, 1983; Wassermann, Dorner & Kao, 1990; Wassermann et al., 1983). Diese *korrelative Methode* hat aber eine Reihe von Problemen (Busemeyer, 1991; s.a. Waldmann, 1994): So korrelieren beispielsweise die Probandenurteile nicht nur mit einer Regel hoch, sondern mit mehreren, weil die verschiedenen Regeln in vielen Fällen zu den gleichen Einschätzungen führen. Außerdem liegen keine Nachweise dafür vor, dass diejenige Regel, die am höchsten korreliert, tatsächlich die Regel ist, die dem Verhalten zu Grunde liegt.

Eine andere Methode, das so genannte *regel-analytische Verfahren* besteht darin, den Probanden Problemaufgaben vorzulegen, die diagnostisch für bestimmte Regeln sind (z.B. Arkes & Harkness, 1983; Shaklee & Hall, 1983; Shaklee & Mims, 1981, 1982; Shaklee & Tucker, 1980; Shaklee & Wasserman, 1986). Dabei wird von der Annahme ausgegangen, dass einzelne Aufgaben mit allen Regeln richtig bearbeitet werden können (z.B. deterministische Fälle, in denen nur Zelle  $a > 0$  ist), während andere Aufgaben nur unter Zuhilfenahme komplexerer Strategien, wie z.B. der  $\Delta P$ -Regel, lösbar sind. An Hand der von den jeweiligen Probanden gelösten Problemaufgaben wurde dann die von den Probanden eingesetzte Strategie rekonstruiert. Löst ein Proband beispielsweise Aufgaben, die nur mit Hilfe der  $\Delta P$ -Regel "richtig" bearbeitet oder gelöst werden können, dann wird angenommen,

dass dieser Proband die  $\Delta P$ -Regel angewendet hat. Löst ein Proband hingegen nur Aufgaben die auch mit der *Zelle a-Strategie* bearbeitet werden können, aber keine Aufgaben, die den Einsatz komplexerer Strategien erforderlich machen, dann wird von dieser Person angenommen, dass sie von der *Zelle a-Strategie* Gebrauch gemacht hat.

Auch das regel-analytische Verfahren hat seine Probleme: So ist die Zuordnung von Probanden zu Strategien nicht klar, wenn die Probanden fehlerbedingt von einer Strategie abweichen oder wenn sie ihre Strategie von Aufgabe zu Aufgabe verändern. Mit der korrelativen Methode teilt das regel-analytische Verfahren das Problem, dass eine korrekte Zuordnung nur dann gelingen kann, wenn man tatsächlich alle von den Probanden eingesetzten Regeln kennt. Es ist schließlich denkbar, dass der Versuchsleiter die von den Probanden eingesetzten Strategien gar nicht kennt oder bei der Konstruktion der Problemaufgabe nicht berücksichtigt. Zudem gehen sowohl die korrelative als auch die regel-analytische Methode davon aus, dass die Probanden die vorgegebenen Zellinformationen gleich gewichten. Eine unterschiedliche Gewichtung von Zellinformationen muss nicht beinhalten, dass die Probanden von einer bekannten Strategie abweichen.

Ein Manko bei der bisherigen Suche nach den von den Probanden eingesetzten Strategien besteht in unseren Augen darin, dass die Probanden selbst bislang kaum zu den von ihnen eingesetzten Strategien befragt wurden. Zwar weisen mündliche oder schriftliche Befragungen ihrerseits erhebliche methodische Mängel hinsichtlich der Testgütekriterien (Objektivität, Reliabilität, Validität) auf (s. z.B. Bortz, 1984), doch können qualitative Selbstauskünfte der Probanden quantitative Kausal- oder Kontingenzzurteile zusätzlich bestätigen und darüber hinaus dabei helfen, zu entscheiden, ob ein Proband seine Strategie über mehrere Problemaufgaben hinweg gewechselt hat und welche Strategie ein Proband eingesetzt hat, falls mehrere Strategien zu gleichen Vorhersagen führen. Zudem können idiosynkratische Strategien, deren Ergebnisse von keiner der bekannten Strategien vorhergesagt werden, besser identifiziert werden.

Zusammenfassend kann festgestellt werden, dass die Suche nach einer einheitlichen, allgemein gültigen Regel, die dem Kontingenzzurteil zu Grunde liegt, methodische Probleme beinhaltet und insgesamt betrachtet erfolglos geblieben ist. Cheng & Novick (1992, s.a. Waldmann, 1994) konnten zudem zeigen, dass alle vorgestellten Regeln mit Ausnahme der  $\Delta P$ -Regel das so genannte "Basisratenproblem" teilen, das heißt das die Regeln bei unterschiedlichen Basisraten für die Ereignisse (die in Frage kommende Ursache und Wirkung) zu unterschiedlichen Ergebnissen über den Kausalzusammenhang kommen. In Kapitel 4.1 und Kapitel 5 soll heraus gearbeitet werden, dass die Wahl der Regel unter

anderem von situativen, aufgabenbezogenen und medialen Einflussfaktoren, so wie von individuellen Unterschieden und der Gedächtnisbeanspruchung abhängt.

*Bewertung:* Unabhängig von den zahlreichen Untersuchungen, die belegen, dass die Urteile von Versuchsteilnehmern in mehr oder weniger systematischer Weise von den tatsächlichen Kontingenzen abweichen (Allan, 1993; Allan & Jenkins, 1980; Anderson & Sheu, 1995; Arkes & Harkness, 1983; Baker, Barbier & Vallée-Tourangeau, 1989; Buehner & Cheng, 1997; Jenkins & Ward, 1965; Kao & Wasserman, 1993; Schustack & Sternberg, 1981; Shaklee & Tucker, 1980; Shanks, 1987, 1993; Shanks et al. 1996; Smedslund, 1963; Wasserman, Dorner & Kao, 1990, Wasserman, Elek, Chatlosh & Baker, 1993), herrschte in Forscherkreisen bis Mitte der 90er Jahre Übereinstimmung darüber, dass die  $\Delta P$ -Regel die normative, d.h. statistisch angemessene Herangehensweise bei der Berechnung der Stärke von Ursache-Wirkungs-Zusammenhängen darstellt (Jenkins & Ward, 1965; Allan, 1980, 1993). Ein Vorteil des Kontingenzmaßes  $\Delta P$  ist es, dass es direkt aus Beobachtungsdaten ermittelt werden kann. Demgegenüber hat das Kontingenzmaß  $\Delta P$  aber den Nachteil, dass es sich mit jeder weiteren Beobachtung ändert, so dass sich dadurch auch der Kausalzusammenhang ständig ändern müsste. Zudem liegt der Berechnung des Kontingenzmaßes  $\Delta P$  ein statistisch-probabilistischer Kausalitätsbegriff zu Grunde. Dementsprechend hat die Berechnung der Kontingenz als Maß für die Stärke von Ursache-Wirkungs-Zusammenhängen dieselben Vor- und Nachteile wie die probabilistische Theorie der Kausalität (siehe Kapitel 2.2, S.12 ff.).

Eine zentrale Kritik an einfachen Kontingenztheorien beruht darauf, dass nicht jede Form eines Zusammenhangs zwischen zwei Variablen und nicht jede Wahrnehmung einer Aufeinanderfolge zwischen zwei Ereignissen kausal sein muss: Kovariation impliziert nicht immer Kausalität (Cheng, 1997), d.h. das berechnete Maß für die Kontingenz differenziert nicht zwischen Kausalzusammenhängen und beobachtbaren, aber nicht kausalen statistischen Zusammenhängen, den so genannten Scheinkorrelationen.

Waldmann (1994) bemängelt an einfachen Kontingenztheorien den Umstand, dass es sich dabei nicht um Prozesstheorien handelt, die den Lernverlauf modellieren, sondern um Kompetenztheorien, die das Urteil an der Lernasymptote erklären. Bei einfachen Ursache-Wirkungs-Zusammenhängen kommen assoziationalistische und Kontingenztheorien allerdings zu ähnlichen Vorhersagen (Chapman & Robbins, 1990).

Ein weiterer Kritikpunkt an einfachen Kontingenztheorien wurde von Cheng & Novick (1992) vorgetragen. Er beinhaltet, dass die Ermittlung der Kontingenz zwischen zwei Ereignissen niemals von den jeweils gegebenen Rahmenbedingungen unabhängig ist. In Kapitel 2.2 (S.10 ff.) wurde in Verbindung mit dem bedingungsanalytischen Ansatz von

Mackie (1974) das Beispiel diskutiert, dass eine eisglatte Fahrbahn einen Autounfall verursacht hat. Die Untersuchung der Auswirkungen einer eisglatten Fahrbahn auf die Entstehung von Autounfällen ist aber von bestimmten Wetterbedingungen (Temperaturen unter dem Gefrierpunkt, vorhandene Feuchtigkeit oder Nässe) sowie der Fahrweise (Geschwindigkeit des Fahrzeugs, Bremsart), der Streckenführung (Kurve), der Fahrbahnbeschaffenheit (glatt, nicht gestreut, etc.), dem Fahrzeug selbst (Konstruktion, Zustand der Bremsen, Reifen etc.) und dem Fahrer (Verfassung, z.B. müde, zerstreut, abgelenkt, nervös, unruhig, hektisch, betrunken) abhängig. Cheng & Novick (1992) verdeutlichen die Abhängigkeit der Kontingenz von den gegebenen Rahmenbedingungen an dem Beispiel zweier Forschergruppen, die in Edinburgh und Los Angeles untersuchen wollen, inwieweit Regen von der Gravitation abhängt. Obwohl beide Gruppen zu dem Ergebnis kommen, dass die Wahrscheinlichkeit von Regen ohne Gravitation gleich Null ist, ermitteln sie unterschiedliche Kontingenzen, da aufgrund der unterschiedlichen Regenmengen in Edinburgh ein sehr starker, und in Los Angeles ein sehr geringer Einfluss von Gravitation gemessen wird. Wenn also Klima und Ort unberücksichtigt bleiben, dann wird der Kausalzusammenhang zwischen Regen und dem eigentlich konstanten Faktor Gravitation unterschiedlich beurteilt. Bei Anwendung einer einfachen Kontingenzformel bleiben unterschiedliche Rahmenbedingungen unberücksichtigt.

Darüber hinaus wurde von Cheng (1997) für die beiden Extremsituationen, dass die Wirkung unabhängig von der An- oder Abwesenheit der Ursache immer ( $P(E|C) = P(E|\neg C) \cong 1$ ) oder gar nicht ( $P(E|C) = P(E|\neg C) \cong 0$ ) auftritt, der spezielle Einwand formuliert, dass Probanden bei nicht kontingenten Ursachen ( $\Delta P = 0$ ) in Abhängigkeit von der Gerichtetheit der Ursache (präventiv oder generativ) unterschiedliche Kausalurteile bilden.

So wird für den einen Extremfall, dass eine Wirkung unabhängig von der An- oder Abwesenheit der Ursache *immer* gegeben ist (Deckeneffekt), der Einfluss einer *generativen* Ursache als unsicher angesehen, weil die Probanden trotz der Nullkontingenz den Eindruck haben, dass sie keine Aussage über die Wirksamkeit der interessierenden Ursache machen können. Beispielsweise würden sich Botaniker über die Wirkung eines Düngemittels im Unklaren befinden, wenn alle Pflanzen mit oder ohne Düngemittel wachsen. Hinsichtlich des in der Kausalliteratur weit seltener untersuchten Einflusses von *präventiven* Ursachen (Shanks, 2002) würden die Probanden hingegen zu dem Schluss gelangen, dass eine nicht kontingente Ursache keinen hemmenden Einfluss auf die Wirkung hat. Ein Beispiel hierfür wäre eine Gruppe von lauter Schmerzpatienten, deren Schmerzen unabhängig von der Einnahme oder Nichteinnahme eines Medikaments gleich bleiben.

Für den anderen Extremfall, dass eine Wirkung unabhängig von der An- oder Abwesenheit einer nicht kontingenten Ursache *gar nicht* gegeben ist (Bodeneffekt), verhält es sich dagegen genau umgekehrt: Hier würden die Probanden nach Cheng (1997) zu dem Urteil kommen, dass eine nicht kontingente *generative* Ursache keinen förderlichen Einfluss auf die Wirkung hätte (Pflanzen wachsen mit oder ohne Düngemittel nicht), während der Einfluss einer *präventiven* Ursache als unsicher oder nicht einschätzbar angesehen werden würde (schmerzfremde Probanden bleiben mit oder ohne Medikament schmerzfrei).

Chengs Annahmen zur unterschiedlichen Beurteilung des Einflusses von nicht kontingenten generativen und präventiven Ursachen in den beiden Extremsituationen werden durch eine Untersuchung von Wu & Cheng (1999), so wie durch Konditionierungsexperimente mit Tieren (Baker et al., 1996; Williams, 1996) gestützt. Die Ergebnisse von Wu & Cheng (1999) zeigen, dass Probanden unter bestimmten Bedingungen nicht kontingente Ursachen nicht als unwirksam einschätzen. Verfügen die Probanden bei der Beurteilung des Kausalzusammenhangs über die Option, Unsicherheit auszudrücken, so erweisen sie sich als sensitiv für den Unterschied zwischen Kovariation und Verursachung.

Kontingenzmodelle und assoziationsistische Modelle wie das RWM (Rescorla & Wagner, 1972; s.a. Kapitel 3.2) hingegen ignorieren den Unterschied zwischen Kovariation und Verursachung und gehen von der falschen und unsinnigen Annahme aus, einen statistischen Zusammenhang als unmittelbares Maß für die Stärke eines Kausalzusammenhangs heranziehen zu können. Kontingenzmodelle machen keine unterschiedlichen Vorhersagen hinsichtlich der Beurteilung von nicht kontingenten Ursachen in Abhängigkeit von der Richtung oder dem Typ (generativ oder präventiv) der Ursache und des Urteilskontextes oder der Situation (Decken- oder Bodeneffekt). Sind Erfolgs- und Basisrate gleich groß ( $P(E|C) = P(E|\neg C)$ ), ergibt sich nach der  $\Delta P$ -Regel eine Kontingenz von Null, d.h. die Probanden müssten unabhängig vom Ursachentyp und Situation den Einfluss nicht kontingenter Ursachen als unwirksam einschätzen.

Die zahlreichen Nachteile, die in Verbindung mit einfachen Kontingenztheorien vorgebracht wurden, haben dazu geführt, dass diese Theorien weiter entwickelt wurden und/oder andere Maße zur Ermittlung der Stärke von einzelnen Ursache-Wirkungs-Zusammenhängen vorgeschlagen wurden.

### Komplexe, multifaktorielle Kontingenztheorien

Komplexe, multifaktorielle Kontingenztheorien gehen davon aus, dass stets mehrere Faktoren kausal relevant für ein Ereignis sind. Dadurch kann es vorkommen, dass die ermittelte Kontingenz nicht den tatsächlichen Kausaleinfluss wiedergibt. Deshalb wurde von einigen

Autoren (s. Hagmayer & Waldmann, 2006, für einen Überblick) vorgeschlagen, weitere kausal relevante Faktoren zu berücksichtigen und an Stelle der unbedingten bedingte Kontingenzen zu verwenden. Bedingte Kontingenz bedeutet, dass bei der Bestimmung der Kontingenz das Vorliegen anderer kausal relevanter Faktoren konstant gehalten wird.

### *Empirische Befunde und Kritik*

Mehrere Studien (s. Hagmayer & Waldmann, 2006, für einen Überblick) haben gezeigt, dass Personen zusätzlich vorliegende Faktoren bei der Bestimmung von Kausalrelationen berücksichtigen. Weitere Untersuchungen deuten aber darauf hin, dass zusätzliche Faktoren nur dann berücksichtigt werden, wenn diese salient und die Anforderungen an die Verarbeitungskapazität gering sind (Fiedler, Brinkmann, Betsch & Wild, 2000; Schaller & O'Brian, 1992; Spellman et al., 2001; Waldmann & Hagmayer, 2001, vgl. Hagmayer & Waldmann, 2006).

Darüber hinaus können multifaktorielle im Gegensatz zu einfachen Kontingenztheorien die im Zusammenhang mit assoziationalistischen Theorien bekannten Blockierungs- und Überschattungseffekte erklären (Waldmann, 1996, 2000, 2001; vgl. auch Hagmayer & Waldmann, 2006). Unter bestimmten Zusatzannahmen liefern multifaktorielle Kontingenztheorien Lösungen für die Probleme der einfachen Kontingenztheorien. Sie haben aber wie diese das Problem, dass kausal relevante nicht von kausal irrelevanten Faktoren unterschieden werden können. Einen Ausweg aus diesem Problem sehen Hagmayer & Waldmann (2006) darin, multifaktorielle Kontingenztheorien durch Annahmen aus der Theorie der Kausalmodelle (Hagmayer, 2001; Hagmayer & Waldmann, 2006; Reips, 1997; Waldmann, 1994, 1996, 2000; Waldmann & Holyoak, 1992) zu ergänzen.

### **3.1.2 Das Probabilistische Kontrastmodell**

Das Probabilistische Kontrastmodell (Cheng & Holyoak, 1995; Cheng & Novick 1991, 1992; Cheng, Park, Yarlas & Holyoak, 1996), kurz PCM (englisch: Akronym für *Probabilistic Contrast Model*), ist ein einfaches kovariationsbasiertes Modell zur Berechnung der Kontingenz zwischen Ursache und Wirkung. Von einfachen Kontingenztheorien unterscheidet sich das PCM durch die Annahme, dass die Bestimmung von Kontingenzen innerhalb einer ausgewählten Teilmenge von Ereignissen vorgenommen wird. Mit multifaktoriellen Kontingenztheorien hat das PCM die Sichtweise gemeinsam, dass am kausalen Geschehen vielfältige Prädiktoren oder Hinweisreize beteiligt sind. Darüber hinaus beschreibt das PCM aber näher, dass es vor der Berechnung des Kontrastes von  $P(E|C)$  und  $P(E|\neg C)$  zu einem kontextsensitiven Auswahlprozess (Oestermeier, 1997) kommt, in dem an

Hand der situativen Umgebung und des inhaltlichen Vorwissens einer Person festgelegt wird, welche Menge von Faktoren und Ereignissen überhaupt als kausal relevant angesehen werden. Cheng & Novick (1991, 1992) sprechen in diesem Zusammenhang von so genannten "focal sets" und meinen damit diejenige ausgewählte Menge von Ereignissen, auf welche Personen bei der Berechnung der Kovariation Bezug nehmen.

Innerhalb des PCM ist die Unterscheidung zwischen Ursachen und ermöglichenden Bedingungen ("enabling conditions") grundlegend. Ermöglichende Bedingungen sind notwendig, damit eine Ursache einen Effekt hervorrufen kann. Sie werden von Cheng & Novick (1991, 1992) folgendermaßen definiert: Bedingung *i* (z.B. Sauerstoff) ist dann eine ermöglichende Bedingung für eine Ursache *j* (z.B. brennende Zigarette), wenn *i* konstanter Bestandteil in einem fokalen Set (z.B. in Verbindung mit der Erklärung von Waldbränden) einer Person ist, aber in einem anderen fokalen Set (z.B. bei Laborversuchen, in denen Sauerstoff nicht konstant vorhanden gegeben ist) mit dem Effekt *e* (Feuer) kovariiert, und wenn *j* in einem weiteren fokalen Set (z.B. in einem luftleeren Raum) nicht mit *e* kovariiert, wenn *i* abwesend ist. Ob ein bestimmtes Ereignis oder eine vorliegende Bedingung eine Ursache, eine ermöglichende Bedingung oder gar kausal irrelevant ist, hängt also vom jeweils ausgewählten fokalen Set ab. Beispielsweise sind Temperaturen unter dem Gefrierpunkt bei der Verursachung eines Autounfalls durch Glatteis auf der Fahrbahn eine ermöglichende Bedingung, beim Gefrieren von Wasser aber eine Ursache und beim Ausbruch einer Leberzirrhose nach einem hohen, langjährigen Alkoholkonsum kausal irrelevant.

Durch die Festlegung des fokalen Sets entscheiden die Personen, welche Ereignisse als mögliche Ursachen näher untersucht werden und welche der vorliegenden Bedingungen als konstante Hintergrundbedingungen aufzufassen sind. Durch die Festlegung eines fokalen Sets wird eine Konditionalisierung der Ursache auf bestimmte zusätzliche Faktoren vorgenommen. Innerhalb komplexer Kausalgefüge im Rahmen eines ausgewählten fokalen Sets wird die Kovariation der möglichen Ursachen mit dem Effekt in Form von Kontrasten bestimmt. Wie bei der Bestimmung eines einfachen Kausalzusammenhang mit Hilfe des Kontingenzmaßes  $\Delta P$ , sind diese Kontraste definiert als die Differenz zwischen der bedingten Auftretenswahrscheinlichkeit eines Effekts nach Erscheinen der möglichen Ursache und der bedingten Auftretenswahrscheinlichkeit eines Effekts ohne vorheriges Erscheinen der möglichen Ursache. Neben Haupteffektkontrasten, mit welchen der Einfluss möglicher Ursachen erfasst werden soll, lassen sich auch Interaktionskontraste ermitteln. Diese geben an, inwieweit der Effekt von der Konjunktion mehrerer Ursachen abhängig ist. (s. Cheng & Novick, 1992, sowie Cheng, 1993, für Details)

*Empirische Befunde und Kritik*

Das PCM kann als eine Erweiterung oder Verallgemeinerung kontingenztheoretischer Ansätze auf Situationen verstanden werden, in denen vielfältige Prädiktoren oder Hinweisreize am kausalen Geschehen beteiligt sind, und wo der Hintergrund für einen bestimmten Zielreiz nicht konstant ist.

Damit aus der Berechnung von Kontingenzen nach der  $\Delta P$ -Regel (s. z.B. Allan, 1980, 1993) ein Modell generiert werden kann, dass in der Lage ist, den Prozess des kausalen Schließens auch in komplexeren Situationen, wie z.B. natürlichen Umwelten, abbilden zu können, müssen Zusatzannahmen gemacht werden. Dabei erweist sich das Konzept der fokalen Sets als sehr bedeutsam, weil durch diese Zusatzannahme Situationen erklärbar oder vorhersagbar werden, in denen das menschliche Denken scheinbar von statistisch angemessenen Strategien wie der einfachen  $\Delta P$ -Regel abweicht (Baker et al., 1996).

Das PCM zeichnet sich dadurch aus, dass es alle Befunde (wie z.B. Blockierung, Überschattung, Berücksichtigung zusätzlicher Ursachefaktoren) vorhersagen kann, die auch durch multifaktorielle Kontingenztheorien vorher gesagt werden können (s. Cheng, 1993, 1997; Cheng & Holyoak, 1995; Cheng & Novick, 1992; Cheng, Park, Yarlas & Holyoak, 1996). Das PCM ist dazu in der Lage, weil die innerhalb eines fokalen Sets bestimmbar Kontraste den bedingten Kontingenzen zwischen den jeweiligen Ursachefaktoren und dem Effekt ereignis entsprechen.

Untersuchungen zum PCM (Cheng & Novick, 1990, 1992; Cheng & Holyoak, 1995) nehmen vor allem auf Situationen Bezug, in denen die Kontingenzzurteile von Probanden von den Vorhersagen einer einfachen  $\Delta P$ -Regel abweichen. In einer Reihe von Arbeiten aus Sozialpsychologie (Nisbett & Ross, 1980), Philosophie (Mackie, 1974) und kognitiver Psychologie (Jenkins & Ward, 1965; Schustack & Sternberg, 1981; Shaklee & Tucker, 1980) konnten Abweichungen von statistischen Vorgehensweisen beim kausalen Schlussfolgern belegt werden und boten zu Zweifeln an der These Anlass, dass kausale Schlussfolgerungen auf statistischen Wahrscheinlichkeitsberechnungen beruhen. Cheng & Novick (1990, 1991, 1992) versuchten, der Ansicht entgegenzutreten, dass kausales Schlussfolgern sich eher als eine Ansammlung von unterschiedlichen Heuristiken und systematischen Fehlerverzerrungen (englisch: *biases*) charakterisieren lässt. Sie führen die Abweichungen von normativen kontingenzbasierten Ansätzen im Einklang mit dem PCM darauf zurück, dass sich die systematischen Berechnungen der Probanden nicht auf dieselben fokalen Sets von Ereignissen bezogen haben, wie es die Untersucher angenommen hatten. Darüber hinaus konnten Cheng & Novick (1990) den Nachweis erbringen, dass Probanden dazu in der Lage sind,



angemessene Kausalattributionen ohne Bias vorzunehmen, wenn sie über Informationen bezüglich der Wahrscheinlichkeit eines Effektereignisses bei allen möglichen Ursache kombinationen verfügen.

Mit Hilfe unterschiedlicher fokaler Sets erklärt das PCM auch die Unterscheidung, die Personen zwischen einer ermöglichenden Bedingung und einer Ursache vornehmen. Cheng & Novick (1991) setzten ihren Probanden vier mögliche Ursachefaktoren für einen Effekt vor. Zwei Faktoren waren konstant gegeben, wobei der eine relevant und der andere irrelevant für das kausale Geschehen war. Der dritte Faktor wies eine positive, der vierte eine negative Kontingenz im Verhältnis zum Effekt auf. Dabei waren die Probanden in der Lage, den relevanten, konstant vorhandenen Faktor als notwendige Bedingung anzusehen und den relevanten, variierenden Faktor je nach Kontrast als Ursache oder Inhibitor. Der irrelevante Faktor wurde als irrelevant beurteilt. Assoziationistische Modelle oder einfache kontingenztheoretische Ansätze können solche Ergebnisse nicht vorhersagen.

Shanks (1993; Shanks et al., 1996) wendet gegen die Focal Set-Theorie der  $\Delta P$ -Regel ein, dass diese den Kontingenzzurteilen der Probanden nur in einigen, aber bei weitem nicht in allen Fällen gerecht wird. Während Shanks (1993; Shanks et al., 1996) davon ausgeht, dass die Kontingenzzurteile besser durch assoziationistische Mechanismen erklärt werden können, vertritt Spellmann (1996) ähnlich wie Cheng und ihre Mitarbeiter (Cheng & Novick, 1990, 1992; Cheng & Holyoak, 1995; Cheng, Park, Yarlus & Holyoak, 1996) die Ansicht, dass im Falle der Existenz mehrerer möglicher Ursachen Abweichungen von der  $\Delta P$ -Regel das Ergebnis eines normativen Prozesses sind, bei welchem die Kausalurteile im Hinblick auf die Existenz der anderen möglichen Ursachen relativiert werden, indem bedingte an Stelle von unbedingten Kontingenzen berechnet werden.

Von einigen Autoren (z.B. Shanks, 1993; Young, 1995) wurde die Annahme fokaler Sets kritisiert, weil innerhalb des PCM die Kriterien für die Bildung, die Auswahl und den Gebrauch eines oder mehrerer fokaler Sets relativ wenig festgelegt sind. Dadurch ist es praktisch unmöglich, die Theorie gegenüber anderen Theorien zu überprüfen. So ist beispielsweise bei der Auswahl von fokalen Sets davon auszugehen, dass Probanden nur Ereignisse berücksichtigen, die ihnen auch bekannt sind. Dementsprechend unterschiedlich kann der Auswahlprozess von Individuum zu Individuum ausfallen.

Chengs (1997) spezieller Einwand, dass Probanden bei einer Kontingenz von Null ( $\Delta P = 0$ ) in Abhängigkeit vom Kontext, in dem die Nullkontingenz auftritt, unterschiedliche Kausalurteile bilden, trifft nicht nur kontingenztheoretische Modelle im allgemeinen, sondern auch das PCM im Besonderen, da auch mit Hilfe des PCM keine unterschiedlichen Vorhersagen für

nicht kontingente generative oder präventive Ursachen bei extrem hohen ( $P(E|C) \cong 1$ ) oder niedrigen ( $P(E|C) \cong 0$ ) Basisraten gemacht werden.

### 3.1.3 Die Power PC Theorie

Die so genannte Causal Power Theorie des Probabilistischen Kontrastmodells (englisch: *the causal power theory of the probabilistic contrast model*, Cheng, 1997; Novick & Cheng, 2004), kurz *Power PC Theorie* oder *Theorie der kausalen Power*, baut auf der Kritik an kontingenztheoretischen Ansätzen auf und ist eine Weiterentwicklung des PCM. Sein theoretisches Fundament erhält die Power PC Theorie aber durch den Power-Ansatz. Dieser wiederum hat seine Wurzeln in der kantischen Vorstellung, dass irgendeine Form von a priori Wissen gegeben sein muss, welches es uns ermöglicht, bestimmte Informationen über Ereignisabfolgen kausal zu interpretieren. Kant zu Folge besteht dieses a priori Wissen darin, dass alle Ereignisse eine Ursache haben. Von psychologischer Seite (s. Cheng, 1997, für einen Überblick) wurde die kantische Sichtweise oft dahin gehend interpretiert, dass Menschen nur dann ein Ereignis, ein Objekt oder einen Sachverhalt als Ursache von einem anderen ansehen, wenn sie über Wissen oder über eine Wahrnehmung von einer spezifischen erzeugenden (generative) Quelle, von einem kausalen Mechanismus oder einer kausalen Kraft (englisch: *causal power*) verfügen, durch welchen oder welche die in Frage kommende Ursache mit der Wirkung verknüpft wird. Kausale Kraft oder Power ist der "intuitive Begriff" (Cheng, 1997) dafür, dass ein Objekt, ein Ereignis oder ein Sachverhalt etwas anderes verursacht, indem dabei eine Kraft oder Energie von der Ursache auf die Wirkung ausgeübt wird. Entsprechend dem Power-Ansatz ist es nicht einfach so, dass Ursachen ihren Wirkungen zeitlich voraus gehen, sondern eher so, dass Ursachen ihre Wirkungen erzeugen oder hervorbringen.

Nach Cheng (1997) hat der reine Power-Ansatz für sich genommen zwei zentrale Schwachstellen: Zum einen kann die Stärke eines kausalen Zusammenhanges nicht berechnet werden, zum anderen sind die für den Power-Ansatz grundlegenden Begriffe wie "Kraft", "Energie", "Mechanismus", "Erzeugung" oder "hervorbringen" zirkulär, weil sie Wissen über kausale Regularitäten und Kausalität voraussetzen (zur Zirkularität des Power-Ansatzes s. Cheng, 1993; Oestermeier, 1997). Der Power-Ansatz geht von der Annahme aus, dass Begriffe wie "Kraft" oder "Power" viel grundlegender für kausale Schlüsse sind als Begriffe wie "statistische Bedeutsamkeit" (englisch: *statistical relevance*), "statistischer Zusammenhang" oder "Kausalität". Dieser Sichtweise, dass kausale Schlussfolgerungen auf unserem Wissen über kausale Kräfte basieren, entgegnet Cheng (1997), dass "kausale Kräfte"

per definitionem kausal sind und Wissen über kausale Kräfte bereits Wissen über Kausalität voraussetzt.

Die Power PC Theorie wurde von Cheng (1997) entwickelt, um einige der Probleme zu überwinden, die sowohl mit dem Power-Ansatz als auch mit dem PCM verbunden sind. In Verbindung mit dem Power-Ansatz können sowohl die fehlende Möglichkeit, die Stärke eines Kausalzusammenhangs zu berechnen als auch der Zirkularitätsvorwurf durch die Zusatzannahme überwunden werden, dass zwischen dem Power-Ansatz und einer kontingenzbasierten Bestimmung der Stärke von Ursache-Wirkungs-Zusammenhängen kein Widerspruch bestehen muss. In der Power PC-Theorie wird der Power-Ansatz als theoretische Erklärung für eine formalisierte Berechnung der kausalen Stärke implementiert ("*causal power is to covariation as theory is to model*", Cheng, 1997, S. 369). Nach Cheng (1997) liefert der Power-Ansatz die theoretische Grundlage für ein Verständnis darüber, ob kovariierende Ereignisse in einem kausalen Zusammenhang stehen oder nicht. Im Gegensatz dazu ignoriert das PCM die Tatsache, dass Menschen intuitiv wissen, dass Kovariation nicht Kausalität beinhalten muss. Cheng (1997) zu Folge setzen Menschen beobachtete Kovariationen nicht einfach mit kausalen Beziehungen gleich, sondern sie interpretieren diese mechanistisch im Sinne des Power Ansatzes als Manifestationen für das Wirksamwerden von unsichtbaren kausalen Kräften (englisch: *causal powers*). Menschen wissen intuitiv, dass der Sonnenaufgang nicht durch das Krähen eines Hahnes auf dem Bauernhof verursacht wird, obwohl beide Ereignisse konstant miteinander verknüpft sind und die Dämmerung dem Sonnenaufgang zeitlich immer vorausgeht.

Um die Stärke dieser verborgenen Kausalkräfte auch rechnerisch einschätzen zu können, hat Cheng (1997) das kontingenzbasierte PCM weiter entwickelt und mathematische Formeln zur Berechnung der kausalen Stärke von generativen und präventiven Ursachen bereit gestellt. Die Berechnung der kausalen Stärke bezieht dabei die Berechnung des Kontingenzmaßes  $\Delta P$  mit ein und erfolgt je nach Ursachentyp verschieden, weil für generative und präventive Ursachen teilweise unterschiedliche Vorhersagen gemacht werden (Cheng, 1997). Weil sich zudem unterschiedliche Basisraten unterschiedlich auf die Beurteilungen der Stärke von Ursachen auswirken (z.B. Dichte-Bias), wird das Kontingenzmaß  $\Delta P$  bei der Berechnung der kausalen Stärke für generative und präventive Ursachen um den Einfluss der jeweils beobachteten Basisrate korrigiert.

Für generative Ursachen ( $\Delta P > 0$ ) gilt, dass es bei wachsender Basisrate immer schwieriger wird, die tatsächliche Stärke einer Ursache zu beurteilen. Tritt der Effekt auch ohne die in Frage kommende Ursache immer auf ( $P(E|\neg C) = 1$ ), kann der Beurteiler kein Urteil über

deren Stärke bilden. Um diesen Einfluss der Basisrate bei der Berechnung der kausalen Stärke berücksichtigen zu können, wird folgende Gleichung eingesetzt:

$$p = \frac{\Delta P}{1 - (E | -C)} \quad (4)$$

Das Maß  $p$  ( $p$  ist das Kürzel für das englische *power*) steht für die kausale Stärke oder Kraft der Ursache  $C$ . Gleichung (4) veranschaulicht, dass  $p$  bei konstantem  $\Delta P$  und anwachsender Basisrate wegen des kleiner werdenden Nenners größer wird. Erreicht  $P(E|-C)$  bei Nullkontingenz ( $\Delta P = 0$ ) einen Wert von 1, lässt sich  $p$  nicht mehr eindeutig bestimmen, weil der Werte im Zähler und Nenner Null betragen. Bei einer Basisrate von Null, kommen das PCM und die Power PC Theorie zu denselben Vorhersagen, weil der Wert im Nenner Eins beträgt.

Für präventive Ursachen ( $\Delta P < 0$ ) gilt hingegen, dass es bei kleiner werdenden Basisraten immer schwieriger wird, die tatsächliche Stärke einer Ursache zu beurteilen. Ist die Basisrate  $P(E|-C) = 0$ , ist es einem Beurteiler überhaupt nicht mehr möglich, den Einfluss einer präventiven Ursache zu ermitteln. Zur Bestimmung der Stärke einer präventiven Ursache hat Cheng (1997) die folgende Formel aufgestellt:

$$p = \frac{\Delta P}{P(E | -C)} \quad (5)$$

Bei konstantem  $\Delta P$  und kleiner werdender Basisrate wächst die präventive Kraft  $p$ , wie ein das Einsetzen entsprechender Werte in Formel (5) vor Augen führt. Erreicht  $P(E|-C)$  bei Nullkontingenz ( $\Delta P = 0$ ) einen Wert von 0, lässt sich  $p$  nicht mehr eindeutig definieren, weil auch hier die Werte in Zähler und Nenner jeweils Null betragen. Um es in einem Beispiel zu formulieren: Wenn eine Krankheit ohne Impfstoff nie auftritt ( $P(E|-C) = 0$ ), so kann auch nicht festgestellt werden, ob ein Impfstoff das Auftreten einer Krankheit verhindern würde. Bei einer Basisrate  $P(E|-C) = 1$  kommen das PCM und die Power PC Theorie zu denselben Vorhersagen.

Für den Fall, dass die Anzahl der Fälle bei gegebener und nicht gegebener Ursache gleich groß ist ( $P(C) = P(-C)$ ), können die Stärken für generative und präventive Ursachen wesentlich einfacher als in den Gleichungen (4) und (5) an Hand der Zelhäufigkeiten in den Gleichungen (6) und (7) berechnet werden:

$$p_{\text{generativ}} = (a - c) / d \quad (6)$$

$$p_{\text{präventiv}} = (a - c) / c \quad (7)$$

Die konzeptuellen wie rechnerischen Unterschiede zwischen der Power PC Theorie und dem PCM lassen sich an einem einfachen Zahlenbeispiel (s.a. Hagmayer, 2001), in dem bewusst eine hohe Basisrate zu Grunde gelegt wurde, verdeutlichen: Wenn eine Wirkung auch ohne Einfluss der Zielursache häufig auftritt (d.h. die Basisrate  $P(E|-C)$  ist hoch), dann kann die Zielursache nur selten ihre Wirkung zeigen. Tritt die Wirkung beispielsweise in 8 von 10 Fällen ohne die Zielursache auf ( $P(E|-C) = 0.8$ ), dann hat die Zielursache nur in zwei Fällen Gelegenheit, ihre Wirkung zu zeigen. Wenn davon ausgegangen wird, dass die Wirkung also bei gegebener Zielursache in 9 von 10 Fällen eintritt ( $P(E|C) = 0.9$ ), so bedeutet dies, dass die Zielursache ihre Wirkung in einem von zwei Fällen, in denen sie ihre Wirkung unter Beweis stellen konnte, auch tatsächlich gezeigt hat. Während die Kontingenzt in diesem Rechenbeispiel mit einem  $\Delta P$  von 0.1 sehr schwach ist, ergibt sich für die kausale Stärke  $p$  ein Maß von 0.5 (1 : 2).

Für den Fall, dass mehrere potentielle Ursachen in Betracht zu ziehen sind, nimmt die Theorie der kausalen Power an, dass diese Ursachen unabhängig voneinander sind und auch unabhängig auf den Effekt wirken. Ob diese Annahmen von Probanden auch tatsächlich so gemacht werden, wurde bislang noch nicht untersucht (Hagmayer & Waldmann, 2006). Novick & Cheng (2004) haben die Theorie der kausalen Power für Interaktionen zwischen zwei Ursachen vorgeschlagen, wodurch die Beschränkung auf unabhängige Ursachen aufgegeben wurde. Von Lien & Cheng (2000) wurden Vorschläge entwickelt, wie Erkenntnisse über den Einfluss von kausalem Vorwissen durch eine Erweiterung der Theorie erklärt werden können.

### *Empirische Befunde und Kritik*

Wie die kontingenzttheoretischen Ansätze und das PCM, so ist auch die Theorie der kausalen Power eine normative Theorie, die auf wahrscheinlichkeitstheoretischen Überlegungen beruht. Damit aus empirischen Daten die kausale Stärke einer Ursache geschätzt werden kann, muß eine Reihe von Vorannahmen erfüllt sein. Beispielsweise ist die Bestimmung der kausalen Power an Rahmenbedingungen wie die Abwesenheit von Decken- oder Bodeneffekten gebunden. So sagt die Theorie der kausalen Power vorher, dass die Stärke einer generative Ursachen nicht bestimmt werden kann, wenn der Effekt auch in Abwesenheit der Ursache immer gegeben ist ( $P(E|-C) = 1$ ). Ebenso lässt sich die kausale Power einer präventiven Ursache nicht bestimmen, wenn der Effekt in Abwesenheit der hemmenden Ursache nie auftritt ( $P(E|-C) = 0$ ). Wu & Cheng (1999) konnten für solche Extremsituationen

empirisch nachweisen, dass die Probanden trotz der Nullkontingenz den Eindruck haben, dass sie den Kausaleinfluss der interessierenden Ursache nicht sicher einschätzen können, sofern sie über die Option verfügen, Unsicherheit auszudrücken.

Ein Vorzug der Theorie der kausalen Power besteht darin, dass sie im Gegensatz zu kontingenztheoretischen Ansätzen den empirisch immer wieder nachgewiesenen Density-Bias (Allan, 1993; Allan & Jenkins 1980, 1983) bei generativen Ursachen gut erklären kann. Dieser besagt, dass Kausalzusammenhänge mit zunehmender Wahrscheinlichkeit des Effekts stärker beurteilt werden. Setzt man zwei unterschiedlich hohe Basisraten im Nenner von Gleichung (4) ein, so ergibt sich bei gleicher Kontingenz im Zähler und höherer Basisrate im Nenner ein höherer Wert für die kausale Stärke  $p$ .

Weiter sagt die Theorie der kausalen Power bei konstant gehaltenem  $\Delta P$  für generative und präventive Ursachen einen entgegen gesetzten Einfluss der Basisrate voraus. Während bei generativen Ursachen höhere Basisraten zu höheren Werten für die kausale Stärke  $p$  führen, verhält es sich bei präventiven Ursachen genau umgekehrt: Hier ziehen geringere Basisraten höhere Werte für die kausale Stärke  $p$  nach sich, wie ein Einsetzen unterschiedlich hoher Basisraten im Nenner von Gleichung (5) veranschaulicht.

Einen unterschiedlichen Basisrateneinfluss in Abhängigkeit von der Gerichtetheit der Ursache (generativ oder präventiv) sagen weder das RWM noch kontingenztheoretische Ansätze vorher. Buehner & Cheng (1997, s.a. Buehner, 2001) verglichen nun die Theorie der kausalen Power mit dem RWM und dem PCM, indem sie das Kontingenzmaß  $\Delta P$  über verschiedene Bedingungen bei variierender kausaler Stärke hinweg konstant hielten. Für generative Ursachen fanden sie dabei, dass die Probanden die Kausalstärken bei steigenden Basisraten höher einschätzten. Buehner, Cheng & Clifford (2003, Experimente 2 bis 4) konnten mit der gleichen Vorgehensweise gegensätzliche Basisrateneffekte für generative und präventive Ursachen finden. Für generative Ursachen ergaben sich bei höheren Basisraten höhere Beurteilungen der Kausalstärken, bei präventiven Ursachen sanken die Beurteilungen der Kausalstärken bei höheren Basisraten. All diese Ergebnisse werden von der Power PC Theorie ebenso vorhergesagt wie der Befund, dass unterschiedliche Basisraten bei Nullkontingenzen keinen Einfluss auf die Kausalurteile haben.

Andere Studien, in denen ebenfalls die Theorie der kausalen Power mit anderen Ansätzen verglichen wurde, kamen allerdings zu Ergebnissen, die der Theorie der kausalen Power widersprechen (Allan, 2003; Lober & Shanks, 2000; Perales & Shanks, 2003; Vallée-Tourangeau, Murphy & Drew, 1997; Vallée-Tourangeau, Murphy, Drew & Baker 1998; White, 2002c, 2003c, 2004). Die Befunde von Vallée-Tourangeau et al. (1997, 1998) stützen

weder die Theorie der kausalen Power, noch das PCM und das RWM. Anders als Buehner & Cheng (1997) variierten diese Autoren nicht das Maß für die kausale Stärke  $p$  und hielten  $\Delta P$  konstant, sondern hielten  $p$  konstant und variierten  $\Delta P$ . Lober & Shanks (2000), Perales & Shanks (2003), so wie Buehner et al. (2003) setzten beide Paradigmen ein. Dabei kamen Lober & Shanks (2000) in 5 ihrer 6 Experimente zu dem Schluss, dass die Probanden eher Kontingenzen oder das RWM für die Bestimmung der kausalen Stärke einer Ursache heranziehen. Lediglich die Ergebnisse von Experiment 3, welches einen ähnlichen Aufbau hatte wie das von Buehner & Cheng (1997) beschriebene Experiment 1, stimmten mit den Vorhersagen der Power PC Theorie überein. In den Studien von Perales & Shanks (2003) wurde die Theorie der kausalen Power nur durch die Ergebnisse von Experiment 1 gestützt, auch hier wurde  $\Delta P$  konstant gehalten und  $p$  variiert. Dagegen widersprechen die Ergebnisse aus den Experimenten 2 und 3 den Vorhersagen der Theorie der kausalen Power klar (s.a. Experiment 1 von Buehner et al., 2003), unter anderem weil die Theorie der kausalen Power keinen Density-Effekt (Allan, 1993; Allan & Jenkins, 1983) bei Nullkontingenzen vorhersagt. Insgesamt kommen Perales & Shanks (2003) zu dem Schluss, dass ihre Ergebnisse am besten durch das assoziationsistische Modell von Pearce (1987) und das Meinungsrevisionsmodell von Catena et al. (1998) vorhergesagt werden. Beide Modelle haben den Vorzug, dass sie den Einfluss unterschiedlicher Basisraten sowohl in Abhängigkeit von der Gerichtetheit der Ursache (generativ oder präventiv) als auch bei nicht kontingenten Ursachen vorhersagen können.

Buehner et al. (2003) replizierten in ihrem Experiment 1 die Ergebnisse der Experimente 2 und 3 von Perales & Shanks (2003). Anders als Perales & Shanks (2003) sehen sie die Theorie der kausalen Power dadurch aber nicht als widerlegt an. Sie führten das Scheitern der Theorie der kausalen Power auf zwei Gründe zurück: Zum einen auf Ambivalenzen in der Aufgabenstellung für die Probanden, zum anderen auf Gedächtnisbegrenzungen durch eine fehlerhafte Darbietung der Informationen über Ursache-Wirkungs-Zusammenhänge. In den Experimenten 2 bis 4 wurde die Gedächtnisbeanspruchung durch die Wahl einer summarischen Darbietungsform reduziert, zudem wurde eine grafische Darstellungsform gewählt, um die Datenpräsentation so klar und verständlich wie möglich zu machen. Bezüglich der Aufgabenstellung sollten die Probanden keine vage Einschätzung der Stärke eines Kausalzusammenhangs abgeben, sondern sie wurden kontrafaktisch danach befragt, in wie vielen von 100 Fällen, in denen die Wirkung nicht zu Tage trat, nun die Wirkung auftreten würde, wenn die fragliche Ursache eingeführt werden würde. Durch diese Veränderungen wollten Buehner et al. (2003) herausfinden, ob Menschen unter so weit wie

möglich "störungsfreien" Bedingungen dazu in der Lage sind, kausale Beziehungen entsprechend den Vorhersagen der Theorie der kausalen Power zu entdecken. Ungeachtet dieser im Grunde zirkulären Vorgehensweise (die Aufgabenstellung gibt den Probanden vor, die Ergebnisse zu zählen und den Einfluss der fraglichen Ursache zur Basisrate in Beziehung zu setzen und auf eine Skala von 0 bis 100 zu übertragen), kamen Buehner et al. (2003) zu dem Ergebnis, dass die Basisrate entsprechend den Vorhersagen der Theorie der kausalen Power bei  $\Delta P = 0$  keinen Einfluss auf die Beurteilung der Kausalstärken durch die Probanden hatte. Zudem hatte - anders als in den Studien von Lober & Shanks (2000), Perales & Shanks (2003), Vallée-Tourangeau et al. (1997, 1998) - die Variation von  $\Delta P$  bei konstanter kausaler Power ebenfalls keinen Einfluss auf die Kausalurteile. Die Ergebnisse in den Experimenten 2 bis 4 können nur durch die Theorie der kausalen Power vollständig vorhergesagt werden; alternative Ansätze wie ein klassisch kontingenztheoretisches Modell, eine gewichtete  $\Delta P$ -Regel (Anderson & Sheu, 1995), verschiedene assoziationsistische Modelle (RWM, Modifikation des RWM durch Van Hamme & Wasserman (1994), Pearce-Modell) und das lineare Kombinationsmodell von Schustack & Sternberg (1981) scheiterten.

Allan (2003) wendet in einem kritischen Überblick gegen Cheng (1997) ein, dass diese zur Stützung ihrer Theorie der kausalen Power Studien herangezogen habe, die entweder in der Tat die Theorie der kausalen Power widerlegen (Allan & Jenkins, 1983) oder bislang noch niemals veröffentlicht wurden (Fratianne & Cheng, 1995; Park & Cheng, 1995). Zudem seien andere Studien, die die Theorie der kausalen Power sowieso widerlegen würden (z.B. Kao & Wasserman, 1993; Levin et al., 1993; Wasserman et al., 1990), von Cheng (1997) erst gar nicht erwähnt worden. Nach Allan (2003) weichen sogar selbst die veröffentlichten Ergebnisse von Cheng und ihren eigenen Mitarbeitern (Buehner, 2001; Buehner & Cheng, 1997; Wu & Cheng, 1999) teilweise von den Vorhersagen der Theorie der kausalen Power ab. Außerdem, so argumentiert Allan (2003) weiter, kann die Theorie der kausalen Power Befunde nicht vorhersehen, die besagen, dass bei einer Einzelfalldarbietung die Stärke der Kausalurteile in Abhängigkeit vom Zeitpunkt der Urteilsbildung (Allan, 1993; Shanks, 1985, 1987, 1993) und von der Abfolge der einzelnen Durchgänge (Dennis & Ahn, 2001; López, Shanks, Almaraz & Fernández, 1998; Yates & Curley, 1986) variiert.

Der Überblick über die empirischen Studien zeigt, dass durch die Power PC Theorie wie schon zuvor durch die  $\Delta P$ -Regel die Ergebnisse zur kausalen Urteilsbildung nicht vollständig vorhergesagt werden können. Griffiths & Tenenbaum (2005) weisen zudem auf Ergebnisse von White (1998, 2002c, 2003c) hin, die belegen, dass menschliche Kausalurteile in Abhängigkeit von der Anzahl beobachtbarer Fälle in einer Vierfeldertafel variieren können.



Da sowohl für die Berechnung der kausalen Stärke  $p$  als auch bei der Berechnung des Kontingenzmaßes bedingte Wahrscheinlichkeiten ermittelt werden, können weder die  $\Delta P$ -Regel, noch die Power PC Theorie diesen Effekt in Abhängigkeit von der Anzahl beobachteter Fälle erklären. Ebenso wenig können die  $\Delta P$ -Regel oder die Power PC Theorie erklären, dass Probanden sich auch in Situationen, in denen nicht genügend Informationen zur Berechnung der bedingten Wahrscheinlichkeiten verfügbar sind, als fähig erweisen, kausale Zusammenhänge aus unvollständigen Datensätzen einzuschätzen (Griffiths & Tenenbaum, 2005). Anderson & Sheu (1995) konnten zudem zeigen, dass Menschen zum Erfassen kausaler Zusammenhänge keine Kontingenzen benötigen, sondern dass dafür auch schon einfache Verhältnisdaten genügen. Weder die  $\Delta P$ -Regel, noch die Power PC Theorie lassen sich auf Verhältnisdaten anwenden (Griffiths & Tenenbaum, 2005). Ausgehend von ihrer Kritik haben Griffiths & Tenenbaum (2005) ein neues Modell, das *Causal Support Modell*, entwickelt, welches vorhersagen soll, wie Menschen Wissen über die Existenz eines monokausalen Zusammenhangs erwerben. Das *Causal Support Modell* basiert auf Bayes-Netzwerken (Gopnik, Glymour, Sobel, Schulz, Kushnir & Danks, 2004; Gopnik, Sobel, Schulz & Glymour, 2001; Pearl, 2000; Spirtes, Glymour & Scheines, 1993) und ist im Gegensatz zur  $\Delta P$ -Regel oder der Power PC Theorie dazu in der Lage, sowohl die Abhängigkeit menschlicher Kausalurteile von der Anzahl beobachteter Fälle, als auch Kausalurteile bei unvollständigen Datensätzen und aus Verhältnisdaten zu bilden. Der zentrale Unterschied zwischen der  $\Delta P$ -Regel und der Power PC Theorie einerseits und dem *Causal Support Modell* andererseits besteht nach Griffiths & Tenenbaum (2005) darin, dass mit Hilfe der  $\Delta P$ -Regel und der Power PC Theorie die Stärke eines Ursache-Wirkungs-Zusammenhangs ermittelt wird, während mit Hilfe des *Causal Support Modells* die kausale Struktur eingeschätzt werden kann: Anders als bei der  $\Delta P$ -Regel oder der Power PC Theorie wird also in Verbindung mit dem *Causal Support Modell* nicht danach gefragt, wie stark ein Kausalzusammenhang ist, sondern danach, ob ein Kausalzusammenhang existiert. In ihren eigenen Untersuchungen kamen Griffiths & Tenenbaum (2005) zu dem Ergebnis, dass Menschen eher Wissen über kausale Strukturen erwerben und danach fragen, ob ein Kausalzusammenhang existiert.

Zusammenfassend lässt sich feststellen, dass die Power PC Theorie von Cheng (1997; Novick & Cheng, 2004) seit ihrem Erscheinen im Jahr 1997 eine breite Resonanz gefunden und viel Anlass zu kontroversen Diskussionen (s.a. Cheng & Novick, 2005; Luhmann & Ahn, 2005; White, 2005b) geboten hat. Es ist damit zu rechnen, dass die Power PC Theorie auch in den kommenden Jahren empirische und theoretische Kontroversen hervorrufen wird. Wie bei

anderen theoretischen Ansätzen, ist die empirische Befundlage zur Power PC Theorie insgesamt widersprüchlich, so dass nicht davon ausgegangen werden kann, dass Menschen unter allen Umständen Kausalstärken entsprechend den Vorhersagen der Theorie der kausalen Power bestimmen. Die Studien von Buehner, Cheng & Clifford (2003) belegen in unseren Augen eindrucksvoll, wie durch bestimmte Manipulationen von Aufgabenstellung, Darbietungs- und Präsentationsform die Vorhersagen der Theorie der kausalen Power erfüllt werden können. Möglicherweise lassen sich also die widersprüchlichen Befunde zu den verschiedenen theoretischen Ansätzen zumindest teilweise darauf zurückführen, dass den Probanden in verschiedenen Untersuchungen unterschiedliche experimentelle Rahmenbedingungen vorgegeben wurden.

### **3.1.4 Bestimmung der kausalen Stärke nach der Theorie der empirischen Evidenzevaluation**

In der Theorie der Evidenzevaluation von White (2000b, 2002a, 2002c, 2003a, 2003b, 2003c, 2004, 2005a) werden Informationen über das gemeinsame Auftreten von Ursachen und Effekt als Evidenzen konzipiert. Einzelfälle können entweder als Bestätigung (konfirmatorische Information) oder als Nicht-Bestätigung (diskonfirmatorische Information) für die Hypothese angesehen werden, dass die potentielle Ursache das Auftreten des Effekts verursacht. Der Theorie der empirischen Evidenzevaluation (kurz: EE-Theorie) zufolge werden Kausalurteile aus dem Gewicht der Evidenz abgeleitet. Demnach wird ein Kausalzusammenhang stärker beurteilt, wenn der Anteil der konfirmatorischen im Verhältnis zur diskonfirmatorischen Evidenz anwächst. Wie bei der  $\Delta D$ -Regel auch, werden Fälle, in denen Ursache und Effekt gemeinsam auftreten oder gemeinsam ausbleiben, als konfirmatorisch gewertet (Zellen a und d aus der Vierfeldertafel in Abbildung 1), und Fälle, in denen entweder nur die Ursache oder nur der Effekt auftreten, als diskonfirmatorisch angesehen (Zellen b und c aus der Vierfeldertafel in Abbildung 1). Im Gegensatz zur  $\Delta D$ -Regel wird die Differenz zwischen konfirmatorischer und diskonfirmatorischer Evidenz aber noch zusätzlich zur Anzahl aller Einzelfallbeobachtungen ins Verhältnis gesetzt. Formal lässt sich der Anteil konfirmatorischer Evidenz folgendermaßen berechnen:

$$p_{CI} = \frac{(a + d - b - c)}{(a + d + b + c)} \quad (8)$$

Das Maß  $pCI$  repräsentiert den Anteil konfirmatorischer Beobachtungen (englisch: *proportion of confirmatory instances*). Der Anteil konfirmatorischer Evidenz kann der Theorie zufolge für drei Arten von Urteilen heran gezogen werden: (i) Urteile über das Zutreffen der Hypothese, dass ein Kausalzusammenhang vorliegt; (ii) Urteile über die Wahrscheinlichkeit, dass die Ursache den Effekt hervorruft bzw. in einem bestimmten Fall hervorgerufen hat; (iii) Urteile über die Stärke des Kausalzusammenhangs.

In jüngerer Zeit hat White (2003a, 2003c, 2004) eine Modifikation der  $pCI$ -Regel vorgeschlagen. In Gleichung (8) werden die vier Zellen a, b, c und d gleich gewichtet. Um dem immer wieder replizierten Befund Rechnung zu tragen, dass die vier Zellen einer 2 x 2 Kontingenztafel bei der kausalen Urteilsbildung ungleich gewichtet werden (Anderson & Sheu, 1995; Kao & Wassermann, 1993; Mandel & Lehmann, 1998; Wasserman et al. 1990, 1993, 1996), wurde eine verallgemeinerte Version der  $pCI$ -Regel entwickelt, die in Gleichung (9) formalisiert ist:

$$pCI = \frac{w_a[s(x_a)] + w_d[s(x_d)] - w_b[s(x_b)] - w_c[s(x_c)]}{(a + d + b + c)} \quad (9)$$

Bei den Größen  $s[x_a]$ ,  $s[x_b]$ ,  $s[x_c]$  und  $s[x_d]$  handelt es sich um subjektive Evaluationen der Zellen a, b, c und d, die Werte zwischen +1 und -1 annehmen können.  $w_a$ ,  $w_b$ ,  $w_c$  und  $w_d$  sind Gewichte, die den subjektiven Evaluationen  $s[x_a]$ ,  $s[x_b]$ ,  $s[x_c]$  und  $s[x_d]$  zugeschrieben werden und ebenfalls +1 und -1 variieren können. Durch diese verallgemeinerte Modellversion können individuelle Unterschiede bei der kausalen Urteilsbildung besser vorher gesagt werden (White, 2003b). White (2000a) konnte nachweisen, dass individuelle Urteilstendenzen sehr gut durch idiosynkratische Regeln vorhergesagt werden konnten, die in vielen Fällen signifikant von normativen Regeln wie der  $\Delta P$ -Regel abweichen.

Die ungewichtete  $pCI$ -Regel und die  $\Delta P$ -Regel machen dieselben Vorhersagen, wenn die Wahrscheinlichkeiten des Auftretens und des Nicht-Auftretens der Ursache gleich groß sind (White, 2003c), d.h. wenn  $p(C) = p(-C) = 0,5$ . Nur bei einer Abweichung von dieser Ursachenrate werden die Vorhersage-Unterschiede zwischen den beiden Regeln deutlich (White, 2004).

*Empirische Befunde und Kritik*

Für das EE-Modell spricht nach White (2005a), dass es eine Reihe von Phänomenen zur kausalen Urteilsbildung vorhersagen kann. So wurde das Modell in Studien zur kausalen Urteilsbildung mit einer Ursache (White 2002c, 2003a, 2003b, 2003c, 2004), bei der Beurteilung von Interaktionen zwischen zwei in Frage kommenden Ursachen (White 2002b), in Studien zur Kausalattribution nach Kelley (White 2002a) und zur Vorhersage von *cue interaction effects* (Auswirkungen von Wechselwirkungen zwischen Hinweisreizen oder von konkurrierenden Ursachen; White 2005a) erfolgreich angewendet und experimentell bestätigt. Beispielsweise macht die EE-Theorie einige Vorhersagen, an Hand derer sie von Vorhersagen der  $\Delta P$ -Regel oder der Power PC Theorie unterschieden werden kann. In eine Reihe von Experimenten, in denen  $\Delta P$  und seine Konstituenten  $P(E|C)$  und  $P(E|\neg C)$ , sowie die Basisraten des Auftretens von Ursache  $P(C)$  und Effekt  $P(E)$  konstant gehalten und unterschiedliche Darstellungs- (tabellarisch, Liste oder einzelfallweise) und Antwortformate eingesetzt wurden, konnte White (2003c) die Vorhersage stützen, dass die Kausalurteile in Abhängigkeit von  $pCI$  variierten. Umgekehrt variierten die Kausalurteile nicht, wenn  $\Delta P$  bei konstantem  $pCI$  variiert wurde. Diese Ergebnisse sprechen für die EE-Theorie und können auch nicht durch das RWM, die Power PC Theorie, eine gewichtete  $\Delta P$ -Regel oder die  $\Delta D$ -Regel adäquat vorhergesagt werden. Zudem konnte White (2003c) durch Reanalysen von Daten aus älteren Studien (Allan & Jenkins, 1983; Kao & Wasserman, 1993; Levin, Wasserman & Kao, 1993; Wasserman, Kao, Van Hamme, Katagiri & Young, 1996) zeigen, dass Kausalurteile bei konstantem  $\Delta P$  in Abhängigkeit von  $pCI$  variieren. Gemeinsam ist all diesen Studien, dass  $p(C) \neq 0,5$  ist. Auch die häufig gefundene Präferenz von Probanden für die  $\Delta D$ -Regel, sowie die gefundenen Abweichungen von der Kontingenzregel bei sehr hohen oder niedrigen Basisraten der Ursache können durch die EE-Theorie erklärt werden.

Die Theorie der Evidenzevaluation wurde bislang nur von White empirisch untersucht und bestätigt, andere Forschergruppen haben dieser Theorie noch kaum Beachtung geschenkt. Dies ist eine Schwäche dieser Theorie, da an anderen Ansätzen zu beobachten ist, dass diese durch ihre jeweiligen Vertreter meist gestützt werden und durch andere Untersucher nicht. So wird beispielsweise die Theorie der kausalen Power durch Cheng und Kollegen gestützt (Cheng, 1997; Buehner & Cheng, 1997; Buehner, Cheng & Clifford, 2003), während Vertreter assoziationalistischer Modelle (Lober & Shanks, 2000; Perales & Shanks, 2003; Vallée-Tourangeau, Murphy; Drew & Baker, 1998) oder eben White (2002c, 2003c, 2005a) als Vertreter der EE-Theorie die Theorie der kausalen Power experimentell widerlegen.

Ähnlich wie bei Prozessen der Entscheidungsfindung oder der Eindrucksbildung (Catena et al., 1998; Hogarth & Einhorn, 1992) werden Menschen in der EE-Theorie zur kausalen Urteilsbildung als naive Hypothesentester modelliert (White, 2004). In der Regel lautet die zu testende Hypothese einfach, dass eine potentielle Ursache die Wirkung verursacht. Die naive Überprüfung dieser Hypothese erfolgt dadurch, dass einzelne Beobachtungen über Ursache-Wirkungszusammenhänge in Evidenzen umgewandelt werden und anschließend die Differenz aus konfirmatorischer und disconfirmatorischer Evidenz zur Gesamtzahl der Beobachtungen ins Verhältnis gesetzt werden. Hagmayer & Waldmann (2006) kritisieren dabei die geringe Anbindung der EE-Theorie an normative Modelle. Normative Modelle legen exakt fest, wie eine Hypothesenprüfung erfolgen sollte und machen sie nicht von beobachteten Evidenzen abhängig wie die EE-Theorie. In einem engen Zusammenhang damit steht, dass die EE-Theorie derzeit nur eine rein deskriptive Theorie kausalen Denkens ist und somit keinen erklärenden Charakter hat (Hagmayer & Waldmann, 2006).

Die EE-Theorie kann eine Vielzahl empirischer Befunde erfassen und teilweise neue Befunde vorhersagen. Allerdings wird dieser Vorzug dadurch erkauft, dass die EE-Theorie eine Vielzahl freier Parameter aufweist. So können viele Befunde durch eine entsprechende individuelle Gewichtung der Evidenzen erklärt werden. Das hat zur Folge, dass die EE-Theorie nur schwer zu falsifizieren ist (Hagmayer & Waldmann, 2006).

### **3.1.5 Über den kognitiven Status regel-basierter Strategien. Oder: wie sind regel-basierte Strategien mental repräsentiert?**

Obwohl Cheng & Novick (2005) die Aufgabe, Wissen über kausale Zusammenhänge zu erwerben, als Problemlöseprozess auffassen, scheint in der Literatur zur Kausalkognition die Vorstellung weit verbreitet zu sein, dass es sich auch bei den regel-basierten Ansätzen um verschiedene Varianten eines kognitiven Mechanismus handelt (s. z.B. Allan, 1993; Hagmayer & Waldmann, 2006), der "intuitiv statistisch" abläuft und mit assoziationalistischen Mechanismen um die Vorherrschaft ringt. Die Mechanismus-Metapher mag zwar auf assoziationalistische Ansätze passen, zur Charakterisierung regel-basierter Ansätze erscheint sie uns aber eher irreführend zu sein, weil der Mechanismus-Gedanke impliziert, dass die einem Mechanismus zu Grunde liegenden Prozesse automatisiert, unflexibel und ohne bewusste Steuerung ablaufen.

Schon am Ende von Kapitel 2.1 (*Zur Ausschließlichkeit der unterschiedlichen Konzeptionen des Kausalitätsbegriffs nach Hume und Kant*, S. 8 ff.) haben wir andeutungsweise die Sichtweise vertreten, dass unser Wissen über regel-basierte Ansätze in Form von kausalen

Schemata im Gedächtnis abgespeichert ist. Weiter vermuten wir, dass Skripts die Ausführung der einzelnen Operationen steuern, die für die Berechnung von  $\Delta P$ , der kausalen Stärke  $p$  oder von  $pCI$  erforderlich sind (Lit. zu Schemata und Skripts: Anderson, 1983; Ballstaedt, Mandl, Schnotz & Tergan, 1981; Rumelhart, 1975, 1986; Schank & Abelson, 1977). Das lässt die Frage offen, wie ein solches regel-basiertes Kausalskript beschaffen sein könnte: In Anlehnung an Crocker (1981; s.a. Alloy & Tabachnik, 1984), der für den Bereich der Sozialen Wahrnehmung detailliert fünf Schritte beschrieben hat, wie Kovariationsurteile auf eine statistisch angemessene (normative) Weise gebildet werden, sollen die einzelnen Aufgabenschritte erläutert werden, welche eine Person auszuführen hat, wenn sie am Ende einer einzelfallweisen Darbietung von Informationen über Ursache-Wirkungs-Zusammenhänge regel-basiert die Stärke eines monokausalen Zusammenhangs (eine Ursache, eine Wirkung) beurteilen möchte. Die nachfolgend beschriebenen Schritte zur Aufgabenbearbeitung könnten in etwa dem Ablaufplan in einem regel-basierten Kausalskript entsprechen:

- (i) *Datensuche*: Zunächst einmal muss eine Entscheidung darüber getroffen werden, welche Informationen für die Bildung eines Kontingenzurteils relevant sind.
- (ii) *Datenauswahl*: Wenn entschieden ist, welche Informationen überhaupt relevant sind, dann müssen daraus diejenigen Informationen entnommen werden, die später als Evidenz verwendet werden soll.
- (iii) *Datenzuordnung*: Wenn klar ist, an Hand welcher Informationen eine Urteilsbildung vorgenommen werden soll, dann muss jede einzelne Information einer der vier Zellen aus der Kontingenztafel in Abbildung 1 zugeordnet werden. Bildhaft gesprochen müssen entsprechend den vier Zellhäufigkeiten mental vier Körbe bereitgestellt werden. Für jede neue Einzelfallinformation muss richtig entschieden werden, in welchen Korb sie gehört. Jede der vier Zellhäufigkeiten muss bis zur abschließenden Urteilsbildung im Gedächtnis behalten werden.
- (iv) *Datenaggregation*: Mit jedem neuen Fall wird eine der vier Zellhäufigkeiten verändert und muss neben den drei unveränderten Zellhäufigkeiten neu im Gedächtnis behalten werden. Es kommt so im Laufe der Einzelfalldarbietung zu einer schrittweisen Häufigkeitsbildung. Die Schritte (iii) und (iv) laufen lokal (d.h. bei jedem Einzelfall) seriell,

global (d.h. über alle Einzelfälle hinweg) aber parallel ab. Bildhaft gesprochen, wird für jeden neuen Einzelfall erst der Korb identifiziert, in den der neue Einzelfall gehört, dann wird er zu der im Korb vorhandenen Menge hinzugezählt. Bevor aber ein neuer Einzelfall dargeboten wird, ist der Prozess von Datenzuordnung und –aggregation für den vorherigen Einzelfall abgeschlossen.

- (v) *Urteilsbildung*: Für die Urteilsbildung müssen am Ende der Datenaggregation die vier Zellohäufigkeiten aus dem Gedächtnis abgerufen werden und entsprechend einer der Gleichungen (2), (4), (5) oder (8) miteinander verrechnet werden.

Die Schritte (i) und (ii) werden den Probanden in einem Experiment üblicherweise durch den Experimentator abgenommen, der die zu beurteilenden Fälle bereits zusammengestellt und angeordnet hat. In natürlichen Umwelten haben die Personen aber selbst die Aufgabe Daten zu sammeln und die Auswahl eines fokalen Sets (Bezugssystems) vorzunehmen (Baker et al., 1996).

Während ein *regel-basiertes Kausal-Schema* das deklarative Wissen enthält, das im Zusammenhang mit einer regel-basierten Beurteilung der Stärke eines Kausalzusammenhanges relevant ist (z.B. Wissen über Ursachen und Wirkungen, Erfolgs- und Basisraten, Zellohäufigkeiten, konfirmatorische und diskonfirmatorische Evidenz, Wissen über Anwendungskontexte, etc.) setzt sich das korrespondierende *regel-basierte Kausal-Skript* aus dem drehbuchartigen Ablaufplan oder aus der Handlungsanweisung für die einzelnen Arbeitsschritte zusammen, die für eine regel-basierte Beurteilung der Stärke eines bestimmten Kausal-Zusammenhanges erforderlich sind.

Eine andere, inhaltlich ähnliche Konzeptualisierung über den repräsentationalen Status regel-basierter Ansätze beinhaltet die Vorstellung, dass das deklarative und operative Wissen, das notwendig ist, um regel-basiert die Stärke eines bestimmten Kausal-Zusammenhanges zu bestimmen, mental als kausale Variante eines *regel-basierten Repräsentations-* oder eines *Produktionssystems* (Opwis & Lüer, 1996) abgespeichert ist. Modellvorstellungen über regel-basierte Repräsentationssysteme enthalten zusätzlich neben einem *deklarativen Datenspeicher* (der Faktenwissen enthält) und einem *Produktionen- oder Regelspeicher* (der Wissen über Regeln, Anwendungs- und Nutzungsbedingungen, Handlungen, Operationen und Prozeduren enthält) noch einen so genannten *Interpreter* als weitere Komponente. Der Interpreter wertet beispielsweise aus, ob die Bedingungen für die Anwendung von Operationen/Prozeduren für eine regel-basierte Urteilsbildung erfüllt sind (i); er wählt

einzelne Operationen/Prozeduren aus der Menge anwendbarer Operationen/Prozeduren aus (ii); und er steuert die Ausführung der einzelnen Operationen/Prozeduren (iii).

Die Frage, ob die mentale Repräsentation einer regel-basierten Strategie zutreffender als Schema plus Skript oder doch als kausale Variante eines Produktionssystems beschrieben werden kann, kann empirisch nicht entschieden werden und soll hier nicht weiter verfolgt werden. Unabhängig von der Frage, ob regel-basierte Strategien mental als Produktionssysteme oder als Schemata plus korrespondierende Skripts repräsentiert sind, ist es uns wichtig, hervorzuheben, dass regel-basiertes Kausalwissen nicht auf einer peripher-mechanistischen Ebene angesiedelt werden sollte, sondern dass es sich dabei vermutlich um Wissensrepräsentationen handelt, die aus deklarativen, operativen und prozeduralen Anteilen zusammengesetzt und dabei höheren Bewusstseins- und Überwachungsprozessen zugänglich sind.

Eine andere Frage im Zusammenhang mit der mentalen Repräsentation regel-basierter Strategien lautet, ob dieses Wissen *bereichsspezifisch* oder *bereichsübergreifend* (Shanks, 1991) ist. Wir vertreten in dieser Arbeit die These, dass es sich bei regel-basierten Strategien vermutlich erst einmal um bereichsspezifisches Kausalwissen<sup>1</sup> handelt, das von einigen Individuen während der individuellen Lerngeschichte erworben wird. Die Vermutung, dass es sich bei regel-basierten Strategien um während der Lerngeschichte erworbenes Kausalwissen handelt, wird durch den Nachweis belegt, dass die Häufigkeit des Einsatzes regel-basierter Strategien von Alter und Entwicklungsstand abhängt (Shaklee & Elek, 1988; Shaklee & Mims, 1981); es kann auch vermutet werden, dass kulturelle Unterschiede dabei eine Rolle spielen, ob und viele Individuen innerhalb eines bestimmten Kulturkreises regel-basierte Strategien einsetzen. Im Laufe der Zeit können durch Generalisierung des Gelernten auf andere Situationen aus diesem erworbenen bereichsspezifischen Kausalwissen bereichsübergreifende kausale Schemata oder Skripts entstehen, die dann auch bei der Beurteilung neuartiger Ursache-Wirkungs-Zusammenhänge herangezogen werden können.

Regel-basiertes Kausalwissen kann, sofern es im Gedächtnis repräsentiert ist, in Situationen, in denen die Stärke von Ursache-Wirkungs-Zusammenhängen bestimmt werden soll, entweder aus dem Gedächtnis abgerufen werden, oder es kann auch neu generiert bzw. erworben werden. Vermutlich unterscheiden sich Individuen darin, ob und in wie elaborierter

---

<sup>1</sup> Im Zusammenhang mit regel-basierten Strategien präferieren wir den Begriff *Wissen* gegenüber dem von Shanks verwendeten Begriff *Überzeugung* (englisch: *belief*). Shanks (1991) selbst hielt *Überzeugungen* für zutreffender, weil er die Möglichkeit aufrechterhalten wollte, dass Überzeugungen auch falsch sein können – einer Möglichkeit, der wir nicht widersprechen, da auch Wissen subjektiv sein kann. In Verbindung mit regel-basierten Strategien wie der  $\Delta P$ -Regel oder der Bestimmung der kausalen Stärke nach der Power PC Theorie wird aber gleichzeitig immer wieder deren normative Gültigkeit (z.B. Allan, 1993) betont.



Form sie über regel-basiertes Kausalwissen verfügen. Während davon ausgegangen werden kann, dass eine Reihe von Studenten und vor allem Wissenschaftler wie Patricia Cheng, Peter White oder Ward & Jenkins über dieses regel-basierte Wissen verfügen, ist dies bei den Schwesternschülerinnen in den Untersuchungen von Smedslund (1963) oder bei den Kindern in der Studie von Shaklee & Mims (1981) eher unwahrscheinlich. In Kapitel 4 bis 6 soll auf einzelne Bedingungen eingegangen werden, die einen Einsatz regel-basierter Strategien oder, anders formuliert, einen Abruf oder Neuerwerb einer regel-basierten Wissensrepräsentation wahrscheinlicher machen.

Am Ende dieses Abschnitts zur mentalen Repräsentation von regel-basierten Strategien soll noch einmal kurz auf die Unterscheidung zwischen regel-basierten Strategien und einfachen Heuristiken eingegangen werden (Kahnemann & Tversky, 1996; Kahneman, Slovic & Tversky, 1982; Tversky & Kahneman, 1973, 1974). *Heuristiken* ("Finderegeln") bestehen "aus vereinfachenden Annahmen, mit deren Hilfe ein Problem schneller gelöst werden kann als ohne Vereinfachung" (Dorsch, 2004, S.400). Sie entstehen entweder dadurch, dass eine regel-basierte Strategie fehlerhaft ausgeführt wird, oder aus eigenständigen Lösungsversuchen, auf die eine Person in Ermangelung besserer Lösungsstrategien zurückgreift. Im Bereich der Kausalkognition lassen sich Heuristiken von regel-basierten Strategien dadurch abgrenzen, dass es sich bei ihnen um ungültige, statistisch unangemessene Lösungsversuche handelt. Eine interessante Interpretation zur Entstehung von Heuristiken haben Kahnemann & Tversky (1996) vorgestellt. Normalerweise werden im Bereich der Sozialkognition Abweichungen von normativen Standards der Ignoranz des Urteilenden gegenüber normativen Regeln zugeschrieben. Nach Kahnemann & Tversky (1996) ist eine solche Interpretation aber nur gerechtfertigt, wenn die Aufgabeninstruktion den Inhalt und die Struktur des vorgegebenen Problems klar vermittelt. Weicht die Problemrepräsentation des Urteilenden von der intendierten Problemrepräsentation des Versuchsleiters ab, dann sind Abweichungen von der Urteilsnorm mehrdeutig, weil sie entweder die Anwendung unangemessener Regeln widerspiegeln oder ein Missverstehen der auszuführenden Aufgabe.

### **3.2 Assoziationalistische Ansätze**

Assoziationalistische Ansätze wurden von zahlreichen Autoren (z.B. Allan, 1993; Dickinson, 2001; Dickinson, Shanks & Evenden, 1984; Shanks, 1995, 1993; Shanks & Dickinson, 1987; Shanks, Lopez, Darby & Dickinson, 1996; Wasserman et al., 1993; Wasserman et al., 1996) zur Erklärung der Prozesse beim Erwerb von Wissen über kausale Zusammenhänge vorgeschlagen. Vertreter assoziationalistischer Ansätze gehen von der Annahme aus, dass die

Prozesse der klassischen Konditionierung bei Tieren und beim menschlichen Erwerb von Kausalwissen auf denselben grundlegenden Mechanismen beruhen.

Das Prinzip, auf dem alle assoziationalistischen Theorien beruhen, ist sehr einfach: Wenn zu einer bestimmten Gelegenheit (z.B. bei einem Lerndurchgang) zwei Stimuli oder ein Stimulusreiz (englisch: *cue*) und eine Reaktion (englisch: *outcome*) gemeinsam auftreten, dann wächst die Verknüpfung (Assoziation) zwischen ihnen. Tritt ein Reiz ohne einen anderen Reiz oder eine Reaktion in Erscheinung, dann nimmt die Verknüpfung zwischen ihnen ab. Ebenso verhält es sich, wenn eine Reaktion ohne einen dazugehörigen Stimulus eigenständig in Erscheinung tritt.

Assoziationalistische Ansätze gehen weder davon aus, dass Kausalurteile aus intuitiv durchgeführten statistischen Analysen resultieren, noch dass sie aus irgendwelchen im Gedächtnis abgespeicherten Regeln ermittelt oder berechnet werden. Stattdessen spiegelt ein Kausalurteil nicht mehr als die Stärke der relevanten Verknüpfung zwischen den mentalen Repräsentationen von einer Ursache und einem Effekt wider. Mentale Repräsentationen der Stärke von kausalen Zusammenhängen sind nicht das Produkt eines retrospektiven Denkprozesses, sondern das Ergebnis einer kontinuierlich aktualisierten Verknüpfung zwischen in Frage kommenden Ursachen und Wirkungen.

Das bekannteste assoziationalistische Lernmodell ist das Rescorla-Wagner-Modell (RWM, Rescorla & Wagner, 1972). Mit Hilfe des RWM kann nun bestimmt werden, wie sich die assoziative Stärke zwischen einem bedingten Reiz (CS = Conditioned Stimulus oder Hinweisreiz) und einem unbedingten Reiz (UCS = Unconditioned Stimulus oder Ergebnis/Reaktion/Outcome) bei wiederholter Darbietung beider Reize verändert. Die Gleichungen (10) und (11) beschreiben diesen Mechanismus:

$$\Delta V_{CS} = \alpha_{CS} * \beta_{UCS} * (\lambda - \Sigma V) \quad (10)$$

$$V_{CS(n+1)} = V_{CS(n)} + \Delta V_{CS} \quad (11)$$

In Gleichung (10) steht das Maß  $\Delta V_{CS}$  für die Veränderung der assoziativen Stärke zwischen CS und UCS nach jedem Versuch. Bei der Größe  $\alpha_{CS}$  handelt es sich um einen Lernratenparameter, der die Salienz des bedingten Reizes CS widerspiegelt, die Größe  $\beta_{UCS}$  ist ein Parameter, der die Salienz des unbedingten Reizes UCS repräsentiert. Beide Parameter variieren zwischen Werten von 0.0 und 1.0. Bleibt der bedingte Reiz CS aus, wird  $\alpha_{CS}$  auf 0 gesetzt.  $\beta_{UCS}$  kann für die Fälle, in denen der unbedingte Reiz UCS anwesend oder abwesend

ist, zwei verschiedene Werte ungleich Null annehmen. Das Maß  $\lambda$  repräsentiert die maximale assoziative Stärke die der unbedingte Reiz UCS, das Ergebnis, aufbringen kann.  $\lambda$  wird normalerweise auf einen Wert von 1.0 gesetzt, wenn der unbedingte Reiz UCS anwesend ist, und auf einen Wert von 0, wenn der unbedingte Reiz UCS bei dem gerade durchgeführten Durchgang ausbleibt.  $\Sigma V$  bezeichnet schließlich die Summe der assoziativen Stärken aller bedingten Reize CS, die bei einem Versuch oder Durchgang anwesend sind.  $\Sigma V$  wird als das "erwartete Ergebnis" bei einem vorgegebenen Durchgang interpretiert. Ist die Differenz zwischen  $\lambda$  und  $\Sigma V$  groß, dann wird auch das Maß  $\Delta V_{CS}$  groß werden. Nimmt die Differenz zwischen  $\lambda$  und  $\Sigma V$  aber ab, dann wird auch das Maß  $\Delta V_{CS}$  abnehmen und sich schließlich asymptotisch an Null annähern.

Gleichung (11) spezifiziert, wie eine neue assoziative Stärke nach  $n$  Durchgängen oder Versuchen berechnet wird. In Gleichung (11) steht das Maß  $V_{CS(n+1)}$  für die assoziative Stärke bei Versuch (Durchgang)  $n + 1$ ,  $V_{CS(n)}$  steht für die assoziative Stärke des bedingten Reizes bei Versuch (Durchgang)  $n$  und  $\Delta V_{CS}$  bezeichnet die Veränderung, die eintritt, nachdem der bedingte und der unbedingte Reiz bei Versuch (Durchgang)  $n$  gemeinsam in Erscheinung getreten sind.

Um das Rescorla-Wagner Modell auf den Erwerb von Wissen über Ursache-Wirkungs-Zusammenhänge bei Menschen anwenden zu können, wird der bedingte Reiz CS durch die zu beurteilende Ursache C und der unbedingte Reiz UCS durch den Effekt E ersetzt. Dementsprechend kann Gleichung (10) folgendermaßen umgewandelt werden:

$$\Delta V_C = \alpha_C * \beta_E * (\lambda - \Sigma V) \quad (12)$$

In Gleichung (12) ist das Maß  $V_C$  ein Prädiktor für den geschätzten Kausalzusammenhang.

Neben dem bekannten Rescorla-Wagner Modell existieren noch weitere assoziative Ansätze, die auf den Erwerb von menschlichem Kausalwissen angewendet worden sind (s. Dickinson et al., 1984; Vallée-Tourangeau et al., 1998), wie z.B. das Pearce-Hall Modell (Pearce & Hall, 1980) oder das Konfigural-Modell von Pearce (1987, 1994).

### *Empirische Befunde und Kritik*

Assoziationalistische Ansätze zur Erklärung des Erwerbs von Wissen über kausale Zusammenhänge beim Menschen hatten ihre Blütezeit in den 80er Jahren des vergangenen Jahrhunderts. Ein Vorzug assoziationalistischer Ansätze ist, dass ihre grundlegenden Prinzipien relativ einfach sind und vom Organismus nur wenig Gedächtniskapazität verlangen, weil Erfahrung lediglich als eine geringe Anzahl assoziativer Stärken gespeichert wird.

Weil assoziationalistische Theorien an sukzessive Veränderungen der assoziativen Stärken gebunden sind, unterliegen sie der Einschränkung, dass sie einer Datenmodellierung nur dann zu Grunde gelegt werden können, wenn die Informationen über die zu lernenden Ursache-Wirkungs-Zusammenhänge einzelfallweise präsentiert werden. Sie sind nicht anwendbar, wenn die Informationen über Ursache-Wirkungs-Zusammenhänge in irgendeiner summarischen Form präsentiert werden (Alloy & Tabachnik, 1984; Shanks, 1991).

Assoziationalistische Theorien können eine Reihe von Befunden aus dem Bereich des kausalen Lernens erklären. (Allan, 1993, 2003; Shanks, 1993, 2007; Shanks et al., 1996; s.a. Hagmayer & Waldmann, 2006). So sind menschliche Kausalurteile in Analogie zu experimentellen Befunden mit Tieren sensitiv für die Kontingenz zwischen Ursache und Wirkung. Wasserman et al. (1990) konnten zeigen, dass nicht nur ein regel-basierter Ansatz diese Sensitivität vorhersagen kann, sondern auch das RWM. Der Density-Bias (Alloy & Abramson, 1979; Dickinson et al., 1984; Shanks, 1987) kann erklärt werden, wenn der Lernprozess vor dem Erreichen der Lernasymptote heran gezogen wird. Auch den Befund, dass die Einschätzung der Stärke eines Zusammenhangs mit einem zunehmenden zeitlichen Abstand zwischen Ursache und Wirkung abfällt (Einfluss der temporalen Kontiguität) kann von assoziationalistischen Ansätzen modelliert werden (Shanks, Pearson & Dickinson, 1989; Wasserman & Neunaber, 1986). Weiter spricht das Vorliegen von Lernkurven dafür, dass kausales Lernen assoziationalistischen Prinzipien folgt. Typischerweise sehen Kurven für kausales Lernen so aus, dass die Einschätzungen der Stärke positiver Kontingenzen inkrementell ansteigen und die von negativen Kontingenzen abfallen (Baker, Mercier, Vallée-Tourangeau, Frank & Pan, 1993; López, Almaraz, Fernández & Shanks, 1999; Shanks, 1987). Das RWM lässt sich besonders eindrucksvoll in Untersuchungen zum kompetitiven Lernen bestätigen und kann zur Erklärung von *cue interaction effects* (Auswirkungen von Wechselwirkungen zwischen Hinweisreizen oder von konkurrierenden Ursachen; s. Allan, 1993; De Houwer & Beckers, 2002; Shanks, 1995) herangezogen werden. Beim kompetitiven Lernen stehen mehrere Hinweisreize im Wettbewerb zueinander. Ein Beispiel für kompetitives Lernen und *cue interaction effects* ist der aus Tierversuchen bekannte *Blockierungseffekt* (Kamin, 1968, 1969). Dieser konnte auch in Untersuchungen zum menschlichen Kausallernen (Beckers, De Houwer, Pineño & Miller, 2005; Chapman, 1991; Chapman & Robbins, 1990; Dickinson et al., 1984; De Houwer & Beckers, 2002; Shanks, 1985) repliziert werden, in denen die Probanden in einer ersten Lernphase einen perfekten Zusammenhang zwischen einem konditionierten Stimulus (CS1), der Ursache (z.B. ein Licht), und einem nachfolgenden unkonditionierten Stimulus (US), der Wirkung (z.B. Speichelfluss)

lernten. In Phase 2 trat CS1 (Licht) dann immer gemeinsam mit einem zweiten CS (CS2), also einer zweiten Ursache (z.B. ein Ton), auf, gefolgt von US (Speichelfluß). Typischerweise lernen Menschen wie Tiere in solchen Blockierungsexperimenten nur einen Zusammenhang zwischen CS1 (Licht) und US (Speichelfluß), der Erwerb von Wissen über einen Zusammenhang zwischen CS2 (Ton) und US (Speichelfluß) wird hingegen im Vergleich zu einer Kontrollgruppe, in der in Lernphase 1 kein Zusammenhang zwischen CS und US gelernt wird, blockiert.

Der so genannte *retrospektive Reevaluationeffekt* (englisch: *retrospective revaluation* – dt. *retrospektive Um- oder Neubewertung bzw. -abschätzung*; Chapman, 1991), ein weiterer *cue interaction effect*, kann nur durch eine modifizierte Form des RWM (Van Hamme & Wasserman, 1994) oder verwandte Modelle, wie das *Standard Operating Procedures Modell* (Aitken, Larkin & Dickinson, 2001, 2000; Dickinson & Burke, 1996; Larkin, Aitken & Dickinson, 1998) erklärt werden, nicht aber durch das RWM selbst. In Experimenten zur retrospektiven Reevaluation (vgl. Chapman, 1991; De Houwer, Beckers & Glautier, 2002; Dickinson & Burke, 1996; Larkin, Aitken & Dickinson, 1998; Melchers, Lachnit, & Shanks, 2004; Wasserman & Berglan, 1998) wird, anders als bei Nachweisen des Blockierungseffekts, die Reihenfolge der beiden Lernphasen umgedreht: In Lernphase 1 treten also die beiden potentiellen Ursachen Licht (CS1) und Ton (CS2), gefolgt von der Wirkung Speichelfluss (US), gemeinsam auf, in Lernphase 2 geht nur eine der beiden Ursachen (Licht zum Beispiel) dem Speichelfluss voraus. Anders als beim Blockierungseffekt, revidieren die Versuchsteilnehmer typischerweise ihre Einschätzung der beiden Ursachen in Phase 2. Während in Phase 2 das Licht als kausal bedeutsam für das Auftreten von Speichelfluss gesehen wird, schätzen die Versuchsteilnehmer den nur in Phase 1 präsentierten Ton als weniger kausal bedeutsam ein. Es findet also eine retrospektive Blockierung statt, auch wenn der Effekt meist kleiner ausfällt als im Standardparadigma.

Mit dem RWM und seinen Weiterentwicklungen lassen sich auch eine Reihe anderer Befunde und *cue interaction effects* erklären, wie z.B. präventives Lernen und Super-Learning (Aitken, Larkin & Dickinson, 2000), Überschattung (Baker, Mercier, Vallée-Tourangeau, Frank & Pan, 1993; Baker, Vallée-Tourangeau, & Murphy 2000; Price & Yates 1993), Auswahleffekte (Shanks, 1993; Shanks et al., 1996), konditionierte Inhibition, Akquisition und Übererwartungseffekt (Allan, 1993).

Ein wesentlicher Kritikpunkt an assoziationalistischen Theorien beruht darauf, dass der Einfluss früherer Erfahrungen und von Vorwissen aus dem episodischen Gedächtnis bei der Berechnung der assoziativen Stärke ignoriert wird (Baker, Murphy & Vallée-Tourangeau,

1996). Zudem existieren mittlerweile eine Reihe von Alternativmodellen, wie z.B. die Power PC Theorie (Cheng, 1997; Novick & Cheng, 2004), die Theorie der Evidenzevaluation (White, 2002a, 2002c, 2003a, 2003b, 2003c, 2004, 2005a), Bayes-Netzwerke (Gopnik, Glymour, Sobel, Schulz, Kushnir & Danks, 2004; Gopnik, Sobel, Schulz & Glymour, 2001; Griffiths & Tenenbaum, 2005) und deduktive Ansätze (De Houwer, Beckers & Glautier, 2002; Lovibond, 2003; Lovibond, Been, Mitchell, Bouton & Frohardt, 2003) die ebenfalls in Lage sind, *cue interaction effects* vorherzusagen. Deduktive Ansätze gehen beispielsweise davon aus, dass sich *cue interaction effects* besser durch Prozesse des schlussfolgernden Denkens als durch einfache assoziative Prinzipien erklären lassen (De Houwer & Beckers, 2003; De Houwer, Beckers & Vandorpe, 2005; Lovibond, 2003; Lovibond, Been, Mitchell, Bouton & Frohardt, 2003). Im Zusammenhang mit Blockierung konnte zudem gezeigt werden, dass der Blockierungseffekt davon abhängt, ob die Probanden annehmen, dass der redundante CS in Lernphase 1 abwesend ist oder lediglich nicht beobachtet wurde (De Houwer, 2002; Waldmann, 2000). Eine Blockierung trat in diesen Studien nur dann auf, wenn die Probanden davon ausgingen, dass der redundante CS in Lernphase 1 wirklich abwesend war. Kritisch für assoziationalistische Ansätze sind auch Studien, in denen gezeigt werden konnte, dass Probanden komplexe Reevaluationen vornehmen können (Macho & Burkart, 2002). Assoziationalistischen Ansätzen zufolge lassen sich Blockierung und retrospektive Reevaluation nicht durch die Heranziehung von Ereignissen aus der Vergangenheit oder durch Prozesse des schlussfolgernden Denkens erklären. Stattdessen wird angenommen, dass Lernen in aktuellen und zukünftigen Durchgängen aufgrund der Erwartung von aktuell abwesenden Hinweisreizen modifiziert wird. Entsprechend haben assoziationalistische Modelle Probleme mit Befunden, die echte Retrospektion belegen (Hagmayer & Waldmann, 2006).

In der Theorie der Kausalmodelle haben Waldmann und Mitarbeiter (Waldmann, 1994, 1996; Reips, 1997; Hagmayer, 2001; Hagmayer & Waldmann, 2006) die gegen assoziationalistische Ansätze gerichtete These vertreten, dass Menschen Wissen über Kausalzusammenhänge nicht nur datengesteuert erwerben, sondern dass dieser Wissenserwerb durch kausales Vorwissen, durch kausale Schemata, Strukturen oder Modelle, beeinflusst wird. Darüber hinaus lassen assoziationalistische Ansätze grundlegende Eigenschaften von Kausalität, wie zum Beispiel das *Prinzip der kausalen Gerichtetheit* (s. Mackie, 1974) außer Acht. Dieses Prinzip besagt, dass Ursachen Effekte beeinflussen und nicht umgekehrt. Nach Waldmann und Mitarbeitern (Waldmann & Hagmayer, 2001; Reips, 1997; Waldmann, 2000; Waldmann & Holyoak, 1992, 1997) sind Menschen dazu in der Lage, zwischen einer prädiktiven (Bakterien verursachen eine Krankheit) und einer diagnostischen (Schließen von Symptomen auf eine Krankheit und

ihre Ursache) Lernrichtung beim Erwerb von Wissen über kausale Zusammenhänge zu unterscheiden. Assoziationalistische Theorien wie das RWM machen keine unterschiedlichen Annahmen in Abhängigkeit von der kausalen Direktionalität.

Von Cheng (1997) wurde kritisiert, dass assoziationalistische Theorien die Beziehung von Kovariation und kausaler Power falsch konzeptualisieren, indem sie beides irrtümlicherweise gleich setzen. So kann beispielsweise die heilende Wirkung eines Medikaments nicht eingeschätzt werden, wenn alle Patienten gesund sind (Bodeneffekt). Analog kann der schädigende Einfluss von Rauchen auf die Atmung nicht erfasst werden, wenn ohnehin schon alle Patienten aus anderen Gründen an Atembeschwerden leiden (Deckeneffekt). Für Cheng (1997) kann in beiden Fällen von Nullkontingenz nicht, wie Vertreter assoziationalistischer Theorien es tun müssten, darauf geschlossen werden, dass kein kausaler Zusammenhang vorliegt, weil nach ihrer Auffassung immer die Basisrate des Effekts mitberücksichtigt werden muß. Neben dem Einfluss kausaler Modelle haben assoziationalistische Theorien noch mit weiteren Problemen zu kämpfen, für die aber Lösungen aus assoziationalistischer Sicht formuliert werden können (Shanks et al., 1996). Allan (2003) wendet gegen Cheng (1997) ein, dass diese wesentliche Stärken und Weiterentwicklungen des RWM, wie das modifizierte RWM (Van Hamme & Wassermann, 1994) ignoriert und zu Unrecht kritisiert habe.

### **3.3 Weitere Ansätze: Die Rolle von Vorwissen**

Es existieren noch eine Reihe von weiteren Ansätzen, die sich mit der Fragestellung auseinandergesetzt haben, wie Menschen Wissen über kausale Zusammenhänge erwerben (s. Hagemayer & Waldmann, 2006, für einen Überblick). Von besonderer Bedeutung sind dabei der mechanistische Ansatz und die Theorie der Kausalmodelle. Beiden Modellen ist gemeinsam, dass sie die Rolle des Vorwissens beim kausalen Wissenserwerb hervorheben. Da diese Theorien im empirischen Teil dieser Arbeit nicht untersucht wurden, werden sie hier nur knapp dargestellt.

#### **3.3.1 Der mechanistische Ansatz**

Vertreter einer mechanistischen Sichtweise (Ahn et al., 1995; Koslowski, 1996, Schlottmann, 1999) gehen davon aus, dass das Wissen über Mechanismen die Grundlage für den Erwerb von Wissen über kausale Zusammenhänge darstellt. Im Gegensatz zu assoziationalistischen und regel-basierten Ansätzen wird also die Wichtigkeit inhaltlichen Vorwissens in einem bestimmten Bereich betont. Seine empirischen Wurzeln hat der mechanistische Ansatz in Arbeiten zum Denken und Problemlösen; in der Forschung zur Repräsentation von

Wissensstrukturen in Form von Skripts und Schemata, sowie in der kognitiven Entwicklungspsychologie (Hagmayer & Waldmann, 2006). Erkenntnisse aus der kognitiven Entwicklungspsychologie sind deshalb interessant, weil sie zum einen zeigen, dass Wissen über kausale Mechanismen mit dem Alter zunimmt (Bullock et al., 1982). Zum anderen zeigen sie aber auch, dass bereits drei- bis vierjährige Kinder in einfachen Aufgaben kausale Schlussfolgerungen nach den gleichen Prinzipien wie Erwachsene ableiten (Bullock et al., 1982). Diese Prinzipien beinhalten beispielsweise die Annahmen, dass jedes Ereignis eine Ursache hat, dass es sich bei Ursachen um Ereignisse handelt, die dem Effekt zeitlich vorausgehen, und dass ein Effekt durch einen kausalen Mechanismus, eine kausale Kraft zustande kommt. Weitere Studien konnten nachweisen, dass nicht nur Kinder, sondern auch Erwachsene mit Collegeerfahrung beim Lernen und Denken auf intuitive, häufig auch unvollständige und rudimentäre Theorien in verschiedenen Domänen zurückgreifen (Rosenblit & Keil, 2002; Wilson & Keil, 1998).

Koslowski (1996) referiert eine Vielzahl von Studien, in denen Einfluss des Vorwissens auf die Beurteilung der Stärke von einfachen Ursache-Wirkungs-Zusammenhängen untersucht wurde. Dabei stellte sich heraus, dass bei gleicher Kontingenz Kausalzusammenhänge, in denen eine Ursache mit einem plausiblen Wirkmechanismus vorgegeben war, deutlich stärker eingeschätzt wurden als Kausalzusammenhänge, in denen ein plausibler Wirkmechanismus fehlte. In einer Studie von Alloy & Abramson (1979) konnte der Einfluss von Vorannahmen über die eigene Wirksamkeit von Handlungen demonstriert werden. Dort waren viele Probanden der Ansicht, dass das Drücken eines Schalters das Aufleuchten einer Lampe beeinflusste, obwohl zwischen beiden Ereignissen keinerlei Zusammenhang existierte.

Der mechanistische Ansatz wird weiter durch eine Reihe von Experimenten (Ahn & Baillenson, 1996; Ahn, Kalish, Medin & Gelman, 1995; s.a. Ahn & Kalish, 2000) gestützt, denen ein *Informationssuche-Paradigma* zu Grunde gelegt wurde. Dabei werden die Probanden mit bestimmten Situationen konfrontiert und sollen erklären, warum ein bestimmtes Ereignis (z.B. ein Flugzeugabsturz) eingetreten ist. Weiter werden sie danach gefragt, welche Art von Informationen sie benötigen, um die Ursache eines Ereignisses zu ermitteln. Beispielsweise wollten Ahn et al. (1995) herausfinden, ob Probanden eher Fragen zur Kovariation zwischen verschiedenen Ereignissen stellten oder ob sie eher Hypothesen über bestimmte Mechanismen überprüften. Das Ergebnis der Studie war, dass sich die Fragen mehrheitlich auf Mechanismen (ca. 65 %) und nur zu einem kleinen Teil auf Kovariationen bezogen (zwischen 7 und 13 %). Ahn et al. (1995) sehen darin einen Beweis dafür, dass Wissen um Mechanismen wichtiger ist als Wissen über Kovariation. Für Cheng (1997, 1993)



ist diese Schlussfolgerung aber nicht gerechtfertigt. Sie argumentiert (Cheng, 1993), dass Mechanismustheorien erklären, wie einmal erworbenes Wissen über Mechanismen auf neues Lernen übertragen wird, dass sie aber keine plausible Antwort auf die Frage geben, woher das Mechanismuswissen kommt. Cheng (1993) geht davon aus, dass Kovariationen die primäre Quelle für den Erwerb von Neuwissen sind, und dass die Nutzung dieses Wissens auf weiteres Lernen eher nachgeordnet ist.

### 3.3.2 Theorie der Kausalmodelle und Bayes-Netze

Die Theorie der Kausalmodelle ist eine von Waldmann & Holyoak (1992) begründete Theorie des kausalen Wissenserwerbs. Unter Kausalmodellen sind Wissensrepräsentationen von bzw. Annahmen über Ursache-Wirkungszusammenhänge in der Welt zu verstehen (Hagmayer, 2001). Solches Kausalwissen beinhaltet beispielsweise das Prinzip temporaler Priorität, (Ursachen sind zeitlich vor ihren Effekten vorhanden), und das Prinzip der kausalen Gerichtetheit (Ursachen beeinflussen ihre Effekte und nicht umgekehrt). Im wesentlichen haben Kausalmodelle vier Funktionen (Hagmayer, 2001): Sie helfen uns, Phänomene und Sachverhalte zu verstehen und nachvollziehbar zu machen (i), sie erklären Sachverhalte (ii), ermöglichen Vorhersagen über zukünftige Ereignisse (iii) und leiten Handlungen (iv). Obwohl eine Vielzahl von Kausalmodellen existiert, lassen sich diese auf drei einfache Grundmodelle (Gemeinsame-Ursache Modell, Gemeinsame-Effekt Modell, Kettenmodell) zurückführen. Komplexe Kausalmodelle lassen sich durch die Verbindung von Grundmodellen rekonstruieren (Hagmayer, 2001).

In der Theorie der Kausalmodelle wird von der Annahme ausgegangen, dass bereits vorhandenes abstraktes und bereichsübergreifendes Wissen über kausale Eigenschaften des Lernmaterials den Erwerb von Wissen über kausale Zusammenhänge steuert. Dies bedeutet, dass auch wenn kein bereichsspezifisches Wissen vorhanden ist, allgemeines Vorwissen beim Erwerb von Kausalwissen genutzt wird. Weiter beinhaltet die Theorie der Kausalmodelle, dass der Erwerb von Wissen über kausale Zusammenhänge nicht nur auf der Verarbeitung von statistischer Information basiert. Anders als in assoziationalistischen Theorien interagieren beim Erwerb von Wissen über kausale Zusammenhänge datengesteuerte (bottom-up) und wissensgesteuerte (top-down) Prozesse miteinander. Die Grundannahme dabei ist, dass die Struktur der Kausalmodelle einerseits die Interpretation des Lerninputs steuert, andererseits aber wiederum selbst durch den Lerninput modifiziert werden kann.

Die Theorie der Kausalmodelle versucht, komplexe Kausalzusammenhänge erfassbar zu machen, in denen eine große Zahl von Ursachen und Effekten miteinander interagieren.

Dadurch erhebt sie den Anspruch, der Wirklichkeit eher zu entsprechen als die anderen bisher vorgestellten Ansätze, welche sich mit der Situation befassen haben, dass eine oder mehrere Ursachen nur einen gemeinsamen Effekt haben.

Die Theorie der Kausalmodelle bildet keine Annahmen über die Nutzung von bereichsspezifischem Wissen (Wissen über spezifische Mechanismen), doch wird davon ausgegangen, dass solches Wissen - falls es verfügbar ist - bei der Ermittlung der Stärke von kausalen Zusammenhängen berücksichtigt wird (Waldmann & Hagmayer, 2001). Es liegen mittlerweile zahlreiche Studien vor, die belegen, dass Kausalmodelle den Erwerb von Wissen über Ursache-Wirkungs-Zusammenhänge beeinflussen (für einen Überblick s. Hagmayer & Waldmann, 2006). Allerdings existieren auch kritische Befunde, die keine Sensitivität für den Aspekt der Kausalrichtung und die Struktur der Kausalmodelle beim Lernen finden konnten (Cobos et al., 2002; Matute, Arcediano & Miller, 1996; Price & Yates, 1993; Shanks & Lopez, 1996). Hagmayer & Waldmann (2006) nehmen an, dass die Komplexität der Aufgabe, die Plausibilität der Instruktionen und die Art der Testung entscheidend dafür sein kann, ob Versuchsteilnehmer in der Lage sind, Wissen über Kausalmodelle adäquat zu erwerben.

Um Kausalmodelle formal darstellen zu können, lassen sich lineare Strukturgleichungen oder bayesianische Netzwerkmodelle einsetzen. Beide Verfahren schließen sich nicht gegenseitig aus, sondern können ineinander überführt werden (Pearl, 2000). Bayesianische Netzwerkmodelle (Gopnik, Glymour, Sobel, Schulz, Kushnir & Danks, 2004; Gopnik, Sobel, Schulz & Glymour, 2001) sind eine Weiterentwicklung aus probabilistischen und mechanistischen Ansätzen (Hagmayer & Waldmann, 2006; Pearl, 2000). Von Seiten der Forschung zum Erwerb von Kausalwissen haben Bayes-Netzwerke in jüngerer Zeit ein reges Interesse erfahren. So haben sich mehrere Studien damit befasst, das Ausmaß zu untersuchen, in dem menschliches Kausaldenken den qualitativen Annahmen, die den Bayes-Netzen zu Grunde liegen, entsprechen (Gopnik et al., 2004; Glymour, 1998, 2001; Lagnado & Sloman, 2004; Waldmann & Martignon, 1998). Darüber hinaus wurden Bayes-Netze auch dazu eingesetzt, um quantitative Vorhersagen über menschliches Verhalten bei der Beurteilung kausaler Zusammenhänge machen zu können (Griffiths & Tenenbaum, 2005; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003; Tenenbaum & Griffiths, 2001, 2003). Nach Glymour (2001) lässt sich Chengs (1997) Power PC Theorie als ein Spezialfall eines Bayes-Netzwerks formalisieren.

### 3.4 Kritik an monomechanistischen Theorieansätzen: Schließen assoziative und regel-basierte Mechanismen einander aus?

Die bisherigen Ergebnisse zeigen, dass die Frage, ob das Erlernen oder die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen auf der Basis von statistischen Informationen assoziativ oder regel-basiert erfolgt, immer noch kontrovers diskutiert wird (Shanks, 2007). Keines der vorgestellten Modelle ist in der Lage, alle empirischen Befunde zu erklären (Allan, 1993, Baker, Murphy, Vallée-Tourangeau, 1996; Cheng, 1997; De Houwers & Becker, 2002; De Houwers, Beckers & Vandorpe, 2005; Dickinson, 2001; Shanks, 2004, 2007; Shanks, Holyoak & Medin, 1996). Während einige Forschungsarbeiten zu dem Ergebnis kommen, dass Menschen regel-basiert oder normativ vorgehen (z.B. Buehner, Clifford & Cheng, 2003; Cheng, 1997; Cheng & Novick, 1990, 1992; Cheng & Holyoak, 1995; Melz, Cheng, Holyoak & Waldmann, 1993), gelangen andere Arbeiten zu dem Schluss, dass menschliche Kausalurteile auf assoziativen Mechanismen beruhen (z.B. Allan, 1993; Dickinson, 2001; Shanks, 1993; Van Hamme et al., 1993; Van Hamme & Wasserman, 1993). Die widersprüchlichen Ergebnisse haben zu der Kritik Anlass gegeben, dass nicht davon ausgegangen werden kann, dass der komplexe Prozess der kausalen Urteilsbildung nur durch einen einzigen Mechanismus allumfassend erklärt werden kann (s. Baker, Murphy, Vallée-Tourangeau, 1996; De Houwer & Beckers, 2002; De Houwer, Beckers & Vandorpe, 2005; Dickinson, 2001; Fum & Stocco, 2003; Hagmayer & Waldmann, 2006; Lagnado & Shanks, 2002; Catena & Perales, 2002; Perales, Catena & Maldonado, 2004; Perales, Catena, Shanks & González, 2005; Price & Yates, 1995; Shanks, 2007; Tangen & Allan, 2004; Young, 1995). Catena & Perales (2002), heben beispielsweise hervor, dass die Vertreter assoziationalistischer und regel-basierter Ansätze bei der Suche nach den Mechanismen, die der Bildung von Kausalurteilen zugrunde liegen, den Einfluss von Entscheidungsbildungs- oder rationalen Denkprozessen zwar nicht völlig ignoriert, aber doch vernachlässigt haben. De Houwer & Beckers (2002) sprechen sich deshalb für ein *hybrides Modell* aus, in dem Elemente von assoziationalistischen und nicht assoziationalistischen Modellen vereinigt werden und entsprechend davon ausgegangen wird, dass das Erlernen von Ursache-Wirkungs-Zusammenhängen sowohl durch rationales Denken, als auch durch einfache assoziative Mechanismen gesteuert wird. Nach De Houwer & Beckers (2002) greifen Menschen dabei aber nur dann auf rationale Denkprozesse und Schlussfolgerungen zurück voraus, wenn sie dazu motiviert sind und die Gelegenheit dazu haben. Bei fehlender Motivation oder Gelegenheit zum Einsatz rationaler Denkprozesse kommen dagegen eher assoziative Mechanismen zum Einsatz. Catena & Perales (2002) verweisen zum Beleg der These, dass

regel-basierte und assoziative Mechanismen miteinander kompatibel seien, auf andere Bereiche wie beispielsweise den menschlichen Spracherwerb (Markus et al., 1999; Pinker & Prince, 1988). Im Bereich des Erwerbs von Wissen über kausale Zusammenhänge sprechen nach Catena & Perales (2002) Studien, die belegen, dass unterschiedliche Antwort- bzw. Urteilsmodalitäten unterschiedliche Lernprozesse in Gang setzen (Catena, Maldonado, & Cándido, 1998; Collins & Shanks, 2002; Matute, Vegas & De Marez, 2002), für eine Vereinbarkeit von regel-basierten und assoziativen Mechanismen.

Auch Baker, Murphy & Vallée-Tourangeau (1996) erteilen allumfassenden Erklärungsversuchen (so genannten "Omnibus-Theorien") eine klare Absage und haben zur Überwindung der Kontroverse assoziativ oder regel-basiert einen *evolutionären Standpunkt* (s.a. Catena & Perales, 2002) eingenommen. In Anlehnung an Fodors (1983) modulare Theorie gehen Baker, Murphy & Vallée-Tourangeau (1996) davon aus, dass viele verschiedene, wahrscheinlich unabhängige Systeme mit unterschiedlichen Mechanismen bei Menschen und Tieren an der Verarbeitung von Informationen zu kovariierenden oder kausal verknüpften Ereignissen beteiligt sind. Sie halten es dabei weiter für wahrscheinlich, dass Art und Weise der Informationsverarbeitung in diesen separaten Systemen in Abhängigkeit von den "ökonomischen Beschränkungen" (Baker et al., 1996, S. 6) des Systems variieren.

Schließlich wurde zur Kontroverse zwischen assoziationalistischen und regel-basierten Ansätzen von verschiedenen Autoren eingewendet (s. z.B. Cheng & Holyoak, 1995; Lober & Shanks, 2000; Perales & Shanks, 2003), dass der Vergleich von assoziationalistischen Ansätzen mit regel-basierten Theorien problematisch ist, weil beide auf unterschiedlichen Ebenen der Analyse kognitiver Prozesse anzusiedeln sind. Legt man Marrs (1982) Unterscheidung zwischen einer *computationalen* und einer *algorithmischen* Ebene der Analyse kognitiver Prozesse zu Grunde, dann gelten assoziationalistische Ansätze wie das RWM als algorithmische Theorien, die die Mechanismen beschreiben und erklären sollen, die Menschen einsetzen, um Informationen über Ursache-Wirkungs-Zusammenhänge zu verarbeiten. Das PCM und die PPC-Theorie sind hingegen Beispiele für computationale Analysen kognitiver Prozesse, in denen spezifiziert wird, was während des Verlaufs des Erwerbs von Kausalwissen berechnet wird (Lober & Shanks, 2000); oder allgemeiner gesprochen, in denen die auszuführende Aufgabe, das dafür benötigte Wissen und die dazu verwendeten Strategien spezifiziert werden (Strube & Schlieder, 1996). Nach Perales & Shanks (2003) soll eine computationale Theorie zum Erwerb von Kausalwissen festlegen, mit welchen angemessenen Normen oder Standards die Überzeugungen und Verhaltensweisen

beim Erwerb von Kausalwissen verglichen werden können, um herauszufinden, ob Personen bei der Verarbeitung kausaler Informationen rational vorgehen oder nicht.

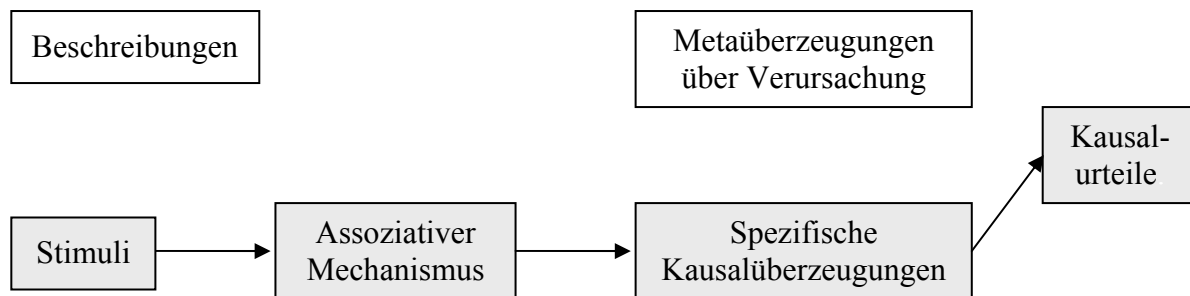
### **3.5 Ein möglicher Rahmen für einen integrativen Ansatz: Die Theorie von Shanks (1991)**

Weil es trotz zahlreicher Untersuchungen unklar geblieben ist, ob die Prozesse bei der Beurteilung der Stärke von kausalen Zusammenhängen besser durch regel-basierte oder durch assoziationalistische Ansätze erklärbar sind, wurden integrative Ansätze von einer Reihe von Autoren immer wieder propagiert (Baker et al. 1996; Price & Yates, 1995; Fum & Stocco, 2003) und vereinzelt auch entwickelt. Einen möglichen Rahmen für einen integrativen Ansatz liefert beispielsweise die Theorie von Shanks (1991), der ebenfalls die Auffassung vertritt, dass regel-basierte und assoziationalistische Ansätze für sich genommen keinen Anspruch auf Allgemeingültigkeit erheben können. So lassen sich beispielsweise textbasierte Beschreibungen kausaler Zusammenhänge nicht mit assoziationalistischen Modellen erklären. Regel-basierte Ansätze haben zwar weniger Probleme damit, kausale Zusammenhänge in erfahrenen und beschriebenen Situationen zu beschreiben, doch liegen, wie wir gesehen haben, eine Reihe von Befunden vor, die sie zu widerlegen scheinen.

Shanks (1991) spricht sich deshalb dafür aus, dass Menschen sowohl über einen assoziativen Mechanismus, als auch über bereichsspezifische Kausalüberzeugungen (englisch: *domain specific causal beliefs*) und bereichsübergreifende Metaüberzeugungen über Verursachung (englisch: *metabeliefs about causation*) verfügen. Der Bildung von Kausalurteilen können also entweder assoziative Mechanismen oder Kausalüberzeugungen zugrunde liegen. Ob ein assoziativer Mechanismus oder Kausalüberzeugungen aktiviert werden, hängt dabei vom jeweiligen Situationstypus ab. Shanks (1991) differenziert hier zwischen zwei Situationstypen, nämlich erlebten und beschriebenen Situationen. Erlebte Situationen sind unter anderem dadurch gekennzeichnet, dass die relevanten Kausalereignisse zeitlich verteilt sind. In beschriebenen Situationen werden die Kausalinformationen entweder in summarischer Form dargeboten, oder es werden sprachlich-textliche Informationen präsentiert, aus denen kausale Schlussfolgerungen gezogen werden sollen.

Weil assoziative Mechanismen a priori die einzelfallweise Anpassung oder Adjustierung von Assoziationsstärken erfordern, können sie nur in erlebten Situationen wirksam werden. Zudem werden assoziative Mechanismen vor allem dann in Gang gesetzt, wenn Individuen bei der Bildung von Kausalurteilen mit unvertrauten Ereignissen konfrontiert sind, in denen

sie wenig oder gar kein Vorwissen anwenden können. Abbildung 2 veranschaulicht, wie es zur Bildung von Kausalurteilen in erlebten Situationen kommt:

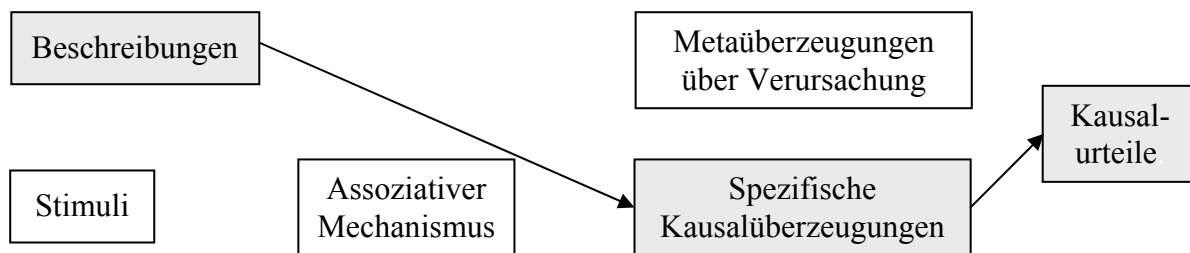


**Abbildung 2:** Bildung von Kausalurteilen in erlebten Situationen (nach Shanks, 1991).

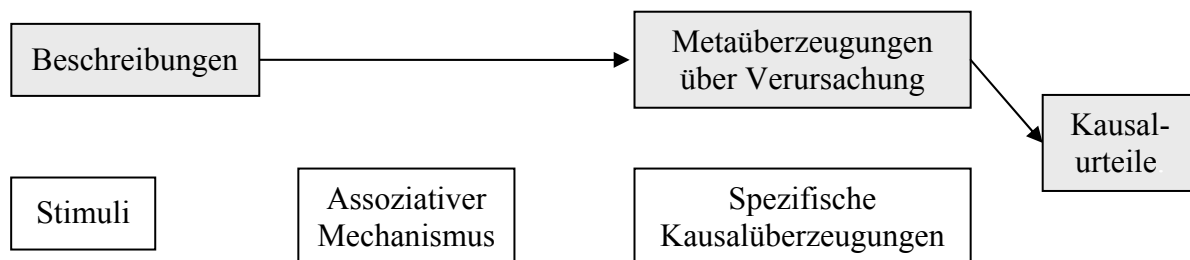
In erlebten Situationen werden die einzelfallweisen Informationen über Ursache-Wirkungs-Zusammenhänge von einem assoziativen Mechanismus in neue spezifische Kausalüberzeugungen umgewandelt, welche wiederum als Grundlage für die Bildung von Kausalurteilen herangezogen werden. Bemerkenswert an Abbildung 2 ist aber, dass Shanks (1991) zu Folge ein assoziativer Mechanismus allein nicht ausreicht, um die Bildung von Kausalurteilen in erlebten Situationen zu erklären.

In beschriebenen Situationen sind Beschreibungen das Ausgangsmaterial bei der Bildung von Kausalurteilen. Shanks (1991) unterscheidet dabei zwei Typen von beschriebenen kausalen Situationen: Bei beschriebenen kausalen Situationen vom Typ I werden summarische Informationen in statistischer Form, in der Regel als Häufigkeiten, dargeboten, in beschriebenen kausalen Situationen vom Typ II werden die kausalen Schlüsse aus linguistischen Informationen gezogen, es handelt sich also um textbasierte beschriebene Situationen. Gemeinsam ist beiden Arten von beschriebenen Situationen, dass kein assoziativer Mechanismus benötigt wird, damit die kausalen Beschreibungen Kausalüberzeugungen aktivieren, welche die Grundlage der Bildung von Kausalurteilen sind. In beschriebenen kausalen Situationen kann sich nun zweierlei ereignen: Entweder führt eine Beschreibung zu einem unmittelbaren Abruf von bereichsspezifischen Kausalüberzeugungen aus dem Gedächtnis (Möglichkeit A), oder es werden für den Fall, dass durch die Beschreibungen keine bereichsspezifischen Überzeugungen aus dem Gedächtnis abgerufen werden können, direkt bereichsübergreifende Metaüberzeugungen aktiviert (Möglichkeit B). Die beiden Möglichkeiten A und B der Bildung von Kausalurteilen in beschriebenen Situationen über Ursache-Wirkungs-Zusammenhänge veranschaulicht Abbildung 3:

## (A) Abruf von spezifischen Kausalüberzeugungen



## (B) Abruf von bereichsübergreifenden Metaüberzeugungen über Verursachung



**Abbildung 3:** Bildung von Kausalurteilen in beschriebenen Situationen (nach Shanks, 1991).

Die Theorie von Shanks (1991) bildet einen Rahmen, der spezifiziert, welche kognitiven Prozesse in Abhängigkeit vom Situationstypus ausgelöst wird. Für die Existenz assoziativer Mechanismen sprechen eine Vielzahl von Befunden (s. Allan, 1993; Dickinson, 2001; Shanks, 1993; Shanks, Holyoak & Medin, 1996) und die Vorstellungen von bereichsspezifischen und bereichsübergreifenden Kausalüberzeugungen lassen sich gut mit mechanistischen Ansätzen (Ahn et al., 1995; Koslowski, 1996, Schlottmann, 1999), der Theorie der Kausalmodelle (Waldmann & Holyoak, 1992; Waldmann, 1996; Hagmayer & Waldmann, 2006) und dem nachgewiesenen Einfluss von Erwartungen und Vorwissen auf die kausale Urteilsbildung (s. Alloy & Tabachnik, 1984; Shanks, 2004) in Einklang bringen. Shanks (1991) selbst verweist darauf, dass die Vorstellung bereichsübergreifender kausaler Grundüberzeugungen (Metawissen) konsistent mit der Behauptung von Cheng & Holyoak (1985) sei, dass Probanden aus erfahrenen Situationen Schemata und Grundüberzeugungen abstrahieren.

So wertvoll die von Shanks (1991) gemachte Unterscheidung zwischen beschriebenen und erlebten Situationen auch sein mag, sie wird der situativen Vielfalt und Komplexität, denen kausal urteilende Individuen in ihrer Lebenswelt ausgesetzt sind, nicht gerecht. Dafür sprechen die vielen unterschiedlichen Theorieansätze zum kausalen Denken (Hagmayer & Waldmann, 2006) und zu kausalen Argumenten (Oestermeier & Hesse, 2000), sowie der in

Kapitel 4 diskutierte vielfältige Einfluss von experimentellen Rahmenbedingungen wie z.B. der Fragestellung, der Informationsdarbietung oder der Informationspräsentation auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen. Darüber hinaus erscheint es nicht nur sinnvoll, die Situationen, in denen kausale Urteile zu fällen sind, mehr zu differenzieren, sondern auch die kausalen Argumente selbst. Oestermeier & Hesse (2000) haben unter anderem zu diesem Zweck eine Taxonomie von 27 kausalen Argumenten entwickelt. So liegen nach dieser Taxonomie unterschiedlichen Theorieansätzen unterschiedliche Typen von kausalen Argumenten zu Grunde. Beispielsweise basieren Kausalurteile bei regel-basierten Ansätzen auf kausalen Argumenten, in denen Angaben über die statistische Kovariation zwischen zwei oder mehreren Ereignissen gemacht werden ("A verursachte B, weil A die Wahrscheinlichkeit/das Risiko/den Prozentsatz von B erhöht"), während in mechanistischen Ansätzen auf der Grundlage von kausalen Erklärungen ("A verursachte C, weil A über den Prozess/den Mechanismus B zu C führte") argumentiert wird (s. Oestermeier & Hesse, 2000, S. 69-70).

Unklar in der Theorie von Shanks (1991) bleibt der Status regel-basierter Ansätze wie der  $\Delta P$ -Regel oder der Bestimmung der kausalen Stärke entsprechend der Power PC Theorie. Shanks (1991) selbst scheint die Auffassung zu vertreten, dass es sich dabei um Kausalüberzeugungen oder kausale Schemata handelt, doch lässt er offen, ob diese bereichsspezifisch oder bereichsübergreifend sind. Für beide Sichtweisen lassen sich Argumente finden. Für die These, dass regel-basierte Ansätze bereichsübergreifende kausale Schemata sind, spricht der hohe Allgemeinheitsgrad regel-basierter Ansätze und die Tatsache, dass regel-basierte Ansätze scheinbar auch dann zur Anwendung gelangen, wenn die Probanden über keinerlei bereichsspezifisches Vorwissen verfügten (Downing, Sternberg & Ross, 1985; Schustack & Sternberg, 1981). Für die These, dass es sich bei regel-basierten Ansätzen um bereichsspezifische kausale Schemata handelt, sprechen die großen individuellen Unterschiede, die immer wieder berichtet werden (White, 1998, 2000a) und die zahlreichen Studien, die belegen, dass der Einsatz regel-basierter Strategien bei der Beurteilung von Ursache-Wirkungs-Zusammenhängen in Abhängigkeit von präsentations- und darbietungsbezogenen Faktoren stark variieren kann (s. Kapitel 4 oder auch Oestermeier & Hesse, 2000).



## **4 Einfluss von Darbietungsart, Informationsart und Präsentationsform auf die kausale Urteilsbildung**

### **4.1 Die Bedeutung situativer und aufgabenbezogener Faktoren**

Bei der Interpretation der widersprüchlichen Befunde zu den bei der kausalen Urteilsbildung zu Grunde liegenden Mechanismen und Informationsverarbeitungsprozessen muss berücksichtigt werden, dass in den einzelnen Studien sehr unterschiedliche Paradigmen eingesetzt wurden. Dies beinhaltet die Möglichkeit, dass die gefundenen Unterschiede auf die gestellte Aufgabe und die verwendeten abhängigen Variablen zurückzuführen sind (Hagmayer & Waldmann, 2006).

Beispielsweise wurde variiert, in welcher Form die Informationen dargeboten und präsentiert wurden, entweder summarisch in Tabellen/Vierfeldertafeln (Anderson & Sheu, 1995; Arkes & Harkness, 1983; Kao & Wasserman, 1993; Ward & Jenkins, 1965; White, 2003c), Tortendiagrammen (Lober & Shanks, 2000), ungeordneten Falllisten (Hagmayer, 2001, Ex 3; Van Hamme & Wasserman, 1993; White, 2003c, Ex 1, 4 und 5) und Bildertafeln (Buehner, Cheng & Clifford, 2003) oder als eine Abfolge von Einzelfällen in Form eines "trial-by-trial" Lernens. Auch hinsichtlich der Vorgehensweise bei der Analyse von Einschätzungen des Kausalzusammenhangs unterschieden sich die Studien, meist wurden entweder Regressionen bzw. Korrelationen zwischen den Probandenurteilen und den Vorhersagen verschiedener Regeln oder Heuristiken berechnet (z.B. Ward & Jenkins, 1965; Anderson & Sheu, 1995), oder es wurden die von den Probanden eingesetzten Strategien für jede Person rekonstruiert und identifiziert (Shaklee, 1983; Shaklee & Tucker, 1980). Deshalb werfen die uneinheitlichen Befunde die Frage auf, unter welchen Umständen jeweils assoziative oder regel-basierte Mechanismen hervorgerufen werden.

Die Bedeutung von situativen und aufgabenbezogenen Faktoren wurde von einer Reihe von Autoren hervorgehoben (Allan, 1993; Baker et al., 1996; Catena et al., 1998; De Houwer & Becker, 2002; Hagmayer & Waldmann, 2006; Melchers, Üngör & Lachnit, 2005; Lagnado & Sloman, 2004; Shanks, 1991; Tangen & Allan, 2004). So nimmt beispielsweise Shanks (1991, s.a. Kapitel 3.5) an, dass wahrgenommene Kausalzusammenhänge in Abhängigkeit vom Situationstypus (erlebte oder beschriebene Situationen) unterschiedlich verarbeitet werden. Während in erlebten Situationen assoziative Lernmechanismen eingesetzt werden, werden in

beschriebenen Situationen bereichsspezifische und bereichsübergreifende Kausalüberzeugungen zur kausalen Urteilsbildung herangezogen.

Baker et al. (1996) werfen in ihrem Forschungsüberblick die Frage auf, welche Arten von Aufgaben jeweils welche informationsverarbeitenden Mechanismen in Gang setzen. Dabei kommen sie zu dem Schluss, dass Menschen bei der kausalen Urteilsbildung eher von assoziativen Mechanismen gesteuert werden, wenn der Prozess der Daten- oder Informationsgewinnung zeitlich ausgedehnt und schwierig ist (wie zum Beispiel bei einzelfallweiser Darbietung und komplexen Kausalgefügen mit mehreren Ursachen). Dagegen gehen Menschen eher regel-basiert vor, (i) wenn ihnen die Daten oder Informationen über die Ursache-Wirkungs-Zusammenhänge in einer vorkategorisierten Art und Weise präsentiert werden, oder (ii) wenn der Datengewinnungsprozess sehr einfach ist (wie zum Beispiel in den Untersuchungen von Buehner et al., 2003, wo den Probanden Informationen über monokausale Ursache-Wirkungs-Zusammenhänge in einer besonders übersichtlichen, gut strukturierten summarischen Form dargeboten wurden). Vorstrukturierte oder einfach zu gewinnende Daten führen vermutlich dazu, dass die Gedächtnisbelastung (englisch: *memory load*) reduziert wird. Ist die Gedächtnisbelastung reduziert, so können bereits existierende kausale Modelle, Schemata oder Strategien einfacher aus dem Gedächtnis abgerufen werden oder es können komplexere Strategien oder Heuristiken bei der Urteilsbildung generiert und angewendet werden.

Allan (1993) misst Variationen der Aufgabe ebenfalls einen gewissen Einfluss auf die Bildung von Kontingenzzurteilen zu. Sie weist darauf hin, dass die Aufgaben in Untersuchungen zur Bildung von Kontingenzzurteilen bei Menschen hinsichtlich mehrerer Dimensionen variierten: Fragestellung (englisch: *question asked*), eingesetzte Antwortskalen (englisch: *rating scale*) oder -typen (englisch: *type of scale*), Präsentationsmodus (englisch: *presentation mode*), Darstellung der Variablen (englisch: *representation of the variables*) und Darbietungsmodus (englisch: *trial procedure*).

Tangen & Allan (2004) haben vorgeschlagen, dass assoziative Lernmechanismen eher bei einzelfallweiser Darbietung und insbesondere bei einer hohen Anzahl von Einzelfallbeobachtungen zum Einsatz kommen. Ob Probanden assoziative Lernmechanismen oder Wissen über Kausalstrukturen einsetzen, hängt zudem von ihrem aufgabenspezifischen Vorwissen ab und davon, wonach sie gefragt werden.

Price & Yates (1995) vertreten ebenfalls die Ansicht, dass es von der Beschaffenheit der Aufgabe abhängt, ob die Beurteilung der Stärke kausaler Zusammenhänge auf assoziativen oder regel-basierten Prozessen beruht. Einer Unterscheidung von Hastie & Park (1986) aus

der Urteilsbildung im Bereich der Sozialkognition folgend, vermuten Price & Yates (1995), dass so genannte *Online-Aufgaben* zum Einsatz assoziativer Lernmechanismen ermutigen, während so genannte *Gedächtnis-gestützte Aufgaben* den Gebrauch regel-basierter Strategien anregen. Online-Aufgaben ermutigen Urteilende dazu, Reizinformationen genau dann zu integrieren, wenn sie präsentiert werden. Gedächtnis-gestützte Aufgaben hingegen ermutigen Urteilende dazu, Reizinformationen in einer Rohform abzuspeichern und sie erst dann zu integrieren, wenn von Ihnen die Bildung eines Urteils auf der Basis dieser Rohdaten verlangt wird. Zudem gehen Price & Yates (1995) ebenso wie Baker et al. (1996) davon aus, dass regel-basierte Strategien vor allem dann eingesetzt werden, wenn die Beanspruchung von Gedächtnis und Informationsverarbeitungsapparat gering ist.

Damit im Einklang steht eine Beobachtung von De Houwer & Becker (2002). Sie stellten in ihrem Forschungsüberblick fest, dass Studien, in denen nicht-assoziative Modelle klar widerlegt wurden (z.B. Dickinson & Burke, 1996; Larkin et al., 1998) sehr komplex aufgebaut waren (viele Hinweisreize oder Ursachen), während Studien, in denen assoziative Modelle klar widerlegt wurden (z.B. De Houwer, 2002; Waldmann, 2000; s.a. Buehner, Cheng & Clifford, 2003), eher einfach aufgebaut waren (wenige Hinweisreize).

Hagmayer & Waldmann (2006; s.a. Tangen & Allan, 2004) kommen in ihrem Literaturüberblick zu dem Fazit, "dass derzeit eine Reihe von Theorien [zum kausalen Denken] miteinander in Konkurrenz stehen. Dennoch zeichnet sich eher ein aufeinander Zugehen ab als ein Auseinanderstreben." ... "Ob es einmal einen integrativen Ansatz geben wird, der allen Befunden gleichermaßen gerecht wird, wird die Zukunft zeigen. Denkbar sind hier auch Entwicklungen, die von unterschiedlichen aufgabenspezifischen Mechanismen ausgehen, die auf verschiedene Aspekte von Kausalität abheben."

## **4.2 Empirische Befunde zum Einfluss von Darbietungsart, Informationsart und Präsentationsform**

Es liegt eine Vielzahl von Studien vor, die belegen, dass unterschiedliche aufgabenbezogene und mediale Faktoren die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen beeinflussen. Bevor im weiteren Verlauf dieses Abschnitts auf die empirischen Belege zum Einfluss der Faktoren *Darbietungsart*, *Präsentationsform* und *Informationsart* eingegangen wird, sollen die Bedeutungen der Begriffe *Aufgabe* und *Medium*, so wie *Darbietungsart*, *Präsentationsform* und *Informationsart* näher beschrieben und voneinander abgegrenzt werden. Was im Einzelnen unter diesen Begriffen verstanden wird, kann sich von Studie zu Studie beträchtlich unterscheiden.

Die Unterscheidung zwischen den beiden Begriffen *Aufgabe* und *Medium* scheint innerhalb der Literatur zur Kausalkognition weitestgehend ignoriert worden zu sein. Dies scheint manchmal so weit zu gehen, dass der viel häufiger verwendete Aufgabenbegriff gelegentlich mit dem selten oder gar nicht verwendeten Begriff Medium gleichgesetzt oder verwechselt wird. So bezeichnet beispielsweise Allan (1993, S.435) so unterschiedliche Dinge wie den aufgabenbezogenen Faktor *Fragestellung* oder den medialen Faktor *Präsentationsform* gleichermaßen als *Aufgabendimensionen*. Der Begriff der *Aufgabe* beinhaltet für uns eine Beschreibung dessen, *was zu tun ist*, während sich der *Medienbegriff* darauf bezieht, *wie etwas ver- oder übermittelt wird*. So einfach diese Unterscheidung zwischen Aufgabe und Medium auf dem Papier auch sein mag, in der Realität ist die Trennung nicht immer ganz einfach, wenn nicht sogar unmöglich, weil sich Aufgabe und Medium wechselseitig beeinflussen und ihre Einflüsse dementsprechend miteinander konfundiert sind. Mit anderen Worten ist es unmöglich, die Auswirkungen eines Mediums von den Auswirkungen einer Instruktionmethode zu trennen (Clark, 1994; Mayer, 2001). Werden beispielsweise Informationen summarisch in einer Tabelle oder einzelfallweise als Dia-Serie dargeboten, so beinhaltet dies nicht nur eine unterschiedliche Übermittlung von Informationen, sondern es verändert sich dadurch auch die auszuführende Aufgabe selbst, weil zur Bewältigung der Aufgabe unterschiedliche Bearbeitungsschritte notwendig werden. Umgekehrt werden unterschiedliche Aufgabenstellungen und unterschiedliche kausale Argumente durch unterschiedliche Medien unterschiedlich gut unterstützt (Oestermeier & Hesse, 2000). Aufgrund dieser begrifflichen Vermengungen kann nicht immer klar zwischen aufgaben- und medienbezogenen Faktoren unterschieden werden, so dass beide Begriffe von uns häufig gemeinsam verwendet werden.

Unter *Darbietungsart* ist zu verstehen, wie die Informationen über Ursache-Wirkungs-Zusammenhänge dargeboten werden. Konkret wurde bei Variationen des Faktors Darbietungsart untersucht, ob die Informationen über Ursache-Wirkungs-Zusammenhänge einzelfallweise oder summarisch zusammengefasst dargeboten werden. Denkbar wäre auch noch, dass die Informationen über Ursache-Wirkungs-Zusammenhänge gruppenweise dargeboten werden, doch hat eine solche Darbietungsart bislang noch kein uns bekanntes Forschungsinteresse hervorgerufen.

Verschiedene Darbietungsformen können sich darin unterscheiden, ob die fallweisen Informationen über Ursache-Wirkungs-Zusammenhänge zeitlich nacheinander oder zeitgleich gemeinsam dargeboten werden. Die Variation des Faktors Darbietungsart erfolgt also an Hand zeitlicher Kriterien. Relevant ist dabei aber nicht, wie groß der zeitliche Abstand

zwischen Ursache und Wirkung ist (siehe hierzu z.B. Anderson & Sheu, 1995; Shanks, Pearson & Dickinson, 1989; Wasserman & Neunaber; 1986), sondern ob alle Fälle simultan angehäuft dargeboten werden, oder ob die Informationen über die zu beurteilenden Ursache-Wirkungs-Zusammenhänge den Probanden einzelfallweise und nacheinander vorgelegt werden. Die Unterscheidung zwischen verschiedenen Darbietungsformen an Hand zeitlicher Kriterien (simultan vs. nacheinander) ist unvermeidlich mit der Häufigkeit der einzelnen Fälle konfundiert: Bei der einzelfallweisen Darbietung werden die Fälle immer sukzessiv und diskret dargeboten, bei der summarischen Darbietung werden alle vorhandenen Informationen über Ursache-Wirkungs-Zusammenhänge entweder zeitgleich aneinander gereiht (Auflistung von Fällen) oder in Form aufaddierter Häufigkeiten dargeboten (Summenwerte entsprechend den Zellen der Vierfeldertafel aus Abbildung 1). Eine summarische Darbietung setzt voraus, dass sich die zu beurteilenden Ursache-Wirkungs-Zusammenhänge schon ereignet haben, bei einer Einzelfalldarbietung können die Einzelfälle entweder unmittelbar in der Gegenwart beobachtet werden (Live-Darbietung), oder sie können sich ebenfalls schon ereignet haben und entweder originalgetreu (Aufzeichnung) oder komprimiert in beschriebener Form (Zusammenfassung) wiedergegeben werden.

Der Begriff der *Darbietungsart* wird in der Literatur zur Kausalkognition gelegentlich mit dem Begriff der *Präsentationsform* gleichgesetzt (s. z.B. Hagmayer, 2001). Das liegt sicherlich an der semantischen Ähnlichkeit zwischen den Begriffen *Darbietung* und *Präsentation* einerseits und *Art* und *Form* andererseits. Es gibt allerdings Gründe, die aus unserer Sicht eine Differenzierung zwischen Darbietungsart und Präsentationsform notwendig machen. Zum einen lassen sich unterschiedliche Präsentationsformen wie zum Beispiel Text oder Bild nicht mit Hilfe von zeitlichen Kriterien begrifflich auseinander halten. Zum anderen ist die Unterscheidung zwischen einzelfallweiser und summarischer Darbietung sehr grob und sagt noch nichts darüber aus, mit welchem Medium eine einzelfallweise oder summarische Darbietung operationalisiert wird. So können sich verschiedene Studien unter anderem darin unterscheiden, ob bei einer summarischen Darbietung Texte, Falllisten, Tabellen, Bilder, Bildertafeln, Stapel- oder Tortendiagramme zum Einsatz kommen und ob eine einzelfallweise Darbietung mit Hilfe von Texten, Zeichnungen, Bildern, Photos, Animationen, Videos, Simulationen oder mit unmittelbar beobachtbaren Objekten, Gegenständen oder Ereignisfolgen vorgenommen wird (s.a. Oestermeier & Hesse, 2000).

Angesichts der Tatsache, dass die Unterscheidung zwischen einzelfallweiser und summarischer Darbietung zu grob ist und sich verschiedene Studien meist beträchtlich darin unterscheiden, wie eine einzelfallweise oder summarische Darbietung operationalisiert wird,

erscheinen bei der Interpretation von Untersuchungen zum Einfluss des Faktors Darbietungsart weitere Differenzierungen im Hinblick auf das eingesetzte Medium oder die eingesetzten Medien sinnvoll. Um für weitere Studien die Möglichkeit zu schaffen, verschiedene Operationalisierungen von ein- und derselben Darbietungsart miteinander zu vergleichen, soll hier der Begriff der *Präsentationsform*<sup>2</sup> als eine neue, von der Darbietungsart unabhängige Kategorie hinzugefügt werden. Unter *Präsentationsform* soll das jeweilige *Medium* verstanden werden, in welchem den Probanden während der Experimentdurchführung die Informationen über Ursache-Wirkungs-Zusammenhänge vermittelt bzw. einzelfallweise oder summarisch dargeboten werden.

Der Medienbegriff wird sehr vielfältig und facettenreich eingesetzt (zum Begriff des Mediums s. Ballstaedt, 1997; Salomon, 1979; Weidenmann, 2006). Gemeinsam ist allen Medien, dass sie Inhalte oder Informationen übermitteln und dabei auf unterschiedliche Darstellungsformen, Zeichensysteme und Technologien zurückgreifen und mit verschiedenen situativen Kontexten assoziiert sind (Salomon, 1979). Von besonderer Bedeutung für die vorliegende Arbeit ist die Verwendung des Medienbegriffs als Oberbegriff für verschiedene externale Darstellungsformen (semiotischer Aspekt des Medienbegriffs). Innerhalb der Wissensvermittlung unterscheidet Ballstaedt (1997) folgende Kategorien von externalen Darstellungsformen: Texte, Charts, Tabellen, Diagramme, Abbilder und Piktogramme. Neben den Präsentationsformen oder -arten unterscheidet Mayer (1997) in Verbindung mit multimedialem Lernen noch die Übertragungsmedien (z.B. Papier und Bleistift, PC, Dias, ...etc.) und die Sinnesmodalitäten (z.B. auditorisch, visuell).

Im Gegensatz zu den Einflüssen der Faktoren *Darbietungsart* und *Präsentationsform* wurde der Einfluss des Faktors *Informationsart* nach unserem Kenntnisstand im Bereich der Kausalkognition bislang noch nicht untersucht. Der Einfluss der Informationsart wurde untersucht, wenn es darum ging zu erklären, warum menschliche Wahrscheinlichkeitsurteile von den Gesetzen der Wahrscheinlichkeit abweichen (Cosmides & Tooby; 1996; Fiedler, 1988; Gigerenzer & Hoffrage, 1995). Gigerenzer & Hoffrage (1995) haben beispielsweise im Zusammenhang mit dem Bayes-Theorem untersucht, ob die Prozentzahl von Probanden, die bei der Lösung statistischer Probleme Bayesianische Algorithmen einsetzten, vom Einfluss der Informationsart abhängt. Der Einfluss des Faktors *Informationsart* wurde dabei variiert,

---

<sup>2</sup> Gelegentlich findet sich in der deutschsprachigen Literatur (Ballstaedt, 1997) auch der Begriff der *Darstellungsform*, den ich hier synonym zum Begriff der *Präsentationsform* verstehe. Da in der englischsprachigen Literatur häufig der Begriff *presentation form* verwendet wird (s. z.B. Oestermeier & Hesse, 2000, S. 81) habe ich die unmittelbarere Übersetzung *Präsentationsform* gewählt.

indem den Probanden die Informationen zu den statistischen Problemen entweder als Wahrscheinlichkeiten oder als Häufigkeiten präsentiert wurden.

Während der Faktor *Präsentationsform* den Einfluss verschiedener Medien auf die Beurteilung der Stärke kausaler Zusammenhänge erfasst, bezieht sich der Faktor *Informationsart* auf die Informationen über die Kausalzusammenhänge selbst und nicht auf deren Vermittlung. Die in einer Untersuchung ausgewählte Informationsart kann nicht unabhängig von der gewählten Darbietungsart gesehen werden. Bei einzelfallweiser Darbietung können die Kausalinformationen logischerweise nur (verbale oder bildhafte) Aussagen über einzelne Kausalzusammenhänge beinhalten, weshalb eine Variation des Faktors *Informationsart* bei Einzelfalldarbietung unmöglich ist. Bei summarischer Darbietung können die Daten aber auf eine unterschiedliche Art und Weise aufbereitet sein, beispielsweise als Auflistung von Einzelaussagen, als Häufigkeiten, als Häufigkeitsanteile oder als Wahrscheinlichkeiten. Bei summarischer Darbietung ist der Einfluss des Faktors *Informationsart* mit dem Einfluss der *Präsentationsform* konfundiert, weil unterschiedliche Präsentationsformate wie zum Beispiel Texte, Bilder, Tabellen, Diagramme eine unterschiedliche Datenaufbereitung beinhalten oder erforderlich machen können. So machen Einzelaussagen im Gegensatz zu Häufigkeiten in einem Stapeldiagramm wenig Sinn oder sind sogar unvereinbar mit diesem Medium, während sie in Protokolllisten durchaus passend sind.

Nach diesem Versuch, die drei Faktoren *Darbietungsart*, *Präsentationsform* und *Informationsart* begrifflich klarer voneinander abzugrenzen, sollen nun empirische Befunde zu den Einflüssen dieser drei Faktoren referiert und diskutiert werden. An dieser Stelle fehlen Ergebnisse und Überblicke zu den Auswirkungen anderer medialer oder aufgabenbezogener Faktoren wie zum Beispiel der *Lernrichtung* (prädiktiv vs. diagnostisch, s. De Houwer & Beckers, 2002; Matute, Arcediano & Miller, 1996; Perales, Catena & Maldonado, 2004; Waldmann, 2000, 2001; Waldmann & Holyoak, 1992); der *Art der Rating-Skala*, der *Fragestellung der Untersuchung und der Wortwahl* (Beyth-Marom, 1982; Buehner, Cheng & Clifford, 2003; Collins & Shanks, 2006; Crocker, 1982; Perales & Shanks, in press; Shanks, 2004; White, 2003b); des *Referenzrahmens oder der Instruktion* (s. Kahneman, Slovic & Tversky, 1982); des *Situationstyps* (erlebt vs. beschrieben, s. Shanks, 1991; Van Hamme & Wassermann, 1993; Hagmayer & Waldmann, 2002; Lagnado & Sloman, 2004); des *Paradigmas* (kontinuierlich vs. deskriptiv, Anderson & Sheu, 1995); des *zeitlichem Abstands zwischen Ursache und Wirkung* (Allan et al., in press; Buehner & May, 2002, 2003a, 2003b, 2004; Hagmayer & Waldmann, 2002; Mendelson & Shultz, 1976; Shanks & Dickinson, 1987;

Shanks et al., 1989; Shultz, 1982); der *Abfolge der Lerndurchgänge* (Collins & Shanks, 2002; Dennis & Ahn, 2001; Hogarth & Einhorn, 1992; Catena, Maldonado, Megías & Frese, 2002; López, Cobos, Cano & Shanks, 1998; López, Shanks, Almaraz, Fernández, 1998; Yates & Curley, 1986); des *Antwortmodus* (Catena, Maldonado & Cándido, 1998; Neunaber & Wasserman, 1986; White, 2003a, 2004); der *Urteilshäufigkeit / des Urteilszeitpunktes* (Catena, Maldonado & Cándido, 1998; Catena, Perales & Maldonado, 2004; Collins & Shanks, 2002; De Houwer & Beckers, 2002) oder der *Salienz von Zellhäufigkeiten* (Arkes & Harkness, 1983, Ex 6; White 2003b, Ex 2 und 3).

Außerdem beschränken wir uns im Rahmen dieser Arbeit weitestgehend auf Studien, in denen die Probanden monokausale Ursache-Wirkungs-Zusammenhänge zu beurteilen hatten, da sich die Probanden in den drei von uns durchgeführten Experimenten ausschließlich mit der Beurteilung der Stärke von monokausalen, probabilistischen Ursache-Wirkungs-Zusammenhängen auseinanderzusetzen hatten.

#### 4.2.1 Einfluss der Darbietungsart

Nur wenige Autoren haben den Einfluss der Darbietungsart auf die Beurteilung der Stärke von kausalen Zusammenhängen in einem Experiment untersucht. Bereits 1965 interessierten sich Ward & Jenkins für die Frage, welchen Einfluss es hat, wenn Informationen über Ursache-Wirkungs-Zusammenhänge einzelfallweise nacheinander oder in summarischer Form dargeboten werden. Die Probanden hatten dabei die Aufgabe, auf einer Skala von 0 bis 100 das Ausmaß von Kontrolle zu beurteilen, welches das Impfen von Wolken auf das Eintreten von Regnen hat. Es zeigte sich, dass bei summarischer Datendarbietung in einer Tabelle 75 % und bei einzelfallweiser Darbietung 17 % der Probanden von der  $\Delta P$ -Regel Gebrauch machten. Wegen der großen individuellen Unterschiede haben Ward & Jenkins (1965) auf Mittelwertvergleiche verzichtet. Auch Kao & Wasserman (1993) kamen in ihrem Experiment 2 zu dem Ergebnis, dass bei summarischer Ergebnisdarbietung mehr Probanden die  $\Delta P$ -Regel einsetzen als bei einzelfallweiser Ergebnisdarbietung. Außerdem fielen die Kausalurteile unter der Einzelfallbedingung ungenauer aus, als unter der summarischen Bedingung. Kao & Wasserman (1993) fanden ebenfalls heraus, dass eine gewichtete  $\Delta P$ -Regel die Kontingenzzurteile der Probanden besser vorhersagt als eine ungewichtete (s.a. Anderson & Sheu, 1995; Levin, Wasserman & Kao; Wasserman, Dorner & Kao, 1990). Kao & Wassermann (1993) konnten dabei feststellen, dass bei einzelfallweiser Darbietung die ungleichen Zellgewichtungen ausgeprägter sind als bei tabellarischen Ergebniszusammenfassungen. Dies spricht dafür, dass bei summarischer Ergebnisdarbietung



mehr Probanden von einer  $\Delta P$ -Regel mit gleichen Zellgewichten Gebrauch machen. In der Tat fanden Kao & Wasserman (1993) heraus, dass bei einzelfallweiser Datendarbietung keiner ihrer 200 Probanden von der  $\Delta P$ -Regel Gebrauch machte.

Es gibt aber auch Studien, die zu belegen scheinen, dass die Darbietungsart keinen signifikanten Einfluss auf die Beurteilung von Ursache-Wirkungs-Zusammenhängen hat (Baker, Berbier & Vallée-Tourangeau, 1989; Hagmayer, 2001, Ex 3; Melchers, Üngör & Lachnit, 2005; Van Hamme & Wassermann, 1993). Dazu muss aber angemerkt werden, dass die Probanden in all diesen Studien bei der Urteilsbildung mehr als eine Ursache zu berücksichtigen hatten, wodurch die Komplexität der Aufgabe erhöht wurde. Außerdem wurden in den Experimenten von Hagmayer (2001) und Van Hamme & Wasserman (1993) Falllisten mit einzeln nacheinander dargebotenen Kausalinformationen verglichen. In Falllisten werden die Ereignisse in der Regel ungeordnet nacheinander präsentiert. Wenn ein Proband dabei die von einer komplexeren Strategie wie zum Beispiel der  $\Delta P$ -Regel Gebrauch machen möchte, dann muss er die einzelnen Fallinformationen erst entsprechend den vier Feldern einer  $2 \times 2$  – Kontingenztafel gruppieren und aufaddieren. Im Vergleich zu anderen summarischen Darbietungsformen wie Tabellen oder Stapeldiagrammen, wo die Informationen über Ursache-Wirkungs-Zusammenhänge bereits aufaddiert und geordnet sind, wird die Informationsverarbeitung bei Falllisten aufgrund der Unstrukturiertheit des dargebotenen Materials erschwert. Zudem ermöglicht eine Fallliste im Gegensatz zu einer  $2 \times 2$  – Kontingenztafel mit Zellhäufigkeiten prinzipiell den Einsatz assoziativer Mechanismen, da die Einzelinformationen nacheinander durchgegangen und die Assoziationsstärken einzelfallweise adjustiert werden können.

Einige Autoren (Arkes & Harkness, 1983; Buehner et al., 2003; Lober & Shanks, 2000; White, 2003c) haben verschiedene Experimente mit gleichem Design durchgeführt, die sich nur hinsichtlich der Darbietungsart voneinander unterschieden, und deren Ergebnisse miteinander verglichen. In der Untersuchungsreihe von Arkes & Harkness (1983) zeigte es sich, dass die Probanden bei der summarischen Informationsdarbietung in Experiment 4 von komplexeren Strategien ( $\Delta D$ : 51 %;  $\Delta P$  28 %) Gebrauch machten als bei der einzelfalldarbietung in Experiment 5 Zelle a: 34 %; a versus b: 26 %;  $\Delta D$ ,  $\Delta P$ , unklar je 13 %). Diese Unterschiede in der Strategiewahl erklären sich Arkes & Harkness (1983) durch die im Vergleich zu Experiment 4 gestiegenen Gedächtnisanforderungen in Experiment 5.

Lober & Shanks (2000) verglichen das RWM, das PCM und die Power PC Theorie miteinander. Hier unterschieden sich die Experimentenpaare 1 und 4, 2 und 5, sowie 3 und 6 nur hinsichtlich der Darbietungsart. Während die Kausalinformationen den Probanden in den

Experimenten 1 bis 3 einzelfallweise dargeboten wurden, erhielten die Probanden in den Experimenten 4 bis 6 eine summarische Darbietung der Daten in zwei nebeneinander liegenden Tortendiagrammen vorgelegt. Lober & Shanks (2000) kamen nach Sichtung aller Ergebnisse zu dem Schluss, dass das RWM nur bei einzelfallweiser Datendarbietung eingesetzt werden kann und dass unter bestimmten Umständen Unterschiede zwischen den beiden Darbietungsarten gefunden werden können. Das PCM konnte die Ergebnisse in allen Experimenten gut vorhersagen, allerdings mussten bei einzelfallweiser Darbietung Erfolgs- und Basisrate unterschiedlich gewichtet werden. Die Power PC Theorie konnte dagegen nur die Ergebnisse in Experiment 3 gut vorhersagen, sonst nicht.

In den Experimenten von Buehner et al. (2003) wurde die Power PC Theorie sowohl bei summarischer Darbietung (Experiment 2) als auch einzelfallweiser Informationsdarbietung (Experiment 3) gestützt, auch wenn sich bei Einzelfalldarbietung ein Dichte-Bias bei der Beurteilung nicht-kontingenter Ursachen einstellte. Buehner et al. (2003) sehen die Richtigkeit der Power PC Theorie dadurch nicht in Frage gestellt, sondern erklärten sich den Dichte-Bias durch eine Fehleinschätzung von  $\Delta P$  bei einer Probandenminderheit. Bedingt wird diese Fehleinschätzung durch die gestiegenen Gedächtnisanforderungen bei einzelfallweiser Darbietung.

White (2003c) verglich in drei verschiedenen Experimenten mit drei unterschiedlichen Darbietungsarten und gleichem Design seine Theorie der Evidenzevaluation mit der Delta P-Regel. Dabei wurden die Zellinformationen den Probanden in Experiment 1 in Falllisten, in Experiment 2 einzelfallweise und in Experiment 3 summarisch in einer Vierfeldertafel vorgelegt. Unabhängig von der Darbietungsart ergaben sich in allen drei Experimenten signifikante Effekte für den Einfluss von Manipulationen des Anteils konfirmatorischer Beobachtungen ( $pCI$ ), während sich für Manipulationen von  $\Delta P$  nur bei summarischer Informationsdarbietung ein signifikanter Effekt ergab.

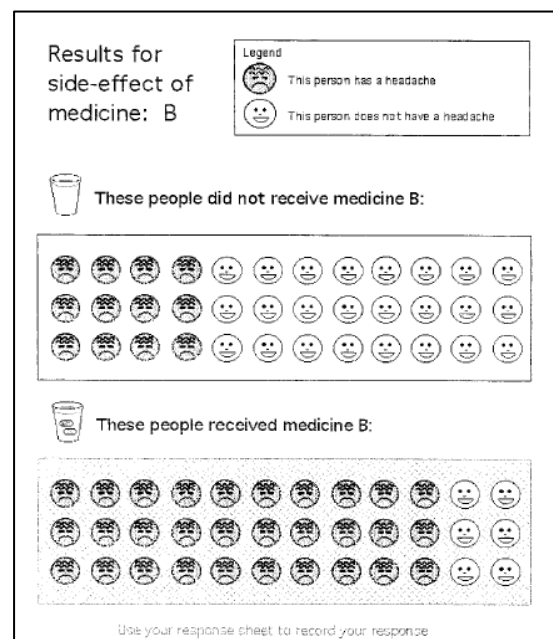
Der Einfluss unterschiedlicher Darbietungsarten auf den Einsatz oder die Urteilsgenauigkeit in Verbindung mit der Power PC-Theorie (Cheng, 1997) oder der Theorie der empirischen Evidenzevaluation (White, 2003c; 2004) erscheint auf der Grundlage der Ergebnisse von Buehner et al. (2003) und White (2003c) unklar und wurde bislang noch nicht in einem Experiment untersucht.

#### *Diskussion der Befunde zum Einfluss des Faktors Darbietungsart*

White (2003c) stellt fest, dass bislang nur wenige Studien existieren, in denen unterschiedliche Arten der Darbietung oder Präsentation von Informationen über Ursache-Wirkungs-Zusammenhänge direkt miteinander verglichen wurden, und dass mehr über die

Gründe für Unterschiede zwischen unterschiedlichen Darbietungs- oder Präsentationsformaten in Erfahrung gebracht werden muss. Die vorliegenden Studien zum Einfluss des Faktors Darbietungsart bringen widersprüchliche Ergebnisse hervor. Einige Studien (Arkes & Harkness, 1983; Buehner et al., 2003; Kao & Wasserman, 1993; Lober & Shanks, 2000; Ward & Jenkins, 1965) sprechen dafür, dass die Probanden bei summarischer Darbietung eher regel-basierte Strategien wie die  $\Delta P$ -Regel oder auch die Power PC-Theorie einsetzen und genauere Kontingenzurteile bilden, während andere (Baker, Barbier & Vallée-Tourangeau, 1989; Hagmayer, 2001; Van Hamme & Wasserman, 1993) zum Ergebnis kommen, dass Einzelfall- und summarische Darbietung sich nicht hinsichtlich ihres Einflusses auf die Beurteilung von Ursache-Wirkungs-Zusammenhängen unterscheiden.

Bei der Interpretation dieser widersprüchlichen Ergebnisse ist aber zum einen zu berücksichtigen, dass – wie schon erwähnt - die Unterscheidung einzelfallweise versus summarisch viel zu grobmaschig ist, weil sich die in den verschiedenen Studien eingesetzten Medien und Präsentationsformate meist beträchtlich hinsichtlich einer Reihe von Merkmalen und Eigenschaften wie zum Beispiel Übersichtlichkeit, Salienz relevanter Informationen, Einsatz grafischer Hilfsmittel oder Struktur unterscheiden. Zum anderen muss beachtet werden, dass in den vorliegenden Studien neben der Darbietungsart meist auch andere aufgaben- und medienbezogene Faktoren, wie z.B. Fragestellung, Antwortmodus oder Präsentationsformat einen komplexen Einfluss auf die Strategiewahl oder die Urteilsgenauigkeit haben. Dadurch ist davon auszugehen, dass unabhängig von der Auswahl der zu manipulierenden Variablen in jeder Studie die Einflüsse mehrerer aufgaben- und medienbezogener Faktoren miteinander konfundiert sind. Besonders gut lässt sich das Zusammenwirken der Faktoren Darbietungsart und Präsentationsform am Stimulus-Material veranschaulichen, das in Experiment 2 aus der Studie von Buehner, Cheng & Clifford (2003) eingesetzt wurde. Dort wurden die summarisch dargebotenen Informationen zur Beurteilung des kausalen Zusammenhangs auf übersichtlich gestalteten Bildertafeln präsentiert. Wie Abbildung 4 zeigt, waren die einzelnen Bildertafeln in zwei Gruppen (Ursache gegeben



**Abbildung 4:**

Schwarz - Weiß - Version des Stimulus-Materials von Buehner et al. (2003, Ex 2).

– Ursache nicht gegeben) mit je gleich vielen Einzelfällen ( $n = 36$ ) unterteilt worden, das Eintreten oder Nichteintreten der Wirkung (Kopfschmerz) wurde entweder durch ein weißes und freundlich lächelndes Gesicht oder durch ein blaues und finster drein schauendes Gesicht symbolisiert. Pro Gruppe waren die Bildertafeln in drei Reihen ( $3 \times 12$  Einzelfälle) angeordnet, zudem befanden sich die finsternen Gesichter alle auf der linken, und die freundlichen Gesichter alle auf der rechten Seite. Außerdem ließen sich die vier Zellgrößen a, b, c und d rechnerisch einfach zueinander ins Verhältnis setzen, weil die diesen Zellgrößen entsprechenden Summenwerte allesamt durch die Zahl 3 teilbar waren.

Unklar ist, *wie* sich der Faktor Darbietungsart auf die Bildung kausaler Urteile, auf die Strategiewahl oder die Urteilsgenauigkeit auswirkt. Einig scheint man sich bislang nur darüber zu sein, dass bei einer summarischen im Vergleich zur einzelfallweisen Darbietung die Anforderungen an das Gedächtnis reduziert sind (s. z.B. Arkes & Harkness, 1983; Baker et al., 1996; Kao & Wasserman, 1993; Shaklee & Mims, 1982). Auch Buehner, Cheng & Clifford (2003) setzten die summarisch dargebotenen und übersichtlich gestalteten Bildertafeln dazu ein, um die Gedächtnisbeanspruchung zu reduzieren.

Studien, in denen kein Unterschied zwischen summarischer und Einzelfalldarbietung gefunden werden konnte (Baker, Barbier & Vallée-Tourangeau, 1989; Hagmayer, 2001; Melchers, Üngör & Lachnit, 2005; Van Hamme & Wasserman, 1993) müssen nicht im Widerspruch zu der Annahme stehen, dass die Gedächtnisbeanspruchung von der Darbietungsart abhängt, da davon auszugehen ist, dass die Gedächtnisbeanspruchung in einem Experiment auch von anderen aufgaben- und medienbezogenen Faktoren beeinflusst wird. In der Tat spricht einiges dafür, dass in den Experimenten von Baker, Barbier & Vallée-Tourangeau (1989), Hagmayer (2001), Melchers, Üngör & Lachnit (2005), sowie von Van Hamme & Wasserman (1993) die Gedächtnisbeanspruchung für die Probanden auch in der summarischen Bedingung relativ hoch war, unter anderem weil bei der Urteilsbildung mehrere Ursachen berücksichtigt werden mussten und die Aufgabenstellung somit relativ komplex war. Zudem wurden den Probanden die Informationen über Ursache-Wirkungs-Zusammenhänge unter der summarischen Bedingung in Falllisten dargeboten. In einer Fallliste sind die Informationen über Ursache-Wirkungs-Zusammenhänge im Vergleich zu anderen summarischen Darbietungsformen wie Diagramm oder Tabelle eher unübersichtlich angeordnet, weil die Daten lediglich aneinandergereiht, aber nicht gruppiert und aufaddiert sind. Für den Nachweis eines Einflusses des Faktors Darbietungsart scheint es also wesentlich zu sein, dass bei einer summarischen Darbietung die Gedächtnisbeanspruchung auch wirklich reduziert wird.

Angesichts der vorliegenden Befunde erscheint die Vorstellung sehr plausibel, dass Variationen der Darbietungsart und wahrscheinlich auch anderer aufgaben- und medienbezogener Faktoren eine unterschiedliche Beanspruchung des Gedächtnisses nach sich ziehen. Es besteht aber auch die Möglichkeit, dass durch Variationen der Darbietungsart völlig unterschiedliche kognitive Prozesse (regel-basiert oder assoziativ) in Gang gesetzt werden können (Shanks, 1991; Lober & Shanks, 2000). Nach einer solchen Interpretation wird die Verarbeitung von Informationen über Ursache-Wirkungs-Zusammenhänge nicht nur - wie bei einer rein gedächtnistheoretischen Fundierung - moduliert, sondern durch die Variation der Darbietungsart oder anderer aufgaben- und medienbezogener Faktoren gesteuert.

Zusammenfassend ergibt sich das Bild, dass bei summarischer Informationsdarbietung mehr Personen regel-basierte Strategien wie die  $\Delta P$ -Regel einsetzen und genauere Kontingenzzurteile erzielt werden. Voraussetzung dafür ist, dass die Komplexität der Aufgabenstellung angesichts der unvermeidlichen Konfundierung mit anderen aufgaben- und medienbezogenen Einflüssen nicht zu groß ist. Wenn die Komplexität der Aufgabenstellung durch eine summarische Darbietung nicht ausreichend reduziert werden kann, führen die Probanden entweder regel-basierte Strategien fehlerhaft aus oder sie setzen einfachere Heuristiken wie die ab-Regel ein. Alternativ dazu können unter bestimmten Bedingungen (z.B. summarische Darbietung in einer Fallliste) auch assoziative Mechanismen in Gang gesetzt werden. Bei einzelfallweiser Informationsdarbietung steigen die Komplexität der Aufgabenstellung und die Gedächtnisbeanspruchung. Das beinhaltet, dass der Einsatz regel-basierter Strategien wie der  $\Delta P$ -Regel im Vergleich zur summarischen Darbietung unwahrscheinlicher, wenn auch nicht völlig ausgeschlossen wird. Einige Studien belegen, dass auch bei Einzelfalldarbietung genaue Kontingenzeinschätzungen vorgenommen werden können (Anderson & Sheu, 1995; Baker, Barbier & Vallée-Tourangeau, 1989; Shanks, 1993; Shanks et al., 1996; Spellman, 1996; Wasserman et al., 1993), doch meist sagt hier eine gewichtete  $\Delta P$ -Regel die Daten besser vorher als eine ungewichtete (Anderson & Sheu, 1995; Kao & Wasserman, 1993; Lober & Shanks, 2000, Experimente 1 bis 3).

#### 4.2.2 Einfluss der Informationsart

Wie bereits erwähnt, wurde der Einfluss des Faktors *Informationsart* nach unserem Kenntnisstand im Bereich der Kausalkognition bislang noch nicht untersucht. Im Bereich der psychologischen Forschung zur Entscheidungsbildung (Jungermann, Pfister & Fischer, 2005)

konnte aber gezeigt werden, dass es eine Rolle spielt, ob Informationen über Wahrscheinlichkeiten und Werte in numerischer oder in verbaler Form angeboten werden. So fanden beispielsweise Jungermann & Fischer (1996) erhebliche Unterschiede in der Beurteilung der Risiken und der Wahl von Medikamenten, je nachdem ob die Häufigkeit der Nebenwirkungen verbal oder numerisch beschrieben war. Dabei zeigte es sich auch, dass die Interpretation verbaler Häufigkeitsausdrücke (wie z.B. "selten", "gelegentlich" oder "häufig") sehr stark vom Kontext abhängt, in dem die Wörter verwendet werden.

Der Einfluss der Informationsart hat sich darüber hinaus auch noch als bedeutsam erwiesen, wenn es darum ging, das Auftreten von Urteilsverzerrungen bei der Lösung statistischer Probleme zu erklären. Eine Reihe von Untersuchungen (Cosmides & Tooby; 1996; Fiedler, 1988; Gigerenzer & Hoffrage, 1995) kam dabei zu dem Ergebnis, dass Urteilsverzerrungen in Verbindung mit dem Bayes-Theorem deutlich reduziert werden können, wenn den Probanden die Probleme nicht – wie üblich – als Wahrscheinlichkeiten präsentiert werden, sondern als Häufigkeiten. In der Studie von Gigerenzer & Hoffrage (1995) wendeten ca. 48 % der Probanden Bayesianische Algorithmen unter der Häufigkeitsbedingung an, in der Wahrscheinlichkeitsbedingung waren es nur rund 22 %. Gigerenzer & Hoffrage (1995) werteten diese Resultate als Beleg dafür, dass mathematisch äquivalente Repräsentationen von Information kognitive Algorithmen mit einer nicht äquivalenten Berechnungsgrundlage zur Folge haben können. Das bedeutet, dass kognitive Algorithmen nicht unabhängig von der Art der dargestellten Information gesehen werden können. Gigerenzer & Hoffrage (1995) vermuten, dass Häufigkeitsformate das Vorkommen von Urteilsverzerrungen nicht nur in Verbindung mit dem Bayes-Theorem verringern können, sondern auch bei anderen statistischen Überlegungen.

Bei einem Vergleich zwischen dem Vorgehen bei der Anwendung regel-basierter Strategien und dem bei Gigerenzer & Hoffrage (1995) beschriebenen "Häufigkeiten-Ansatz" (Lagnado & Shanks, 2002) zur Lösung statistischer Probleme lassen sich Parallelen finden. So sammeln die Probanden auch bei regel-basierten Ansätzen Häufigkeiten und gruppieren diese, ehe dann probabilistische Berechnungen mit den geordneten Informationen durchgeführt werden. Ein wesentlicher Unterschied zum Häufigkeiten-Ansatz bei Gigerenzer & Hoffrage (1995) besteht aber darin, dass bei regel-basierten Ansätzen die Häufigkeiten vor der Differenzbildung zwischen Erfolgs- und Basisrate erst einmal in Wahrscheinlichkeiten umgerechnet werden, während beim Häufigkeiten-Ansatz Regeln eingesetzt werden, die unmittelbar auf die absoluten Häufigkeiten Bezug nehmen.

Zu vermuten ist auch, dass der Faktor Informationsart stark mit der Antwortskala in einer Studie interagiert. Lagnado & Shanks (2002) interessierten sich in ihrer Studie für den Einfluss der Antwortskala. Sie boten ihren Probanden Häufigkeiten dar und fanden heraus, dass Probanden, die Häufigkeitsurteile bilden sollten, kohärentere Urteile bildeten als die Probanden, die Wahrscheinlichkeitsurteile bilden sollten. Anders als bei Lagnado & Shanks (2002) müssen die Probanden in Studien zur Beurteilung der Stärke monokausaler Ursache-Wirkungs-Zusammenhängen allerdings häufig Wahrscheinlichkeitsurteile abgeben, indem sie die Stärke eines Kausalzusammenhangs auf einer Skala von 0 bis 100 (oder -100 bis 100 oder 0 bis -100) einschätzen müssen (Allan & Jenkins, 1980; Alloy & Abramson, 1979; Anderson & Sheu, 1995; Arkes & Harkness, 1983; Buehner, Cheng & Clifford, 2003; Catena, Maldonado, Megías & Frese, 2002; Jenkins & Ward, 1965; Lober & Shanks, 2000; Perales & Shanks, 2003; Schustack & Sternberg, 1981; Vallée-Tourangeau, Murphy, Drew & Baker, 1998; Ward & Jenkins, 1965; Wasserman, 1990; White, 2003b).

Wenn die Annahme zutrifft, dass die Einflüsse von Informationsart und Antwortskala miteinander interagieren, dann wäre zu erwarten, dass die Wahrscheinlichkeitsurteile der Probanden genauer ausfallen, wenn sie Wahrscheinlichkeiten zu beurteilen haben, als wenn sie Wahrscheinlichkeitsurteile auf der Basis von Häufigkeiten abzugeben haben. Der Grund hierfür liegt darin, dass die Probanden bei Wahrscheinlichkeiten im Vergleich zu Häufigkeiten weniger Rechenschritte durchführen müssen, da die Probanden die Häufigkeiten erst verrechnen und in Prozentzahlen umwandeln müssen, um auf einer von 0 bis 100 reichenden Antwortskala ein Urteil oder eine Einschätzung abzugeben. Zusammengefasst lässt sich daraus in Verbindung mit dem Einfluss der Informationsart auf die Beurteilung der Stärke kausaler Zusammenhänge die Vermutung ableiten, dass die Urteilsgenauigkeit von der Ergebnisnähe der Art der zu beurteilenden Information abhängt.

### **4.2.3 Einfluss der Präsentationsform**

Innerhalb der Wissensvermittlung unterscheidet Ballstaedt (1997) folgende Kategorien von externalen Darstellungs- oder Präsentationsformen: Texte, Charts, Tabellen, Diagramme, Abbilder und Piktogramme. Diese Kategorien lassen sich wiederum in Subtypen unterteilen. Trotz der Vielfalt an unterschiedlichen Präsentationsformen, wurde deren Einfluss auf die Beurteilung von Ursache-Wirkungs-Zusammenhängen bislang kaum direkt und unabhängig vom Einfluss der Darbietungsart untersucht. Wasserman & Shaklee (1984, Ex 2 und 4)

verglichen ein Zeitlinienformat<sup>3</sup> mit einem Tabellenformat und fanden heraus, dass die Urteile der Probanden, denen die Daten tabellarisch präsentiert wurden, genauer ausfielen und besser durch die  $\Delta P$ -Regel vorhersagbar sind. Diese Unterschiede können nach Wasserman & Shaklee (1984) nicht durch eine unterschiedliche Gedächtnisbeanspruchung erklärt werden, sondern müssen andere Gründe haben. Allerdings konnte auch die von Wasserman & Shaklee (1984) gelieferte Erklärung, dass die Probanden die Kausalinformationen in den beiden Bedingungen unterschiedlich organisieren, die Ergebnisse nur teilweise erklären. Kritisch zum Experiment von Wasserman & Shaklee (1984) ist anzumerken, dass sich die beiden Experimentalbedingungen nicht nur hinsichtlich des Präsentationsformats unterschieden, sondern auch hinsichtlich der Darbietungsart, da den Probanden, denen eine Zeitlinie präsentiert wurde, die Daten einzelfallweise dargeboten wurden, während den Probanden, die eine Tabelle bearbeiteten, die Daten summarisch dargeboten wurden.

Shaklee & Wasserman (1986) verglichen ebenfalls Zeitlinie und Tabelle miteinander. Allerdings wurde, anders als bei Wasserman & Shaklee (1984), bei der Konstruktion und Auswertung des Experiments ein regel-analytischer Ansatz zugrunde gelegt (Shaklee & Hall, 1983; Shaklee & Mims, 1981, 1982; Shaklee & Tucker, 1980). Insgesamt zeigte sich, dass nur wenige Probanden von der  $\Delta P$ - oder der  $\Delta D$ -Regel Gebrauch machten ( $\Delta P$ : 3,1 %;  $\Delta D$ : 8,1 %). Dennoch unterschieden sich Zeitlinie und Tabelle im Hinblick auf den Einsatz einfacherer Strategien signifikant voneinander. Während in den beiden Tabellenbedingungen 57,5 % der Probanden die a versus b Strategie und 20 % die Zelle a Strategie einsetzten, waren es in den beiden Linienbedingungen nur 18,75 % (a versus b Strategie) und 10 % (Zelle a Strategie). Unklassifizierbare oder keine regel-basierten Strategien kamen dagegen in den beiden Linienbedingungen (60 % der Probanden) häufiger als in den beiden Tabellenbedingungen (11,25 % der Probanden) vor.

Die Unterschiede in Abhängigkeit von der Präsentationsform führen Shaklee & Wasserman (1986) wie Wasserman & Shaklee (1984) auf die Konfundierung mit dem Faktor Darbietungsart zurück. Sie argumentieren, dass die Probandenurteile bei Einzelfalldarbietung in den beiden Linienbedingungen auf Schätzungen basieren, bei summarischer Darbietung in einer Tabelle aber auf einfachen Regeln. Die offene Frage, warum im Gegensatz zu vielen anderen Studien (Arkes & Harkness, 1983; Shaklee & Hall, 1983; Shaklee & Tucker, 1980, Ex 2; Ward & Jenkins, 1965) nur so wenige Probanden auch in den tabellarischen Bedingungen mit summarischer Datendarbietung von anspruchsvolleren Strategien wie der

---

<sup>3</sup> Die Ereignisse wurden den Probanden wie bei einzelfallweiser Darbietung nacheinander dargeboten. Zusätzlich wurden die Ereignisse auf einem Zeitstrahl angeordnet, so dass sie für die Probanden nach der Darbietung einsehbar blieben und das Gedächtnis entlastet wurde.



$\Delta P$ - oder der  $\Delta D$ -Regel Gebrauch machen, beantworten Shaklee & Wasserman (1986) damit, dass in ihrer Studie die tabellarische Präsentationsform und die Fragestellung von anderen Studien abweichen. So wurden beispielsweise in anderen Studien (z.B. Shaklee & Hall, 1983; Shaklee & Tucker, 1980) häufig tabellarische 2 x 2 Kontingenztafeln eingesetzt, während Shaklee & Wasserman (1986) die vier Zellinformationen vertikal von Zelle a nach Zelle d anordneten. Da in einer tabellarischen 2 x 2 Kontingenztafel alle Zellen miteinander benachbart sind, werden Vergleiche aller vier Zellen untereinander eher angeregt als in einer vertikalen Tabelle mit den Zellen a und b in den ersten beiden Reihen. Im Zusammenhang mit der Fragestellung heben Shaklee & Wasserman (1986) hervor, dass ihre Probanden Kausalurteile bilden sollten (der Stimulus verursacht, verhindert oder hat keinen Einfluss auf das Ergebnis), während die Probanden in den Studien von Shaklee & Hall (1983), so wie Shaklee & Tucker (1980) ein Urteil über Wahrscheinlichkeitsunterschiede abgeben sollten (Ergebnis ist wahrscheinlicher, wenn sich der Stimulus ereignet, nicht ereignet, oder es gibt keinen Unterschied).

#### *Diskussion der Befunde zum Einfluss des Faktors Präsentationsform*

Auch wenn die bisherige Befundlage im Bereich der Kausalkognition als äußerst dürftig angesehen werden muss, weisen die Ergebnisse von Shaklee & Wasserman (1986) und Wasserman & Shaklee (1984) darauf hin, dass die Präsentationsform einen Einfluss auf die kausale Urteilsbildung hat. Allerdings sollte dabei auch der Einfluss von und die Konfundierung mit anderen aufgaben- und medienbezogenen Faktoren wie zum Beispiel der Darbietungsart oder der Fragestellung berücksichtigt werden.

Völlig offen im Bereich der Kausalkognition ist die Suche nach einer Theorie, die den nahe liegenden Einfluss der Präsentationsform, sowie der Darbietungsart und der Informationsart erklären helfen könnte. Der Blick auf andere Areale kognitionspsychologischer Forschung, in denen der Einfluss der Präsentationsform mehrfach belegt wurde, bietet hier Möglichkeiten, fündig zu werden. So haben beispielsweise Kleinmuntz & Schkade (1993) eine Theorie entwickelt, in der davon ausgegangen wird, dass Informationsanordnungen (englisch: *information displays*) einen Einfluss auf Prozesse bei der Entscheidungsbildung haben. Diese Theorie wird in Kapitel 6.1 vorgestellt. Dabei wird auch ihre Anwendbarkeit auf den Bereich der kausalen Urteilsbildung diskutiert.



## **5 Individuelle Unterschiede und Gedächtnisbeanspruchung bei der kausalen Urteilsbildung**

### **5.1 Einfluss individueller Unterschiede**

#### *Empirische Befunde*

Nach unserer Auffassung werden die Prozesse bei der Verarbeitung von Informationen über Ursache-Wirkungs-Zusammenhänge nicht nur von situativen, medialen (z.B. Präsentationsform, Informationsart, Darbietungsart) und aufgabenbezogenen (z.B. Fragestellung, Instruktion) Faktoren beeinflusst, sondern auch von individuellen Unterschieden. Eine Reihe von Studien belegt den Einfluss individueller Unterschiede auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen (Alloy & Abramson, 1979; Alloy & Tabachnik, 1984; Inhelder & Piaget, 1958; Shaklee & Elek, 1988; Shaklee & Hall, 1983; Shaklee & Mims, 1981; Shaklee & Tucker, 1980; Wasserman, Chatlosh & Neunaber, 1983; Wasserman & Shaklee, 1984; White, 1998, 2000a). So fanden beispielsweise Alloy & Abramson (1979) heraus, dass depressive Probanden Kontingenzen genauer einschätzen als nicht depressive Probanden.

Mehrere Studien belegen, dass Männer häufiger und genauer regel-basierte Strategien einsetzen als Frauen. (Shaklee & Elek, 1988; Shaklee & Hall, 1983; Wasserman, Chatlosh & Neunaber, 1983; Wasserman & Shaklee, 1984). Basierend auf Befunden von Maccoby & Jacklin (1974) nehmen Shaklee & Hall (1983) an, dass Männer Frauen im mathematischen Denken überlegen sind. Shaklee & Elek (1988) widersprechen dieser Interpretation aber, weil sie herausfanden, dass Männer nur dann komplexere Strategien einsetzten, wenn sie Kovariationsurteile bilden sollten, nicht aber bei der Bildung von Kausalurteilen. Die fehlenden Geschlechterunterschiede bei der Bildung von Kausalurteilen zeigen nach Shaklee & Elek deutlich, dass Frauen ebenso komplexe mathematische Berechnungen anstellen können wie Männer. Bei der Bildung von Kovariationsurteilen hingegen vermuten die beiden Autorinnen, dass Frauen hierbei im Gegensatz zu Männern dem Anfertigen mathematischer Berechnungen weniger Bedeutung beimessen.

Andere Studien (Inhelder & Piaget, 1958; Shaklee & Elek, 1988; Shaklee & Mims, 1981; Shaklee & Tucker, 1980) haben herausgefunden, dass die Komplexität der Strategien zur Beurteilung von Ursache-Wirkungs-Zusammenhängen mit wachsendem Lebensalter kontinuierlich zunimmt. So machten in der Untersuchung von Shaklee & Mims (1981, Ex 1)

38 % aller studentischen Probanden, 27 % aller Probanden aus Klasse 10, 4 % aus Klasse 7 und 0 % aus Klasse 4 von der  $\Delta P$ -Regel Gebrauch, während unklare oder nicht klassifizierbare Strategien bei 46 % der Viertklässler und nur bei 3 % der Studenten auftraten. Diese Ergebnisse legen nahe, dass die Komplexität der eingesetzten Strategien entwicklungsabhängig ist und dass der Gebrauch einer derart komplexen Strategie wie der  $\Delta P$ -Regel nicht angeboren ist, sondern erworben ist.

Neuere Untersuchungen (Anderson & Sheu, 1995; Lober & Shanks, 2000; Perales & Shanks, 2003; Wasserman, 1990) betonen ebenfalls, dass die Probanden keine einheitliche Strategie anwenden, sondern dass es Subgruppen gibt, deren Teilnehmer jeweils auf unterschiedliche Strategien wie die  $\Delta P$ -Regel oder die Berechnung der kausalen Stärke nach der Power PC-Theorie zurückgegriffen haben. Allerdings werden im Gegensatz zu älteren Studien keine Angaben über mögliche Gründe für die gefundenen Unterschiede gemacht. Lober & Shanks (2000) räumen passend dazu ein, dass der Existenz unterschiedlicher Subgruppen theoretisch viel zu wenig Gewicht eingeräumt wurde. Perales & Shanks (2003) bemängeln weiter, dass sich die jüngere Forschung fast ausschließlich auf die Identifizierung eines einzelnen kognitiven Mechanismus konzentriert hat, obwohl singuläre, allumfassende Erklärungsmechanismen nicht in der Lage sind, dem Reichtum und der Vielfalt menschlichen Kausallernens gerecht zu werden. Genauere Analysen zu Experiment 3 bei Perales & Shanks (2003) verdeutlichen, dass die in der Forschung üblich gewordenen Mittelwertvergleiche die Breite individueller Strategien nicht angemessen repräsentieren können.

Die bisher vorgestellten Studien zum Einfluss individueller Unterschiede haben gezeigt, dass Faktoren wie das Vorhandensein einer depressiven Symptomatik, das Geschlecht und der Entwicklungsstand bei der Beurteilung der Stärke kausaler Zusammenhänge eine Rolle spielen können. Zudem wurde deutlich, dass verschiedene Individuen in ein- und derselben Untersuchung jeweils von unterschiedlichen Strategien Gebrauch machen.

Von einer anderen Fragestellung geht White (1998, 2000a) aus. Er interessiert sich weniger dafür, ob unterschiedliche Individuen in einer Studie von einer oder mehreren Strategien angemessen oder unangemessen Gebrauch machen, sondern er versucht, Abweichungen von sowohl regel-basierten als auch assoziativen Ansätzen mit dem idiosynkratischen Gebrauch unterschiedlicher Arten von Zellinformationen als konfirmatorisch und diskonfirmatorisch zu erklären. So werden insbesondere die Zellhäufigkeiten c und d aus einer Vierfeldertafel (s. Abbildung 1) von einigen Individuen als positive Evidenz für die Wirksamkeit einer Ursache angesehen, von anderen als negative (White, 1998; s.a. Arkes & Harkness, 1983, Ex 7). Weder assoziative Modelle, noch die Power PC-Theorie oder das PCM, können das

Vorhandensein dieser individuellen Tendenzen bei der Zellgewichtung erklären. Für White (2000a) ist der Einsatz idiosynkratischer Strategien wohlüberlegt und erfolgt auf einer höheren kognitiven Ebene der Informationsverarbeitung. Die Vorstellungen, die Individuen über die Wertigkeit einzelner Zellinformationen haben, sind im Laufe der Lerngeschichte erworben worden.

Der Einsatz idiosynkratischer Strategien scheint mit den Faktoren Darbietungsart und Präsentationsform zu interagieren. Werden den Probanden die Informationen über Ursache-Wirkungs-Zusammenhänge beispielsweise in einer 2 x 2-Vierfeldertafel summarisch dargeboten, dann setzt nur eine Minderheit idiosynkratische Strategien ein. Wenn sich die Komplexität der Datenpräsentation erhöht, dann bilden dagegen idiosynkratische Tendenzen die Mehrheit (White, 1998).

Der Ansatz von White (1998, 2000a) im Besonderen und von der Relevanz individueller Unterschiede im Allgemeinen ist konsistent mit Befunden, welche die hohe Bedeutsamkeit des Vorwissens bei der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen hervorheben (Alloy & Tabachnik, 1984; Koslowski, 1996; Young, 1995).

### *Diskussion*

Zusammenfassend ist festzustellen, dass die Befundlage zum Einfluss individueller Unterschiede auf die Beurteilung der Stärke von Kausalzusammenhängen eher spärlich und wenig differenziert ist. Der immer wieder gefundene Einfluss individueller Unterschiede steht im Widerspruch zur geringen Forschungsaktivität auf diesem Gebiet (Perales & Shanks, 2003), so dass es unklar bleibt, worin sich Probanden, die regel-basiert vorgehen, von solchen unterscheiden, die nicht regel-basiert vorgehen. Die bislang durchgeführten Studien heben mit dem Alter oder dem Geschlecht Merkmale von geringer theoretischer Aussagekraft hervor. Warum ältere Probanden oder Männer häufiger regel-basierte Strategien anwenden, wird nicht begründet. Die Abhängigkeit der Strategiewahl vom Entwicklungsstand (Shaklee & Mims, 1981; White, 1995) kann als ein Hinweis darauf verstanden werden, dass es sich bei komplexen, regel-basierten Strategien um keine angeborenen, sondern um erworbene kognitive Schemata oder Skripts handelt. Der Einfluss des Geschlechts wird zu pauschal festgestellt, es wird lediglich darüber spekuliert, dass die gefundenen Unterschiede auf unterschiedlichen mathematischen Fertigkeiten beruhen (Shaklee & Hall, 1983). Ob individuelle Unterschiede im Hinblick auf die mathematischen Fertigkeiten oder die numerische Verarbeitungskapazität bei der Beurteilung der Stärke von Kausalzusammenhängen tatsächlich eine Rolle spielen, wurde bislang noch nicht untersucht.

Die mangelnde Berücksichtigung individueller Unterschiede ist grundsätzlich zu kritisieren. Die bislang bekannt gewordenen Ergebnisse machen deutlich, dass sich Individuen selbst in ein- und derselben Untersuchungsbedingung im Hinblick auf ihre Verarbeitungsmechanismen oder ihre Strategiewahl bei der Beurteilung der Stärke von kausalen Zusammenhängen unterscheiden (Anderson & Sheu, 1995; Inhelder & Piaget, 1958; Kao & Wasserman, 1993; Lober & Shanks, 2000; Perales & Shanks, 2003; Shaklee & Elek, 1988; Shaklee & Hall, 1983; Shaklee & Mims, 1981; Shaklee & Tucker, 1980; Ward & Jenkins, 1965; Wasserman, 1990; Wasserman, Chatlosh & Neunaber, 1983; Wasserman & Shaklee, 1984; White, 1995, 1998, 2000a). Dies lässt die weit verbreitete Praxis, Mittelwertvergleiche durchzuführen und individuelle Unterschiede außer Acht zu lassen, fragwürdig erscheinen. Für angemessener halten wir es, wie in älteren Studien (z.B. Inhelder & Piaget, 1958; Shaklee & Elek, 1988; Shaklee & Hall, 1983; Shaklee & Mims, 1981; Shaklee & Tucker, 1980; Ward & Jenkins, 1965) individuell eingesetzte Strategien zu identifizieren oder die Strategiewahl als abhängige Variable mit in Betracht zu ziehen, wenn es um die Frage nach den zugrunde liegenden Prozessen bei der Beurteilung der Stärke von kausalen Zusammenhängen geht. Wir gehen davon aus, dass – anders als bei assoziativen Mechanismen - der Anwendung regelbasierter Strategien Prozesse zu Grunde liegen, die einer höheren kognitiven Ebene entspringen.

## **5.2 Bedeutung von Gedächtnisbeanspruchung und Verarbeitungsaufwand**

Das Gedächtnis ist an einer Vielzahl kognitiver Prozesse beteiligt und spielt insbesondere bei höheren kognitiven Funktionen wie z.B. Lernen, Urteilsbildung, Schlussfolgern, Kopfrechnen und Verstehen eine entscheidende Rolle (Baddeley & Hitch, 1974; Noel, Désert, Aubrun, & Seron, 2001; Oberauer, Demmrich, Mayr & Kliegl, 2001). Da es sich bei der Beurteilung der Stärke von kausalen Zusammenhängen um eine höhere kognitive Funktion handelt<sup>4</sup>, muss davon ausgegangen werden, dass Gedächtnisprozesse auch dabei bedeutsam sind. So sind beispielsweise kausale Schemata und Skripts, Kausalmodelle und spezifisches Wissen über kausale Mechanismen und Ereignisse im Langzeitgedächtnis abgespeichert und werden bei der Verarbeitung von Informationen über Ursache-Wirkungs-Zusammenhänge zugänglich

---

<sup>4</sup> Diese Annahme dürfte nicht unumstritten sein, insbesondere Vertreter rein assoziationalistischer Theorien bezweifeln die Beteiligung von kognitiven Prozessen auf höherer Ebene. Befunde zur Theorie der Kausalmodelle, zu mechanistischen Ansätzen und neuerdings auch zur Überprüfung von assoziationalistischen Lerntheorien (De Houwer, Beckers & Vandorpe, 2005) zeigen aber, dass die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen von Vorwissen, Kausalmodellen und kausalen Schemata abhängig ist.

gemacht. Ob diese Informationen aber auch tatsächlich abgerufen werden können, hängt von der verfügbaren Kapazität des Arbeitsgedächtnisses ab, die je nach Komplexität der Aufgabenstellung und individueller Verarbeitungskapazität variieren kann.

Aktuelle Theorien konzipieren das Arbeitsgedächtnis als ein System, das dafür zuständig ist, die bei der Ausführung einer Aufgabe notwendigen Informationen in einem aktiven Zustand verfügbar zu halten. Darüber hinaus ist das Arbeitsgedächtnis zuständig für die exekutive Kontrolle von Aufmerksamkeitsressourcen bei der Verfolgung von Aufgabenzielen, für die Hemmung oder Abhaltung störender Information und für die Steuerung und Koordination von unterschiedlichen Teilprozessen, die zur Erfüllung einer Aufgabe erforderlich sind (Baddeley, 1986; Miyake & Shah, 1999). Nach Baddeley (1986, 1992, 1996, 1999) funktioniert das Arbeitsgedächtnis als ein System, welches aus mehreren Komponenten besteht, die wiederum von einer zentralen Exekutive kontrolliert und reguliert werden. Unter anderem ist die zentrale Exekutive für die Verteilung begrenzter Ressourcen verantwortlich, die für die Verarbeitung und/oder die Speicherung von Information benötigt werden. Eine Version des Ressourcen-Verteilungs-Modells (Daneman & Carpenter, 1980; Just & Carpenter, 1992; Kahneman, 1973) geht davon aus, dass bei wachsender Gedächtnisbeanspruchung die Gedächtnisleistung sinkt, weil für die Verarbeitung oder Speicherung von Information durch das Arbeitsgedächtnis weniger Kapazität verfügbar ist.

Aus der Gedächtnisforschung ist es schon seit längerem bekannt, dass die Beanspruchung des Arbeitsgedächtnisses bei der Verarbeitung und Abspeicherung von Information unter anderem von Faktoren der Darbietung und der Präsentation der Daten abhängig ist (z.B. Baddeley & Logie, 1999; Goolkasian & Foos, 2002; Smith & Magee, 1980).

Bei der Diskussion zum Einfluss des Faktors Darbietungsart auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen (Kapitel 4.2.1) waren wir ebenfalls zu dem Ergebnis gekommen, dass Variationen der Darbietungsart und wahrscheinlich auch anderer aufgaben- und medienbezogener Faktoren eine unterschiedliche Beanspruchung des Gedächtnisses nach sich ziehen. Aus theoretischer Sicht kann nun angenommen werden, dass durch eine einzelfallweise Darbietung die Gedächtnisbeanspruchung erhöht wird, weil dabei im Gegensatz zur summarischen Darbietung Ressourcen für die Speicherung von Information verloren gehen. Bei summarischer Darbietung wird das Gedächtnis entlastet, weil die vorliegenden Informationen nicht erst noch gespeichert werden müssen, sondern gleich verarbeitet werden können. Da ausreichend Ressourcen zur Verfügung stehen, dürfte die Verarbeitungsgenauigkeit bei summarischer Darbietung weniger beeinträchtigt sein als bei einer einzelfallweisen Darbietung.

*Empirische Befunde*

Empirische Studien, in denen direkt der Einfluss von Verarbeitungsaufwand und Gedächtnisbeanspruchung auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen untersucht wurde, liegen nur sehr vereinzelt vor. Die beiden uns bekannten Studien von Shaklee & Mims (1982), sowie von Waldmann & Haggmayer (2001) unterscheiden sich zudem beträchtlich darin, wie der Einfluss der Faktoren Verarbeitungsaufwand und Gedächtnisbeanspruchung jeweils operationalisiert und untersucht wurde.

Waldmann & Haggmayer (2001) haben den Einfluss des Verarbeitungsaufwandes auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen untersucht, indem sie über drei Experimente hinweg Darbietungsart und Strukturierung des Lerninputs manipulierten (geordnete und ungeordnete Falllisten vs. geordnete und ungeordnete Einzelfalldarbietung). Sie kamen dabei zu dem Ergebnis, dass der Einsatz von Lernstrategien und die Sensitivität für Kausalmodelle bei der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen von Anforderungen an die kognitive Verarbeitungskapazität abhängig sind und dass viele Möglichkeiten existieren, wie statistische Informationen verarbeitet werden können. Für Waldmann & Haggmayer (2001) sind der Einsatz von Lernstrategien und die Beurteilung der Stärke von kausalen Zusammenhängen das Produkt einer komplexen Interaktion. An dieser Wechselwirkung sind im Einzelnen das Vorwissen und die Ziele des Lernenden (Anstrengungsvermeidung vs. korrekte Urteilsbildung), die Struktur des Lerninhalts (geordnet vs. ungeordnet), die Art der Lernaufgabe (z.B. einzelfallweise vs. summarische Darbietung) und Schätzungen eines Trade-Offs zwischen Strategien und Lernaufwand beteiligt.

Im Hinblick auf die Ziele des Lernenden und dem vermuteten Trade-Off zwischen Strategieeinsatz und Lernaufwand gehen Waldmann & Haggmayer (2001) weiter davon aus, dass Menschen darum bemüht sind, die Lernaufgabe kognitiv handhabbar zu machen, d.h. dass sie danach streben, den Verarbeitungsaufwand so gering wie nötig zu halten. Eine interessante Fragestellung für die Zukunft sehen sie darin, herauszufinden, ob Menschen in der Lage sind, den Aufwand zu regulieren, den sie in die Analyse statistischer Beziehungen investieren. Waldmann & Haggmayer (2001) erwähnen in diesem Zusammenhang ebenfalls Forschungsergebnisse aus dem Bereich der Entscheidungsbildung, in denen gezeigt werden konnte, dass Probanden häufig die Komplexität der Informationsverarbeitung in Abhängigkeit von den Kosten-Nutzen-Einschätzungen für eine korrekte Urteilsbildung adaptiv verändern (Payne, Bettman, & Johnson, 1993).



Shaklee & Mims (1982) haben mit einer regel-analytischen Technik untersucht, welche Auswirkungen die Gedächtnisbeanspruchung auf die Beurteilung von Ursache-Wirkungs-Zusammenhängen haben. Der Einfluss der Gedächtnisbeanspruchung wurde dabei durch eine Variation der Darbietungsart untersucht (einzelfallweise vs. summarisch). Nach Shaklee & Mims (1982) kann eine hohe Gedächtnisbeanspruchung infolge einzelfallweiser Informationsdarbietung zwei Auswirkungen nach sich ziehen: (i) Zum einen kann es zu fehlerhaften und/oder verzerrten Kausalurteile kommen, d.h. dazu, dass ein- und derselbe Mechanismus unterschiedlich genau oder fehlerbehaftet ausgeführt wird. (ii) Zum anderen kann eine hohe Gedächtnisbeanspruchung infolge einzelfallweiser Informationsdarbietung dazu führen, dass die Probanden bei zunehmender Gedächtnisbeanspruchung verstärkt auf vereinfachende Heuristiken wie z.B. die ab-Regel zurückgreifen (s. z.B. die Ergebnisse von Arkes & Harkness, 1983, in den Experimenten 4 und 5). Dabei führen die Probanden keine komplexe Strategie fehlerhaft aus, sondern sie greifen gezielt auf vereinfachende Heuristiken zurück, indem sie nur einen Teil (beispielsweise nur die Zelhäufigkeiten a und b) der dargebotenen Informationen abspeichern und verarbeiten.

*Zu (i): Wie entstehen verzerrte oder fehlerhafte Kausalurteile*

Verzerrte oder fehlerhafte Urteile können auf einer fehlerhaften Selektion oder einer ungleichen Gewichtung von Information beruhen. Nach Shaklee & Mims (1982) gehen verzerrte oder ungenaue Kausalurteile auf einen erschwerten Gedächtnisabruf von Häufigkeiten entsprechend den vier Zellen a, b, c und d einer Vierfeldertafel zurück. Die erhöhten Gedächtnisanforderungen führen also zu Fehlern bei der Bestimmung oder Einschätzung der vier Zelhäufigkeiten. Kao & Wasserman (1993) vermuten im Zusammenhang mit der Beurteilung der Stärke von kausalen Zusammenhängen, dass eine begrenzte Gedächtnis- und Verarbeitungskapazität unter Einzelfalldarbietung dazu führt, dass die Probanden bei der Anhäufung der verschiedenen Zellinformationen selektiver vorgehen und sich mehr auf die Zellgrößen a und b konzentrieren, während die Zelhäufigkeiten c und d eher vernachlässigt werden.

Eine im Vergleich zu Shaklee & Mims (1982) umfassendere Theorie zur Erklärung von Urteilsverzerrungen und Fehlern wurde von Fiedler (2000) im Rahmen seines kognitiv-ökologischen Ansatzes vorgestellt. Er macht nicht Prozesse auf der Ebene der Urteilsbildung selbst für das Auftreten von Urteilsverzerrungen und -fehlern verantwortlich, sondern vermutet, dass diese auch durch fehlerhafte oder vereinfachende Strategien auf der grundlegenden Ebene der Informationssuche bedingt sein können. Urteilsfehler und –

verzerrungen entstehen also schon durch Prozesse bei der Datensammlung und der Auswahl der Stichprobe, welche zur späteren Urteilsbildung herangezogen wird.

Nach Fiedler (2000) entstehen Urteilsfehler und –verzerrungen durch eine dynamische Interaktion zwischen umweltbezogenen und personenbezogenen Faktoren. Die Prozesse auf der Ebene der Informationssuche unterliegen also nicht allein der kognitiven Kontrolle eines Individuums. Stattdessen hängen Datensammlung und Stichprobenauswahl im hohen Maß von der Perspektive des Betrachters, sowie von der Verteilung, Verfügbarkeit, Nähe und Salienz von Informationen in der Umgebung ab. Kognitive Prozesse (wie z.B. selektive Aufmerksamkeitsfokussierung und Informationsaufnahme) und Voraussetzungen (wie z.B. Verarbeitungskapazität und -ziele) legen zwar den Wahrnehmungsraum eines Individuums und seine Orientierung in der Umwelt fest, doch werden sie auch von den Beschränkungen und Gegebenheiten der Umwelt geprägt.

*Zu (ii): Warum wechseln Probanden zu einfacheren und ungenaueren Strategien?*

Der Wechsel zu einfacheren und ungenaueren Strategien bei Einzelfalldarbietung wird von Shaklee & Mims (1982) ebenfalls auf eine zu hohe Gedächtnisbeanspruchung als Folge einer zu hohen Aufgabenkomplexität zurückgeführt (s.a. Kao & Wasserman, 1993). Durch die hohe Gedächtnisbeanspruchung erhöht sich also die Tendenz, einfache, aber ungültige Strategien an Stelle von komplexen Regeln bei der Beurteilung des Kausalzusammenhangs einzusetzen. Diese Erklärung kann durch die Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993, s.a. Kapitel 6.1) vervollständigt werden, nach welcher davon ausgegangen werden kann, dass die Probanden bei einzelfallweiser im Vergleich zu summarischer Darbietung einen höheren Aufwand und eine geringere Genauigkeit im Zusammenhang mit dem Einsatz einer komplexen regel-basierten Strategie antizipieren und deshalb zu einfacheren, aber ungültigen Strategien überwechseln.

### *Diskussion*

Zusammenfassend ergibt sich das Bild, dass eine zu hohe Gedächtnisbeanspruchung zwar ein wesentlicher, aber nicht der einzige Faktor zu sein scheint, welcher bei der Beantwortung der beiden Fragen, warum es bei einzelfallweiser Informationsdarbietung dazu kommt, dass (i) fehlerhafte und/oder verzerrte Kausalurteile gebildet werden und (ii) dazu, dass die Probanden einfachere und ungenauere Strategien einsetzen. Innerhalb der Forschung zur Kausalkognition bietet die kognitiv-ökologische Theorie von Fiedler (2000) einen umfassenderen Rahmen zur Erklärung des Auftretens von Urteilsverzerrungen und –fehlern. Aus dem Bereich der Entscheidungstheorie können Kosten-Nutzen-Theorien (z.B. Kleinmuntz & Schkade, 1993, s.

Kapitel 6.1) zur Erklärung des Auftretens von Strategiewechseln hinzugezogen werden. Beiden letztgenannten Theorieansätzen ist gemeinsam, dass sie von einer dynamischen Interaktion zwischen umweltbezogenen (z.B. Informationsanordnung) und kognitiven Faktoren ausgehen. Während Fiedler die basale Ebene der Informationssuche als zentrale Schnittstelle dieser Wechselwirkung begreift, gehen Kleinmuntz & Schkade (1993) davon aus, dass unterschiedliche Informationsanordnungen unterschiedliche motivationale Anreize bei der Strategiewahl schaffen.

In Verbindung mit dem in Kapitel 5.1 diskutierten Einfluss individueller Unterschiede ist festzustellen, dass der Einfluss von Gedächtnisbeanspruchung und Verarbeitungsaufwand auf die Beurteilung der Stärke kausaler Zusammenhänge bislang nur von einer aufgaben- und medienbezogenen Seite aus untersucht wurde. Obwohl davon ausgegangen werden muss, dass auch die Gedächtnisbeanspruchung aus einer Wechselwirkung zwischen aufgaben- und medienbezogenen Faktoren und einer jeweils individuell unterschiedlich ausgeprägten Verarbeitungskapazität resultiert, wurde der Einfluss der individuellen Verarbeitungskapazität auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen bislang völlig ignoriert. Befunde aus der Gedächtnisforschung legen nahe, dass individuelle Fertigkeiten (wie z.B. im numerischen Denken) und Vorerfahrungen aus dem Langzeitgedächtnis herangezogen werden können, um die begrenzte Kapazität des Arbeitsgedächtnisses zu überwinden (Ericsson & Dulaney, 1999; Miyake & Shah, 1999).

In jüngerer Zeit finden sowohl die Bedeutung des Arbeitsgedächtnisses als auch die Rolle individueller Voraussetzungen zunehmende Beachtung. Neben den Befunden von Waldmann & Hagmayer (2001) und Perales & Shanks (2003) berichten auch De Houwer & Beckers (2003) von einer Untersuchung im Zusammenhang mit der Überprüfung assoziationalistischer Lerntheorien, in der die Verfügbarkeit von Ressourcen aus dem Arbeitsgedächtnis durch die Schwierigkeit einer zusätzlichen Aufgabe manipuliert wurde. Dabei stellte heraus, dass das Auftreten von Blockierungseffekten entscheidend von der Verfügbarkeit der Ressourcen aus dem Arbeitsgedächtnis abhängt. In Verbindung mit dem Einfluss individueller Voraussetzungen kamen Vandorpe, De Houwer & Beckers (2005) zu dem Ergebnis, dass die Auswirkungen konkurrierender Ursachen beim kompetitiven Lernen nur bei solchen Probanden zu beobachten war, die im Anschluss an die Untersuchung angemessen über schlussfolgernde Denkprozesse berichten konnten.



## 6 Theorien und Empirie zum Einfluss von Medien auf die menschliche Informationsverarbeitung

Die frühe Medienpsychologie wurde von der Suche nach Medieneffekten dominiert (Clark, 1983; Mayer, 1997; Salomon, 1979, 1994). Dabei ging es vornehmlich um die Frage, ob ein Medium effektiver als ein anderes ist. Die Vernachlässigung von Aspekten der kognitiven Informationsverarbeitung wurde vor allem durch das Buch *Interaction of Media, Cognition and Learning* von Gavriel Salomon (1979) zu Bewusstsein gebracht und in Frage gestellt. Salomon (1979) betont darüber hinaus, dass die Interaktion zwischen einzelnen Medien und kognitiver Informationsverarbeitung maßgeblich durch die auszuführende Aufgabe festgelegt wird. Unterschiedliche Medien haben dabei unterschiedliche Auswirkungen auf die Informationsverarbeitung und setzen unterschiedliche kognitive Prozesse, Module oder Mechanismen in Gang. Die instruktionale Nützlichkeit eines bestimmten Mediums hängt nun jeweils davon ab, ob die Verarbeitungsweisen, die durch dieses Medium aktiviert werden, aufgabenrelevant sind oder nicht.

Auch Mayer (1997) hält die Fragestellung, ob ein Medium effektiver ist als ein anderes, für wenig sinnvoll, aus seiner Sicht ist zumindest eine Relativierung dieser Fragestellung auf unterschiedliche Umstände, unterschiedliche Aufgabenstellungen und unterschiedliche Lerner notwendig. Innerhalb der Instruktions- und Medienpsychologie sprechen sich deshalb eine Reihe von Forschern (Clark, 1994; Jonassen, Campbell & Davidson, 1994; Kozma, 1994; Mayer, 1997; Zhang, Johnson, Malin & Smith, 2002) dafür aus, dass eher lerner- als medienzentriert vorgegangen werden sollte, indem Fragen nach den kognitiven Auswirkungen einer Instruktionmethode oder einer Präsentationsform bei der Wissenskonstruktion durch den Lernenden in den Vordergrund gerückt werden.

Während die frühe Medienpsychologie also die zentralen Aspekte der Informationsverarbeitung und der Wissenskonstruktion durch den Lernenden außer Acht ließ und als *medienzentriert* charakterisiert werden kann, so gilt für die bisherige Forschung im Bereich der Kausalkognition, dass sie bei der Theoriebildung neben individuellen Unterschieden Aspekte der Informationsdarbietung und –präsentation fast völlig ignorierte und sich insofern eher als *kognitionszentriert* bezeichnen ließe. Ähnlich wie in der Medienpsychologie versucht wurde, herauszufinden, welches Medium am besten ist (Salomon, 1979; Mayer, 1997), so wird im Bereich der Kausalkognition immer noch versucht, verschiedene Theorien gegeneinander auszuspielen und nach einer

allgemeingültigen Theorie zu suchen (dieser Versuch kann als gescheitert angesehen werden). Die Frage, welche Theorie oder welcher Ansatz die meisten Befunde vorhersagt, muss auch für den Bereich der Kausalkognition bezüglich des situativen Kontextes, der individuellen Voraussetzungen, der aufgabenbezogenen Faktoren, sowie der Informationsdarbietung und -präsentation relativiert werden.

Auch Oestermeier & Hesse (2000) bemängeln an der bisherigen Forschung im Bereich der Kausalkognition, dass kausales Denken stillschweigend als etwas angesehen wird, was sich ausschließlich im Kopf des Urteilenden abspielt, obwohl in fast allen Experimenten, in denen es um die Beurteilung kausaler Zusammenhänge geht, unterschiedliche Medien oder Präsentationsformen eingesetzt wurden. Nach ihrer Auffassung kann der Prozess der kausalen Urteilsbildung ebenso wie andere informationsverarbeitende Prozesse (z.B. Problemlösen, Entscheidungsbildung) besser als eine Interaktion von einem Individuum (oder mehreren Individuen) mit seiner Informationsumgebung verstanden werden kann (Clark, 1997; Fiedler, 2000; Hutchins, 1995a, b; Zhang, 1997, 2000; Zhang & Norman, 1994). Medien sind demnach mehr als nur eine periphere Stütze bei kognitiven Prozessen; sie beeinflussen die Salienz einzelner Informationsteile, grenzen ein, legen Vergleichsprozesse nahe oder regen Schlussfolgerungen an, ohne dass dies jeweils verbal explizit gemacht werden muss. Mit Ausnahme der Arbeit von Oestermeier & Hesse (2000) ist eine solche Sichtweise innerhalb der Forschung zu den Informationsverarbeitungsprozessen bei der Beurteilung der Stärke von kausalen Zusammenhängen lange vernachlässigt worden.

Oestermeier & Hesse (2000) haben für die ganze Bandbreite von verbalen kausalen Argumenten dargelegt und diskutiert, wie visuelle Medien die Darstellung und die Ausdruckskraft von verbalen kausalen Argumenten einengen können und wie sie argumentenspezifische Schlussfolgerungen (wie z.B. den Vergleich zwischen Erfolgs- und Basisrate bei der Beurteilung der Stärke von kausalen Zusammenhängen) auslösen, unterstützen und externalisieren können.

In Verbindung mit einer regel-basierten Beurteilung der Stärke von kausalen Zusammenhängen sind von den insgesamt 27 bei Oestermeier & Hesse (2000) identifizierten Arten von verbalen kausalen Argumenten vor allem drei Formen bedeutsam. Die erste Form begründet die Verursachung eines Effekts E durch die Ursache C mit dem gemeinsamen Auftreten von C und E. Aus einer Verallgemeinerung der konstanten Verknüpfung von C und E kann dann in der Tradition Humes ein kausaler Zusammenhang geschlossen werden. Bei der Beurteilung der Stärke von Kausalzusammenhängen entspricht dies einfachen Strategien wie der "Zelle a-Strategie" oder der "a versus b Strategie" (s. Kapitel 3.1.1, S. 25 ff. Die

*Suche nach der Regel*). Die zweite Form begründet den Kausaleinfluss von C auf E damit, dass sich das Erscheinen von E mit dem Erscheinen von C verändert. Um zu dieser Schlussfolgerung zu gelangen, muss der Urteilende nicht nur berücksichtigen, was mit E geschieht, wenn C gegeben ist, sondern auch, was mit E geschieht, wenn C nicht gegeben ist. Regel-basiert entspricht dies einem Vergleich der Anzahl der Fälle, in denen C und E gemeinsam auftreten, mit der Anzahl der Fälle, in denen E auftritt, ohne dass C gegeben ist. Die dritte Form ist die probabilistische Variante der zweiten Form, hier wird der Kausaleinfluss von C auf E damit begründet, dass sich durch C die Wahrscheinlichkeit des Auftretens von E erhöht. In Verbindung mit einer regel-basierten Beurteilung der Stärke von kausalen Zusammenhängen entspricht dies einem Vergleich von Erfolgs- und Basisrate bei der Berechnung des Kontingenzmaßes  $\Delta P$ . Nicht berücksichtigt werden von Oestermeier & Hesse (2000) kausale Argumente, die auf der Power PC Theorie von Cheng (1997) aufbauen. Die Hauptaussage von Oestermeier & Hesse (2000), dass visuelle Medien zum einen Daten sichtbar machen (externalisieren) und zum anderen argumentspezifische Schlussmuster auslösen und unterstützen, lässt sich auch für die drei Formen von kausalen Argumenten spezifizieren, die bei einer regel-basierten Beurteilung der Stärke von kausalen Zusammenhängen bedeutsam sind. Hinsichtlich der Schlussfolgerungen, die durch visuelle Medien nahe gelegt oder externalisiert werden können, unterscheiden Oestermeier & Hesse (2000) zwischen assoziationalistischen und perzeptuellen Inferenzen einerseits und reflektierten, bewusstseinsnahen Inferenzen andererseits. Bei Inferenzen, die aus der Anwendung regel-basierter Strategien resultieren, handelt es sich um reflektierte, bewusstseinsnahe Inferenzen. Im Gegensatz zu den automatisiert ablaufenden, assoziationalistischen Interpretationen über die Stärke von kausalen Zusammenhängen, ist bei einer regel-basierten Beurteilung der Stärke von Kausalzusammenhängen eine zusätzliche Unterstützung durch Externalisierungen und visuelle Strukturierungen des Informationsmaterials (wie z.B. in Tabellen oder Diagrammen) erforderlich, damit die dafür notwendigen komplexen mentalen Operationen und Vergleichsprozesse ausgeführt werden können. Individuen, die ohne eine solche Unterstützung auskommen können, müssen neben den kausalen Schemata oder Skripts, die bei einer regel-basierten Vorgehensweise angewendet werden, über eine hohe Verarbeitungskapazität verfügen.

Die fehlende Berücksichtigung einer dynamischen Interaktion zwischen kognitiver Informationsverarbeitung und Informationsumgebung war nicht überall so ausgeprägt wie im Bereich der Kausalkognition. Andere Forschungsareale wie z.B. Wahrnehmung (Gibson, 1979); Problemlösen (Larkin & Simon, 1987; Zhang, 1997, 2000; Zhang & Norman, 1994),

Entscheidungsbildung (Kleinmuntz & Schkade, 1993; Sanfey & Hastie, 1998; Vessey, 1991, 1994), logisches Denken (Stenning & Oberlander, 1994, 1995) und insbesondere die angewandte Kognitionswissenschaft (Hutchins, 1995a, b; Zhang, Johnson, Malin & Smith, 2002) haben diesen Paradigmenwechsel schon längst vollzogen oder eingeleitet.

Nachfolgend sollen nun einzelne Theorien aus anderen Kognitionsarealen und aus der angewandten Medien- und Instruktionspsychologie vorgestellt werden, die helfen sollen, den differenziellen Einfluss unterschiedlicher Medien auf die menschliche Informationsverarbeitung besser zu verstehen und um daraus allgemeine Vorhersagen für den Einfluss von auch im Bereich der Kausalkognition immer wieder eingesetzten Medien auf kognitive Prozesse bei der menschlichen Informationsverarbeitung ableiten zu können. Im Besonderen soll näher auf die Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) und auf die Auswirkungen externaler Repräsentationen in komplexen Informationsverarbeitungsaufgaben (Zhang, 1997, 2000; Zhang & Norman, 1994) eingegangen werden. Daran anschließend werden aus medien- und instruktionspsychologischer Sicht besondere Charakteristika von Texten, Tabellen und Diagrammen, sowie Theorien zur kognitiven Verarbeitung dieser verschiedenen Medien vorgestellt. Kapitel 6 wird durch empirische Untersuchungen aus der Medienwirkungsforschung abgerundet, die zusammengefasst zu dem Ergebnis gelangen, dass es von der in einer Untersuchung jeweils auszuführenden Aufgabe abhängt, welches Medium am besten abschneidet.

## **6.1 Die Kosten-Nutzen-Theorie von Kleinmuntz & Schkade**

Kleinmuntz & Schkade (1993) haben sich mit dem Einfluss von Informationsanordnungen (englisch: *information displays*) auf kognitive Prozesse bei der Entscheidungsbildung befasst. Innerhalb von psychologischen Theorien zur Entscheidungsbildung befasst man sich unter anderem damit, wie Menschen zwischen Regeln oder Strategien der Entscheidungsfindung auswählen und von welchen Faktoren die Wahl einer Regel abhängt (Jungermann, Pfister & Fischer, 2005). Grundsätzlich wird dabei davon ausgegangen, dass die Wahl einer Regel von Merkmalen der Regel selbst und den Merkmalen einer Aufgabe abhängt. Die Merkmale einer Aufgabe werden wiederum (i) von der Komplexität eines Problems und (ii) von der Art des Informationsangebotes bestimmt. Problemkomplexität und Art des Informationsangebotes haben Auswirkungen auf die Art und Weise, in der Entscheidungen getroffen werden. Im Zusammenhang mit der Art des Informationsangebots berichten Payne, Bettmann & Johnson (1993; s.a. Jungermann et al., 2005) von einigen Studien, in denen der Entscheidungsprozess



von der Art und Weise beeinflusst wird, in der die Informationen angeboten oder dargeboten werden. In diesen Studien wurden beispielsweise die *Konkretheit* (Aschenbrenner, 1978), die *Übersichtlichkeit* (Russo, 1977), die *Vollständigkeit* (Ford & Smith, 1987) und die *Informationsart* bzw. das *Darbietungsformat* (Fischer & Jungermann, 1996; Stone & Schkade, 1991) variiert:

*Konkretheit:* Entscheider tendieren offenbar dazu, nur diejenigen Informationen in Betracht zu ziehen, die in der Beschreibung der Wahlmöglichkeiten explizit gegeben sind. Implizit gegebene Informationen, die nur erschlossen werden können, werden kaum berücksichtigt.

*Übersichtlichkeit:* Liegen die Informationen über Optionen nur verstreut oder verteilt vor, werden sie weniger leicht für Entscheidungen verwendet, als Informationen, die gebündelt und in einer übersichtlichen Form vorliegen.

*Vollständigkeit:* Das Entscheidungsverhalten wird auch davon beeinflusst, ob die Informationen über die zur Verfügung stehenden Entscheidungsalternativen vollständig oder nur teilweise (partiell) angeboten werden.

*Informationsart:* Die Wahl einer Handlungsoption kann auch davon abhängen, ob Informationen in numerischer oder in verbaler Form dargeboten werden.

Auch Kleinmuntz & Schkade (1993) gehen in ihrer Variante der Kosten-Nutzen-Theorie von der grundlegenden Annahme aus, dass der Einsatz von Entscheidungsstrategien durch unterschiedliche Informationsanordnungen entweder erleichtert oder erschwert werden kann. Bedeutsam im Zusammenhang mit der Anordnung von Informationen sind aus ihrer Sicht (i) die Form der einzelnen Informationen (verbal, numerisch oder bildhaft); (ii) die Organisation von Informationen in bedeutungstragende Strukturen und Gruppierungen; und (iii) die Abfolge einzelner Informationsbestandteile. Nach Kleinmuntz und Schkade (1993) beeinflussen Variationen hinsichtlich Form, Organisation und Abfolge von Informationen Prozesse bei der Auswahl von Strategien, indem sie zu Veränderungen des antizipierten Aufwands und der antizipierten Genauigkeit jeder verfügbaren Strategie führen. Antizipiert die urteilende Person aufgrund der Informationsanordnung, dass eine verfügbare, aber relativ komplexe Strategie nur ungenau ausgeführt werden kann oder dass ihr Einsatz mit einem hohen Aufwand verbunden ist, dann stellt dies für sie einen Anreiz dar, auf eine einfachere und ebenfalls verfügbare Strategie zurückzugreifen. Nach der Kosten-Nutzen-Theorie von

Kleinmuntz & Schkade (1993) sind die urteilenden Personen bei der Auswahl einer verfügbaren Strategie also um einen Ausgleich zwischen Aufwandminimierung und Genauigkeitsmaximierung bemüht (englisch: *accuracy-effort-trade-off*). Ist der Verarbeitungsaufwand hoch, dann steigen damit auch die Kosten für eine korrekte Urteilsbildung. Ist der Preis für ein richtiges Urteil zu hoch, greifen die Probanden auf vereinfachende Strukturen zurück. Kleinmuntz & Schkade (1993) berichten von eigenen Studien, in denen nachgewiesen werden konnte, dass sich Form, Organisation und Abfolge von Information jeweils unterschiedlich auf Auswahlprozesse auswirken. So hatte die Organisation einen starken Einfluss auf den Erwerb oder die Aufnahme von Information, während sich die Form eher auf das Kombinieren und Beurteilen von Information auswirkte. Die Abfolge von Information hatte zwar ebenfalls Auswirkungen auf die Informationsaufnahme, doch war der Effekt bei Weitem nicht so ausgeprägt wie in Abhängigkeit von der Organisation. Alle drei Informationsmerkmale beeinflussten auch den Anstrengungsaufwand der Entscheider, auch hier war Einfluss von Organisation am stärksten. Die hier beschriebene Variante der Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) weist große Ähnlichkeiten mit dem Kontingenzmodell von Beach & Mitchell (1978) und dem kognitionspsychologischen Modell des adaptiv-kontingenten Entscheidens von Payne, Bettman & Johnson (1993) auf. Auch in diesen Modellen resultieren Entscheidungen über die Auswahl von Strategien aus einer Kosten-Nutzen-Analyse des Informationsverarbeitungssystems (Jungermann, Pfister & Fischer, 2005).

Im *Kontingenzmodell* von Beach & Mitchell (1978) wird die Auswahl einer Strategie sowohl von Merkmalen der Entscheidungsaufgabe als auch durch Merkmale des Entscheiders (z.B. Wissen über Entscheidungsregeln, Fähigkeit zur Anwendung von Regeln, Motivation) bestimmt. Die Merkmale der Entscheidungsaufgabe hängen einerseits von Merkmalen des einer Aufgabe zugrunde liegenden Problems ab und andererseits von den Merkmalen des Entscheidungsumfeldes. Die zentrale Annahme des Kontingenzmodells besteht darin, dass in Abhängigkeit von Merkmalen des Problems und des Umfeldes diejenige Regel oder Strategie ausgewählt wird, die mit den geringsten Kosten zu einer zufriedenstellenden Entscheidung führt. Demnach ist der Entscheider bei der Strategiewahl also um einen Kompromiss zwischen den erwarteten Kosten und dem erwarteten Nutzen bemüht.

Mit dem Modell des *adaptiv-kontingenten Entscheidens* (Payne, Bettman & Johnson, 1993) kann beschrieben werden, unter welchen situativen Bedingungen welche Entscheidungsregel oder Strategie eingesetzt wird. Ob eine bestimmte Regel oder Strategie aktiviert wird, hängt von den Merkmalen der aktuellen Entscheidungssituation ab (Der Begriff "*kontingent*" wird

im Sinne von "bedingungsabhängig" verwendet) und erfolgt *adaptiv* in Anpassung an die Situation. Auch bei Payne et al. (1993) wählen Menschen in gegebenen Situationen Strategien oder Regeln nach Genauigkeit und kognitivem Aufwand aus. Von einigen Autoren wurde vorgeschlagen, den kognitiven Aufwand als *Berechnungsaufwand* (englisch: *computational effort*) zu definieren (Bettmann, 1979; Shugan, 1980). Je genauer eine Regel ist, desto mehr kognitive Anstrengung oder Berechnungsaufwand ist erforderlich; ein *adaptiver Entscheider* (englisch: *adaptive decision maker*) muss also abwägen, ob sich unter den gegebenen Umständen der Einsatz einer aufwendigen Regel lohnt (Payne, Bettmann & Johnson, 1990, 1993).

Nach Kleinmuntz & Schkade (1993) erweist sich die Kosten-Nutzen-Theorie vor allem bei komplexen Entscheidungsprozessen und Aufgabenstellungen als hilfreich. Eine in diesem Zusammenhang ungeklärte Frage ist, ob die Kosten-Nutzen-Theorie auch zur Erklärung der Strategiewahl bei der Beurteilung der Stärke kausaler Zusammenhänge herangezogen werden kann. Dafür müssten zwei Voraussetzungen erfüllt sein: Zum einen müsste - ähnlich wie bei Prozessen zur Entscheidungsbildung - die Wahl von Strategien zur Beurteilung der Stärke kausaler Zusammenhänge auf Prozessen oder Mechanismen basieren, die nicht oder nicht nur rein automatisiert ablaufen, sondern einer höheren kognitiven Ebene entspringen. In Kapitel 3.1.5 haben wir die Auffassung vertreten, dass es sich bei komplexen regel-basierten Strategien wie zum Beispiel der  $\Delta P$ -Regel oder der Berechnung der kausalen Stärke  $p$  nach der Power PC-Theorie vermutlich um mentale Wissensrepräsentationen in Form von regel-basierten Kausal-Schemata handelt, die sich aus deklarativen, operativen und prozeduralen Anteilen zusammensetzen und dabei höheren Bewusstseins- und Überwachungsprozessen zugänglich sind. Dagegen hängen bei einfachen und hoch automatisiert ablaufenden Aufgaben (welche unter Umständen auch in komplexere Aufgaben- und Beurteilungskontexte eingebettet sein können) Aufwand und Genauigkeit bei der Urteilsbildung vermutlich nicht so sehr von Kosten-Nutzen-Analysen ab, sondern hier spielen eher grundlegende Prinzipien der Wahrnehmung (z.B. Gestaltgesetze) eine zentrale Rolle. Beispielsweise kann die räumliche Nähe zweier benachbarter Informationsteile dazu führen, dass diese als zusammen gehörig wahrgenommen werden. Wie räumliche Nähe, so kann auch die Verwendung von gewöhnlichen oder üblichen Farben dazu führen, dass die Belastung des Arbeitsgedächtnisses reduziert wird, weil mentale Operationen, die im Arbeitsgedächtnis ausgeführt werden, durch perzeptuelle Operationen, die mehr oder weniger automatisiert ablaufen, ersetzt werden (Lohse, 1997; Wickens & Carswell, 1995).

Die zweite Voraussetzung, die erfüllt sein muss, damit die Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) auf den Bereich der Kausalkognition übertragen werden kann, beinhaltet die Annahme, dass Kosten-Nutzen-Abwägungen die Prozesse, die bei der Beurteilung der Stärke kausaler Zusammenhänge ablaufen, beeinflussen können oder in irgendeiner Weise daran beteiligt sind. Nach unserem Kenntnisstand wurde der Einfluss von Kosten-Nutzen-Abwägungen auf die Prozesse bei der Beurteilung der Stärke kausaler Zusammenhänge bislang noch nicht untersucht. Es kann aber angenommen werden, dass Kosten-Nutzen-Abwägungen nicht nur innerhalb der Psychologie der Entscheidung (Jungermann, Pfister & Fischer, 2005) bedeutsam sind, sondern auch im Zusammenhang mit den Prozessen, die beim Erwerb von Kausalwissen und bei der Beurteilung der Stärke von Kausalzusammenhängen eine Rolle spielen. Jungermann, Pfister & Fischer (2005) verstehen Entscheidung unter anderem als "... einen Prozess, dessen zentrale Komponenten Beurteilungen (judgments) und Wahlen (choices) sind" (S. 4), als "... Prozess der vergleichenden Beurteilung und Wahl" (S.7) und als etwas, das "... der Auswahl einer von mehreren Handlungsmöglichkeiten" (S. 7) dient. Wir vermuten, dass bei der Beurteilung der Stärke eines Kausalzusammenhangs auch Entscheidungsprozesse wie beispielsweise Kosten-Nutzen-Analysen mitbeteiligt sind, da die urteilende Person dabei in irgendeiner Weise aus mehreren Alternativen eine geeignete Strategie oder Heuristik auswählen muss und Auswahlprozesse gleichbedeutend mit Entscheidungen sind. Waldmann & Hagmayer (2001) sprechen sich ebenfalls dafür aus, dass Entscheidungsbildungsprozesse an der Beurteilung der Stärke von kausalen Zusammenhängen beteiligt sind. Nach ihrer Auffassung sind der Einsatz von Lernstrategien und die Beurteilung der Stärke von kausalen Zusammenhängen das Produkt einer komplexen Wechselwirkung, an der im Einzelnen das Vorwissen und die Ziele des Lernenden (Anstrengungsvermeidung vs. korrekte Urteilsbildung), die Struktur des Lerninhalts (geordnet vs. ungeordnet), die Art der Lernaufgabe (z.B. einzelfallweise vs. summarische Darbietung) und Schätzungen eines Trade-Offs zwischen Strategien und Lernaufwand beteiligt sind.

## 6.2 Auswirkungen externaler Repräsentationen in komplexen Informationsverarbeitungsaufgaben: Die Theorie von Zhang

Aus theoretischer Sicht lässt sich der Einfluss von unterschiedlichen Medien und Präsentationsformen auf kognitive und informationsverarbeitende Prozesse fruchtbar erklären, wenn verschiedene Medien als Formen von externalen Repräsentationen (Zhang & Norman, 1994; Zhang, 2000, 1997) konzipiert werden:

Komplexe Informationsverarbeitungsprozesse werden von Zhang (1997, 2000; Zhang & Norman, 1994) als verteilte kognitive Aufgaben (englisch: *distributed cognitive tasks*) verstanden, weil die zu verarbeitenden Informationen sowohl internalen als auch externalen Repräsentationen entnommen werden. *Externale Repräsentationen* (s. Zhang 1997, S.180) lassen sich definieren als Wissen und Struktur in der Umgebung, als physikalische Symbole, Objekte oder Dimensionen (z.B. geschriebene Symbole, Dimensionen einer Grafik), und als externe Regeln, Beschränkungen oder Beziehungen, die in physikalische Konfigurationen eingebettet sind (z.B. räumliche Beziehungen zwischen schriftlichen Einheiten, visuell-räumliche Anordnungen von Diagrammen). Informationen aus *internalen Repräsentationen* sind im Gedächtnis als Wissen und Strukturen abgespeichert, beispielsweise in Form von Propositionen, mentalen Landkarten, analogen Bildern, mentalen Modellen, semantischen Netzwerken, Schemata, Skripts oder prozeduralem Wissen. Bei der Bearbeitung einer kognitiven Aufgabe wie z.B. Problemlösen, Entscheidungsbildung oder kausaler Urteilsbildung werden die wahrnehmungs- und gedächtnisbasierten Informationen aus externalen und internalen Repräsentationen wechselweise in einer integrativen und dynamischen Art verarbeitet.

Traditionell wurden externale Repräsentationen wie z.B. Bilder, Diagramme und Grafiken lediglich als Input und Reizmaterial oder als bloße Gedächtnisstützen für das kognitive System verstanden (Zhang, 2000). Für Zhang & Norman (1994; Zhang, 2000, 1997) dagegen sind externale Repräsentationen mehr; sie betonen, dass externale Repräsentationen unser kognitives Verhalten steuern, einschränken und sogar festlegen. Im Unterschied zu internalen Repräsentationen haben externale Repräsentationen eine Reihe von Eigenheiten und nehmen vielfältige Funktionen wahr (Zhang, 2000; Zhang, Johnson, Malin & Smith, 2002; Zhang & Norman, 1994). Externale Repräsentationen ...

- *fungieren als externes Gedächtnis*: Externale Repräsentationen dienen als Gedächtnisstützen, sie lagern Informationen aus oder bilden ständig verfügbare

Datenbanken. Dadurch erweitern sie das Arbeitsgedächtnis und reduzieren die interne Gedächtnisbeanspruchung.

- *dienen als Basis für neues Wissen und neue Fertigkeiten:* In Verbindung mit Studien zur Beziehung zwischen mentalen Bildvorstellungen und externalen Bildern konnten Chambers & Reisberg (1985; Reisberg, 1987) zeigen, dass externale Bilder Wissen und Fertigkeiten bereitstellen, die nicht aus internen Repräsentationen entnommen werden können.
- *reduzieren den Suchaufwand:* Externale Repräsentationen reduzieren den Aufwand für die Suche nach Elementen für einen Problemlöseprozess und machen symbolische Beschreibungen überflüssig, in dem Informationen, die gemeinsam verwendet werden müssen, in Diagrammen räumlich nahe beieinander stehen (Larkin & Simon, 1987). Die in Diagrammen veranschaulichten geometrischen Beziehungen reduzieren den Suchraum, der in satzartigen Systemen syntaktisch möglich wäre (De Jong et al., 1998).
- *unterstützen einfach zu verarbeitende perzeptuelle Inferenzen:* Externale Repräsentationen machen Information verfügbar, die unmittelbar wahrgenommen und genutzt werden kann, ohne dass diese Information interpretiert und explizit formuliert werden müsste. Im Zusammenhang mit Studien zum Problemlösen fanden Larkin & Simon (1987) heraus, dass externale Repräsentationen in Diagrammform auf der Wahrnehmungsebene Operationen unterstützen, durch welche Merkmale einfach erkannt und Schlussfolgerungen unmittelbar abgeleitet werden können.
- *strukturieren (steuern) kognitives Verhalten:* Externale Repräsentationen verankern und strukturieren unbewusst kognitives Verhalten. Die physikalischen Strukturen externaler Repräsentationen beschränken den Aktionsraum für unterschiedliche kognitive Verhaltensweisen und steuern eine effiziente Wissensorganisation. Das beinhaltet, dass einige Handlungen durch externe Repräsentationen überhaupt erst möglich werden, während andere verhindert werden (Norman, 1988; Zhang & Norman, 1994).
- *verändern die Beschaffenheit einer Aufgabe:* Externale Repräsentationen verändern die Beschaffenheit einer Aufgabe, indem durch sie effizientere Handlungsabfolgen hervor gebracht und Aufgaben einfacher werden. Aufgaben mit und ohne externale Repräsentationen sind aus der Sicht des Aufgabenbearbeiters völlig unterschiedlich, selbst wenn die abstrakten Strukturen der Aufgaben identisch sind. (Norman, 1991, 1993; Zhang & Norman, 1994).

- *schränken den Abstraktionsgrad einer Aufgabe ein:* Hoch spezifische externe Repräsentationen wie Euler-Venn-Diagramme schränken beim logischen Denken den Abstraktionsgrad ein und erleichtern dadurch die Verarbeitbarkeit, d.h. die Interpretierbarkeit und die Transparenz von Information (Stenning & Oberlander, 1994).
- *beeinflussen die Problemschwierigkeit und die Strategiewahl:* In den Bereichen Entscheidungsbildung (Kleinmuntz & Schkade, 1993) und Problemlösen (Zhang & Norman, 1994) konnte belegt werden, dass unterschiedliche Repräsentationen oder Darstellungen identischer Informationsinhalte einen starken Einfluss auf die Problemschwierigkeit oder die Strategiewahl bei der Entscheidungsbildung haben. Die Analysen von Kleinmuntz & Schkade (1993) zur Entscheidungsbildung legen nahe, dass Unterschiede in der Strategiewahl motivational bedingt sein können, indem die Probanden in Abhängigkeit von der Form, der Organisation und der Abfolge von Informationen einen Trade-Off zwischen der Maximierung von Genauigkeit und der Minimierung von Aufwand vornehmen.
- *tragen zur Evolution des Denkens bei:* Donald (1991) und andere (Goody, 1987; Oestermeier, 2001; Olson, 1996; Zhang, 2000) heben die kulturelle und evolutionäre Bedeutung externer Repräsentationen bei der Entwicklung modernen Denkens hervor. Neue externe Repräsentationen ermöglichen neue mentale Operationen.

Soll der Versuch unternommen werden, die vielfältigen Auswirkungen externer Repräsentationen auf die menschliche Informationsverarbeitung zu reduzieren, so lassen sich drei wesentliche Besonderheiten identifizieren: (i) Externale Repräsentationen bewirken eine *Unterstützung der Informationsverarbeitung*, in dem sie als externes Gedächtnis Informationen auslagern und Suchprozesse reduzieren (Larkin & Simon, 1987), (ii) externe Repräsentationen machen *Strukturierungsangebote*, in dem sie die Informationsverarbeitung strukturieren und kognitives Verhalten steuern (Larkin & Simon, 1987; Stenning & Oberlander, 1995; Zhang, 2000, 1997; Zhang & Norman, 1994), und (iii) externe Repräsentationen unterstützen unmittelbar zu verarbeitende *perzeptuelle Inferenzen* (Larkin & Simon, 1987).

*Zu (i):* Die Unterstützung der Informationsverarbeitung erfolgt beispielsweise dadurch, dass mittels Gruppierung und räumlicher Anordnung Informationen verfügbar gemacht werden, ohne dabei das Arbeitsgedächtnis zu belasten. Bei komplexen Problemstellungen kann darüber hinaus der Abruf von Metakognitionen durch externe Repräsentationen unterstützt werden (Larkin & Simon, 1987).

Zu (ii): Externale Repräsentationen machen Strukturierungsangebote, indem über die Inhalte und die Anordnung externaler Repräsentationen kognitive Aktivitäten und Prozesse gelenkt und gesteuert werden. Suthers (2001) spricht in diesem Zusammenhang von repräsentationaler Führung (englisch: *representational guidance*). Repräsentationale Führung kann beispielsweise dadurch geschehen, dass bestimmte Informationen hervorgehoben oder salient gemacht werden. Externale Repräsentationen können Informationsverarbeitung und kognitive Aktivität aber auch dadurch steuern, dass sie – je nach Art der externalen Repräsentation - zum einen die Handlungsmöglichkeiten einschränken, zum anderen aber zu bestimmten Handlungen oder kognitiven Prozessen auffordern, bzw. diese nahe legen (Greeno, 1989, 1994). In Verbindung mit dem Aufforderungsgehalt externaler Repräsentationen verwendet Zhang (1997) den aus dem wahrnehmungs-ökologischen Ansatz von Gibson (1979) bekannten Begriff der Affordanzen. Unter *Affordanzen* versteht Gibson (1979) die gleich bleibende Information aus der Umgebung, durch welche Handlungsmöglichkeiten eröffnet und das Verhalten der Wahrnehmenden gesteuert wird.

Zu (iii): Perzeptuelle Inferenzen vereinigen in gewisser Weise die Unterstützungs- und die Strukturierungsfunktion externaler Repräsentationen. Zum einen entlasten sie das Arbeitsgedächtnis, in dem kognitive Operationen wie z.B. Vergleiche und Schlussfolgerungen durch die externale Repräsentation auf die Wahrnehmungsebene verlagert werden, wo sie einfacher und unmittelbarer verarbeitet werden können und so Kapazitäten für andere kognitive Operationen freimachen (s.a. Wickens & Carswell, 1995). Zum anderen kann eine durch die Informationsanordnung nahe gelegte Wahrnehmungsinferenz Denkmöglichkeiten erweitern oder eingrenzen und damit strukturieren.

Beispiele für den Einfluss unterschiedlicher externaler Repräsentationen gleicher Informationsinhalte auf kognitive Prozesse stammen von Zhang & Norman (1994): So werden Multiplikationen schneller und effizienter ausgeführt, wenn die Zahlen in arabischer und nicht in römischer Schreibweise präsentiert werden (73 x 27 vs. LXXIII x XXVII). Im Bereich des Problemlösens haben verschiedene Autoren (s. Zhang & Norman, 1994) herausgefunden, dass unterschiedliche Repräsentationen oder Darstellungen eines Problems einen starken Einfluss auf die Problemschwierigkeit haben können.

Zusammengefasst kommt Zhang (2000, 1997) zu dem Ergebnis, dass die Bearbeitung einer Aufgabe umso leichter und fehlerresistenter ist, je mehr Informationen auf externe Repräsentationen verteilt sind. Den Leistungszuwachs durch externe Repräsentationen erklärt sich Zhang (2000, 1997) unter anderem durch eine effizientere Wahrnehmungsverarbeitung, eine Reduktion der Arbeitsgedächtnisbelastung, effizientere



Verarbeitungsroutinen, das Sichtbarwerden von Handlungen und Rückmeldungen, sowie durch eine Strukturierung und Verankerung von Handlungen.

Im Bereich der Kausalkognition unterstützen Medien und externale Repräsentationen beispielsweise komplexere mentale Prozesse wie z.B. Verallgemeinerungen, Vergleiche, kontrafaktisches Denken und kausale Erklärungen (Oestermeier & Hesse, 2000). Zudem vermitteln Medien kulturelle Formen von Information (Messungen, akkumulierte Häufigkeiten, Formeln, Statistiken), die in natürlichen Umgebungen nicht verfügbar sind; sie reduzieren die Informationsüberflutung der realen Welt und heben einzelne Informationen selektiv hervor (Oestermeier & Hesse, 2000).

Buehner & May (2003b) haben im Verbindung mit der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen an regel-basierten Ansätzen kritisiert, dass beim Prozess der Beurteilung der Stärke kausaler Zusammenhänge von der irrigen Annahme ausgegangen wird, dass die für die Berechnung des Kontingenzmaßes  $\Delta P$  erforderlichen bedingten Wahrscheinlichkeiten  $P(E|C)$  und  $P(E|\neg C)$  für den Urteilenden unmittelbar verfügbar sind. Um aber  $P(E|C)$  und  $P(E|\neg C)$  verfügbar zu haben, muss der Urteilende diese auch berechnen können. Aufgabenbezogene und mediale Faktoren beeinflussen die Verfügbarkeit von Informationen massiv und können somit die Berechnung dieser Wahrscheinlichkeiten erleichtern, erschweren oder gar unmöglich machen.

### **6.3 Charakteristika und kognitive Verarbeitung von Diagrammen, Tabellen und Texten**

Wenn davon ausgegangen werden muss, dass medienbezogene Faktoren bzw. externale Repräsentationen Auswirkungen auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen haben, so wirft das die Frage auf, welche Medien - oder genauer: welche Eigenschaften von Medien – welchen Mechanismus und welche Prozesse zur Verarbeitung von Kausalinformationen nahe legen, erleichtern, auslösen, erschweren, blockieren oder unmöglich machen.

Innerhalb der Forschung zur Beurteilung der Stärke von Kausalzusammenhängen wurden bei summarischer Darbietung häufig Tabellen, Falllisten, Grafiken und Diagramme (jeweils mit Textinhalten) eingesetzt, bei einzelfallweiser Darbietung auch Texte, Zeichnungen und Bilder auf Papier, Karteikarten, Dias und PC-Monitoren. Diese Medien unterscheiden sich teilweise sowohl hinsichtlich ihrer Eigenschaften und Charakteristika, als auch hinsichtlich der Art, wie aus ihnen Informationen entnommen und kognitiv verarbeitet werden. Im Folgenden sollen nun einzelne Medien hinsichtlich ihrer Charakteristika und einiger ausgewählter Aspekte der

Informationsverarbeitung vorgestellt und miteinander verglichen werden (s.a. Ballstaedt, 1997; Oestermeier & Hesse, 2000; Schnotz, 2002):

### 6.3.1 Texte

#### *Besondere Kennzeichen*

Ein mündlicher oder schriftlicher Text ist eine Veräußerung von konzeptuellem Wissen (Ballstaedt, 1997). Die Funktion und der wesentliche Vorzug von Texten liegen darin, dass sie besser als andere Darstellungsformen abstrakte Konzepte und Zusammenhänge vermitteln. Texte setzen sich aus Symbolen zusammen, nämlich aus Sätzen, Wörtern, Silben und Lauten bzw. Buchstaben. Symbole sind Zeichen mit einer willkürlich festgelegten Struktur und einem willkürlichen Gegenstandsbezug. Die Abfolge von Lauten ist teils durch sprachspezifische, teils durch sprachübergreifende phonotaktische Regeln (s. Bußmann, 2002, S.516) festgelegt, die Abfolge von Wörtern in einem Satz durch die Syntax (s. Bußmann, 2002, S.676) einer Sprache.

Ein wesentliches Gütekriterium für einen gelungenen Text ist seine Kohärenz, sein innerer Zusammenhang. Inkohärente Texte bereiten insbesondere Lesern mit geringem Vorwissen zu den dargestellten Textinhalten Probleme.

#### *Kognitive Verarbeitung*

Die kognitive Verarbeitung von Texten wurde umfangreich untersucht (s. Ballstaedt, 1997; van Dijk & Kintsch, 1983; Rickheit & Strohner, 1993 für Überblicke). Ballstaedt (1997) unterscheidet die Ebenen basale Verarbeitung (i), semantisch-syntaktische Verarbeitung (ii), elaborative Verarbeitung (iii), reduktive Verarbeitung (iv) und rekonstruktive Verarbeitung (v). Die Prozesse auf diesen fünf Verarbeitungsebenen laufen eher parallel als sequenziell ab und interagieren miteinander.

Bei der weitgehend automatisiert ablaufenden *basalen Verarbeitung* werden Buchstaben und Laute identifiziert und Buchstaben- bzw. Lautfolgen als Wörter erkannt. Die *semantisch-syntaktische Verarbeitung* dient der Herstellung von Kohärenz und führt am Ende über Textoberflächenrepräsentation zur Erstellung einer propositionalen Repräsentation (Schnotz, 2002; Schnotz & Bannert, 1999), welche bei der *elaborativen Verarbeitung* durch eine Verknüpfung mit dem Vorwissen des Lernenden und die zusätzliche Generierung von Vorstellungen (Assoziationen) und mentalen Modellen angereichert, ausgeschmückt und ergänzt wird. Die *reduktive Verarbeitung* beinhaltet, dass Leser oder Zuhörer Texte in der Regel nicht eins zu eins übernehmen und abspeichern, sondern auf das Wesentliche, d.h. auf seine subjektiv gefärbten Kernaussagen, reduzieren. Bei der *rekonstruktiven Verarbeitung*

wird das auf den anderen Prozessebenen konstruierte mentale Wissen nutzbar gemacht. Dabei wird davon ausgegangen, dass die mentalen Repräsentationen eines Textes nicht nur sprachlich und konzeptuell beschaffen sind, sondern dass aus den propositionalen Repräsentationen eines Textes und aus dem Vorwissen eines Textverstehers ein mentales Modell konstruiert wird, das neben sprachlich-begrifflichen Bestandteilen auch bildhaft-anschauliches Material enthält.

### **6.3.2 Tabellen**

#### *Besondere Kennzeichen*

Tabellen präsentieren numerische Daten in einer zweidimensional angeordneten Form. Sie können auch als übersichtlicher Textersatz angesehen werden, weil sich die tabellarischen Daten auch als Text darstellen lassen. Genauer kann eine Tabelle als eine "Darstellung quantitativer Daten in systematischer räumlicher Anordnung als Matrix von Spalten und Zeilen" definiert werden (Ballstaedt, 1997, S. 137). Tabellen ordnen symbolische Informationen räumlich an, doch hat die räumliche Anordnung selbst - im Gegensatz zu den Diagrammen - keine denotative Funktion (Oestermeier & Hesse, 2000). Die wahrscheinlich auch innerhalb der Forschung zur Kausalkognition am häufigsten eingesetzte Tabellenvariante ist die Kreuztabelle oder Kontingenztafel (s.a. Abbildung 1). Die Zellinhalte einer Tabelle visualisieren die Daten nicht, sondern symbolisieren sie. Als Ganzes gesehen, können Tabellen aber als eine visuelle Argumentationsform verstanden werden, weil ihre Zeilen und Spalten visuelle Vergleiche und die Verallgemeinerung von Daten unterstützen (Oestermeier & Hesse, 2000).

Ballstaedt (1997) grenzt Tabellen, die quantitative Daten in übersichtlicher Form enthalten, von den so genannten *tabellarischen Charts* - oder Worttabellen - ab, in denen qualitative Informationen - also Begriffe, Kategorien und Aussagen - übersichtlich angeordnet sind. *(Fall-)Listen* werden dabei als eine Übergangsform zwischen einem gedruckten Text und einem tabellarischen Chart verstanden.

Tabellen dienen einer genauen und übersichtlichen Präsentation von Daten. Damit dieses Ziel erreicht werden kann, liegen Richtlinien für die Gestaltung von Tabellen vor (Ballstaedt, 1997).

#### *Kognitive Verarbeitung*

Nach Ballstaedt (1997) lassen sich bei der kognitiven Verarbeitung von Informationen aus einer Tabelle vier Schritte unterscheiden:

1. Aktivieren eines Tabellenschemas, welches Wissen über die Organisation einer Tabelle enthält.
2. Verstehen der Beschriftungen und damit der Variablen, die miteinander in Beziehung gesetzt werden.
3. Gezieltes visuelles Suchen aufgrund einer Fragestellung und Lesen der betreffenden Einzeldaten.
4. Gezielte Vergleiche verschiedener Daten und Interpretationen der Daten als Ausprägungen der Variablen.

### 6.3.3 Diagramme

#### *Besondere Kennzeichen*

Diagramme wurden intensiver untersucht als Tabellen und haben dank ihrer einfachen Herstellung durch PC-Software eine sprunghafte Verbreitung gefunden. Ausgangsform bei der Erstellung von Diagrammen sind Tabellen und ihre Inhalte. Schnotz (2002) bezeichnet Diagramme als logische Bilder, in denen abstrakte Sachverhalte dargestellt sind, die zum Teil gar nicht unmittelbar wahrnehmbar sind. Sie stimmen in abstrakter Form mit dem dargestellten Gegenstand strukturell überein, wobei die Übereinstimmung zwischen Diagramm und dargestelltem Gegenstand auf einer Analogierelation basiert. An anderer Stelle definiert Schnotz (2002) Diagramme als Bildzeichen, die *"dem dargestellten Zeichen nicht in ihrer konkreten Erscheinungsform ähneln, jedoch mit ihm auf einer abstrakteren Ebene durch gemeinsame Strukturmerkmale verbunden sind"* (S. 65xx).

Charakteristisch für Diagramme ist, dass sie nicht sichtbare quantitative Zusammenhänge (wie z.B. zeitliche, kausale, logische oder statistische Zusammenhänge) zwischen Variablen (Merkmalen, Eigenschaften, Größen) durch räumliche und topologische Beziehungen zwischen einfachen grafischen Elementen (Punkte, Linien, Flächen usw.) darzustellen vermögen (Ballstaedt, 1997; Oestermeier & Hesse, 2000). Diagramme setzen sich aus den vier Bestandteilen Hintergrund, formaler Rahmen, spezifische Information und Kennzeichnungen zusammen (Kosslyn, 1989).

In einem Diagramm werden mindestens zwei Skalen zueinander in Beziehung gesetzt. Tabellen und Diagramme enthalten äquivalente Informationen, sind aber unterschiedlich auszuwerten. Informationen über Zusammenhänge zwischen Variablen lassen sich aus einem Diagramm schneller entnehmen als aus einer Tabelle, weil die Such- und Zuordnungsprozesse

in einer Tabelle mehr Zeit in Anspruch nehmen können. Was sonst aus den Daten erst erschlossen werden muss, kann in einem Diagramm einfach abgelesen werden.

Diagramme haben in erster Linie die Funktion der *Veranschaulichung* (Ballstaedt, 1997), sie machen komplexe Informationen *visuell salient* (Shah, Freedman & Vekiri, 2005). Visuelle Dimensionen wie z.B. Länge, Fläche oder Farbe werden genutzt, um inhaltliche Größen und Beziehungen darzustellen oder um quantitative Zusammenhänge anschaulich und verständlich zu machen (Shah, Freedman & Vekiri, 2005; Shah & Hoeffner, 2002; Tversky, 2001; Winn, 1987). Diagramme können sich als nützlich erweisen, wenn es darum geht, ein quantitatives oder wissenschaftliches Konzept abzubilden. Das gilt insbesondere dann, wenn das Konzept durch die Anordnung von Information explizit gemacht wird (Larkin & Simon, 1987; Pinker, 1990). Der zentrale Prozess beim Auswerten eines Diagramms ist der *visuelle Vergleich* (Tukey, 1990). Allein durch die Anordnung der Daten können Zusammenhänge unmittelbar und ohne Worte sichtbar gemacht werden.

Weitere Vorzüge in Verbindung mit der Veranschaulichungsfunktion von Diagrammen sind die *Verdichtung* von Informationen und die *Exploration von Zusammenhängen* (Ballstaedt, 1997). Informationen, die in einem Diagramm auf einen Blick zugänglich gemacht werden können, lassen sich sprachlich oder in einem Text oft nur umständlich und unzugänglich darstellen.

Ballstaedt (1997, s.a. Zelazny, 1986) unterscheidet fünf Grundtypen von Diagrammen: Kreis- und Tortendiagramme, Balkendiagramme, Stab- oder Säulendiagramme, Linien- oder Kurvendiagramme, sowie Punkte- oder Streudiagramme.

*Kreis- und Tortendiagramme* kommen vor allem dann zu Einsatz, wenn es um die Darstellung der Anteile einzelner Komponenten an einer Gesamtheit geht (Kosslyn, 1994; Wilkinson, 1999). Allerdings sind sie weniger dazu geeignet, um die verschiedenen Aufteilungen eines Ganzen miteinander zu vergleichen.

*Balkendiagramme* eignen sich besonders zur Darstellung einer Rangfolge von einzelnen Werten (Rangfolgevergleich).

Zwei bekannte Untergruppen von *Säulen- oder Stabdiagrammen* sind Histogramme und Zeitreihen. *Histogramme* zeigen an, wie viele Fälle jeweils in verschiedene Kategorien fallen, *Zeitreihen* zeigen Veränderungen über die Zeit. Eine weitere Variante von Säulendiagrammen ist das unterteilte Säulendiagramm, kurz *Stapeldiagramm* genannt. Ein Spezialfall ist das *100%-Säulendiagramm*, bei dem alle Säulen oder Stapel die gleiche Höhe 100 haben. Es eignet sich besonders dafür, Wahrscheinlichkeiten zueinander in Beziehung zu setzen. Nach einer Untersuchung von Simkin & Hastie (1987) aktivierten die meisten Probanden spontan

Vergleiche zwischen den Größenwerten, wenn sie mit einem Säulendiagramm konfrontiert werden. Eine *Gruppierung* (abstandslose Nebeneinanderstellung) einzelner Säulen ist dann sinnvoll, wenn die betreffenden Säulen eine Gemeinsamkeit aufweisen. Im Gegensatz zu Liniendiagrammen sind Säulendiagramme besonders dazu geeignet, diskrete Vergleiche nahe zu legen (Carswell & Wickens, 1987; Shah, Mayer & Hegarty, 1999; Zacks & Tversky, 1999), darüber hinaus können computergesteuert präsentierte Einzeldaten in vertikalen Säulendiagrammen schneller und genauer identifiziert werden als in Liniendiagrammen, horizontalen Säulendiagrammen, unterteilten Säulendiagrammen und Tortendiagrammen (Wilson & Addo, 1994b).

*Linien- oder Kurvendiagramme* sind eine Weiterentwicklung des Säulendiagramms. Kurven und Linien sind besser als Säulendiagramme dafür geeignet, einen Verlauf- oder Trend anzuzeigen (Carswell, Emery & Lonan, 1993; Shah, Mayer & Hegarty, 1999; Zacks & Tversky, 1999).

*Punkte- oder Streudiagramme* stellen Zusammenhänge oder statistische Korrelationen zwischen zwei Variablen dar.

Die verschiedenen Diagrammartens unterscheiden sich darin, welche Informationen, Argumente und Schlussfolgerungen besonders leicht aus ihnen entnommen werden können. Um in einem Diagramm die Genauigkeit von Vergleichen zu maximieren, sind Längen gegenüber Flächen und Volumina vorzuziehen, da Längen besser eingeschätzt werden können. Zur Unterscheidung von kategorialen Variablen wird der Einsatz kontrastierender Farben (z.B. blau-gelb, grün-rot) empfohlen (Ballstaedt, 1997).

Zusammengefasst eignen sich Tortendiagramme zur Veranschaulichung von Zusammensetzungen und relativen Proportionen, Balkendiagramme für die Erstellung einer Rangfolge, Säulendiagramme für diskrete Vergleiche und zur Veranschaulichung von Häufigkeitsunterschieden, Liniendiagramme für das Sichtbarmachen von Trends und Punkt- oder Streudiagramme für die Darstellung von Korrelationen (Zelazny, 1986). Daraus ergibt sich, dass keine Diagrammform allgemein besser wäre als andere. Stattdessen muss von einer Interaktion zwischen Aufgabe und Präsentationsform ausgegangen werden (s.a. Shah, Freedman & Vekiri, 2005; Vessey, 1991; 1994), welche sich gut in Verbindung mit so genannten Kompatibilitäts- oder Verträglichkeitsmodellen wie z.B. dem *Näheverträglichkeitsprinzip* (englisch: *proximity compatibility principle*, s. Carswell & Wickens, 1987; Wickens & Andre, 1990; Wickens & Carswell, 1995) erklären lässt. Dieses Prinzip basiert auf der Annahme, dass sowohl Aufgaben als auch Präsentationsformen entweder über eine geringe oder über eine hohe Nähe oder Entfernung verfügen. Wenig nahe

oder weit entfernte Aufgaben (englisch: *low proximity tasks*) beziehen sich auf den Gebrauch von einzelnen Datenpunkten (z.B. das Ablesen eines Wertes), das Identifizieren einzelner Tatsachen oder das Vornehmen von diskreten Vergleichen; nahe gelegene (englisch: *high proximity tasks*) Aufgaben beinhalten die Anwendung von ganzen Punktkonfigurationen in der grafischen Anordnung (z.B. der Vergleich von Trends oder die Darstellung von Entwicklungsverläufen). In wenig nahen oder weit entfernten Präsentationsformen (englisch: *low proximity displays*) liegen die Werte weit auseinander oder werden separat gezeigt (z.B. Säulen- oder Balkendiagramm, Tabelle); in nahe gelegenen Präsentationsformen (englisch: *high proximity displays*) liegen die Werte nahe beieinander, sie können ineinander übergehen und sich zu visuellen Mustern zusammenfügen (z.B. Liniendiagramme, Anordnung von Objekten). Das Näheverträglichkeitsprinzip geht davon aus, dass die Leistung dann am besten ausfällt, wenn Aufgabe und Präsentationsform/Informationsanordnung miteinander kompatibel sind: Weit entfernte Aufgaben passen zu weit entfernten Informationsanordnungen (z.B. Wert ablesen oder diskrete Vergleiche zu Säulendiagramm oder Tabelle), nahe gelegene Aufgaben zu nahe gelegenen Informationsanordnungen (z.B. der Vergleich von Trends, die Darstellung von Entwicklungsverläufen oder das Herstellen von Zusammenhängen zwischen zwei oder mehreren Variablen zu Liniendiagrammen).

Für die Erstellung und Evaluation von Diagrammen liegen - wie bei Tabellen auch - bestimmte Richtlinien und Akzeptanzkriterien vor (z.B. Ballstaedt, 1997; Kosslyn, 1989; Shah & Hoeffner, 2002; Tufte, 1983, 1990). Dadurch sollen Missverständnisse und visuelle Manipulationen vermieden werden.

### *Kognitive Verarbeitung*

Die kognitive Verarbeitung von Diagrammen und der darin enthaltenen Informationen wurde umfangreicher beschrieben als diejenige von Tabellen, es liegt dazu eine Reihe von Arbeiten vor (Ballstaedt, 1997; Carpenter & Shah, 1998; Freedman & Shah, 2002; Kosslyn, 1989; Lohse, 1993a; Pinker, 1990; Schnotz, 2002; Shah, 2001; Shah, Freedman & Vekiri, 2005; Shah & Hoeffner, 2002; Winn, 1994). Nach Schnotz (2002) konstruiert der Betrachter beim Verstehen eines Diagramms ein mentales Modell des dargestellten Sachverhalts. Freedman & Shah (2002; s.a. Shah, 2001; Shah, Freedman & Vekiri, 2005) haben ein Konstruktions-Integrations-Modell vorgeschlagen, nach welchem das Verstehen von Diagrammen aus einem Zusammenspiel zwischen Merkmalen der Informationsanordnung und dem Vorwissen des Betrachters über den Inhalt des Diagramms und über Diagramme an sich resultiert. Die dafür notwendigen Prozesse der kognitiven Verarbeitung von Diagrammen greifen ineinander. Aufgabenanalysen ergeben, dass beim Verstehen von Diagrammen drei wesentliche

Prozessebenen eine Rolle spielen (Shah, 2001; Shah, Freedman & Vekiri, 2005; Shah & Hoeffner, 2002):

Auf der ersten Prozessebene enkodiert der Betrachter zunächst die visuelle Anordnung und identifiziert die wichtigen visuellen Merkmale (wie z.B. eine Linienkurve). Schnotz (2002) spricht im Zusammenhang mit dieser Prozessebene von einer *präattentiven oder subsemantischen Verarbeitung* (s.a. Ballstaedt, 1997). Die dabei ablaufenden Prozesse bestehen in der Diskrimination, Identifikation und Gruppierung grafischer Komponenten (also z.B. von Balken, Säulen, Linien, Punkten oder Flächen). Diese Prozesse folgen den Gestaltgesetzen und führen zur Wahrnehmung einer entsprechenden grafischen Konfiguration. Das Ergebnis der präattentiven Verarbeitung bezeichnet Pinker (1990) als *visuelle Beschreibung*. Sie enthält Informationen über die visuellen Relationen innerhalb der grafischen Anordnung, ist aber noch bedeutungsfrei.

Dann, auf der zweiten Prozessebene, der Ebene der *attentiven oder semantischen Verarbeitung* (Schnotz, 2002), geht der Betrachter zum Verstehen des Diagramms über, indem die auf der präattentiven Ebene wahrgenommene grafische Konfiguration konzeptgeleitet analysiert und mit Bedeutung versehen wird. Dabei wird ein Zusammenhang hergestellt zwischen den visuellen Merkmalen einerseits und den wesentlichen quantitativen Tatsachen und konzeptuellen Beziehungen andererseits, die durch diese visuellen Merkmale dargestellt werden (Kosslyn, 1989; Pinker, 1990; Shah, Freedman & Vekiri, 2005; Shah & Hoeffner, 2002). Während der attentiven Verarbeitung nimmt der Betrachter Ablesungen, Vergleiche und Berechnungen vor. Beim Ablesen und Vergleichen lassen sich drei Stufen der Informationsentnahme unterscheiden: (i) Ablesen von Einzelwerten (lokale Suche), (ii) Vergleichen von Einzelwerten um Trends und Gruppierungen zu entdecken (Längs- und Querschnittsvergleiche), und (iii) Vergleichen von Trends und Gruppierungen, um Relationen von Relationen zu erfassen.

Die genaue Analyse eines Diagramms erfordert visuelle Fähigkeiten und wird durch Vorwissen gesteuert (Freedman & Shah, 2002; Pinker, 1990; Shah, Freedman & Vekiri, 2005; Shah & Hoeffner, 2002). Mit bereichsspezifischem Wissen werden Erwartungen mit den Diagramminhalten abgeglichen und überprüft. Von besonderer Bedeutung für das Verstehen von Diagrammen und Visualisierungen allgemein sind *Grafikschemata* (englisch: *graph schemas*, s. Pinker, 1990). Sie enthalten das Vorwissen des Betrachters über Diagramme und Grafiken. In Pinkers Modell (1990) mediierten Grafikschemata den Übergang von einer visuellen Beschreibung zu einer konzeptuellen Botschaft oder Repräsentation. Mit Hilfe dieser erworbenen Schemata wird unter anderem die Informationssuche gesteuert und die



visuelle in quantitative bzw. qualitative Information überführt. Falls die auf der präattentiven Ebene wahrgenommene grafische Konfiguration nicht automatisch ein bestimmtes Grafikschemata aktiviert oder auf eine einfache Weise mit dem Vorwissen oder den Vorerwartungen des Betrachters verknüpft wird, dann sind die im Diagramm enthaltenen Informationen schwerer zu verstehen (Cleveland, 1993; Freedman & Shah, 2002; Kosslyn, 1989, 1994; Pinker, 1990; Shah, Mayer & Hegarty, 1999; Tversky, 2001) und die Betrachter begehen eher Fehler bei der Interpretation.

Wickens & Carswell (1995) haben eine Theorie aufgestellt, die erklärt, wie Grafiken und Diagramme die Anforderungen durch Informationsverarbeitungsprozesse im Arbeitsgedächtnis reduzieren. Nach dieser Theorie können mentale Operationen, die durch das Arbeitsgedächtnis ausgeführt werden, durch mehr oder weniger automatisiert ablaufende Wahrnehmungsprozesse ersetzt werden. Grafiken und Diagramme reduzieren die kognitive Belastung, indem ein Teil der kognitiven Last an das visuelle Wahrnehmungssystem abgegeben wird. Dadurch können kognitive Ressourcen für weitere Verarbeitungsprozesse frei gesetzt werden.

Auch nach Shah, Mayer & Hegarty (1999) lässt sich das grafische Verständnis optimieren, wenn (i) die automatisiert ablaufende Mustererkennung und assoziative Prozesse maximiert und (ii) komplexe Schlussfolgerungsprozesse minimiert werden können. Dadurch kann die Belastung des Arbeitsgedächtnisses reduziert werden, weil automatisiert ablaufende Wahrnehmungsoperationen die das Arbeitsgedächtnis belastenden Schlussfolgerungsprozesse auf einer höheren kognitiven Ebene ersetzen und damit überflüssig machen (Wickens & Carswell, 1995). Entscheidend ist dabei, dass die wesentlichen quantitativen Tatsachen und Beziehungen explizit und external in der grafischen Konfiguration repräsentiert sind und dadurch perzeptuell offensichtlich sind (Freedman & Shah, 2002; Larkin & Simon, 1987; Lohse, 1993a; Pinker, 1990; Zhang, 1997, 2000; Zhang & Norman, 1994).

Die Fähigkeit zum Verstehen eines Diagramms ist eine spezifische Kulturtechnik, die erlernt werden muss und expertiseabhängig ist. Betrachter unterscheiden sich hinsichtlich ihres Wissens über die Verknüpfung zwischen unterschiedlichen visuellen Merkmalen und deren Interpretation (Freedman & Shah, 2002; Shah, Freedman & Vekiri, 2005). Abiturienten, die sich während der Oberstufe stärker für Mathematik und Naturwissenschaften interessiert haben, sind besser in der Lage, Tabellen und Diagramme zu interpretieren, als Abiturienten, die ihre Interessenschwerpunkte auf andere Fächer gelegt haben (Maichle, 1992). Weitere Befunde weisen darauf hin, dass Experten bzw. Lernende mit höherem Vorwissen bei der Diagrammanalyse eher nach übergreifenden visuellen Mustern suchen und Vergleiche

zwischen verschiedenen Daten herstellen, während Lernende mit geringem Vorwissen eher eine lokal begrenzte Suche nach Einzelinformationen vornehmen und die Beziehungen zwischen verschiedenen Daten vereinfachen (Maichle, 1994, Shah, Freedman & Vekiri, 2005). Lowe (1993) nimmt an, dass Individuen mit höheren Lernvoraussetzungen über elaboriertere Grafikschemata verfügen.

Wie gut Individuen dazu in der Lage sind, Informationen und Zusammenhänge zwischen Daten aus Diagrammen und Grafiken zu extrahieren, hängt nicht nur vom Vorwissen der einzelnen Betrachter ab, sondern auch von deren Arbeitsgedächtnis, sowie ihren visuell-räumlichen Fähigkeiten und ihren Fertigkeiten im logischen Denken ab (Shah, Freedman & Vekiri, 2005). So ist das Arbeitsgedächtnis notwendig, damit die grafisch präsentierten Informationen verfolgt und integriert werden können, insbesondere dann, wenn die Grafiken komplex sind (Carpenter & Shah, 1998; Lohse, 1997).

Schließlich müssen auf der dritten Prozessebene der kognitiven Verarbeitung von Diagrammen die auf der zweiten Prozessebene hergestellten wesentlichen quantitativen Tatsachen und konzeptuellen Beziehungen zu den dargestellten Inhalten und den Vorerwartungen eines Betrachters in Beziehung gebracht werden (Shah & Hoeffner, 2002). Die Interpretation der Information, die im Diagramm enthalten ist, hängt also nicht nur vom Wissen über Diagramme ab, sondern auch vom Wissen über die dargestellten Inhalte selbst und den Vorerwartungen des Betrachters. Mehrere Studien (Freedman & Smith, 1996; Jennings et al., 1982; Shah, 2001) belegen beispielsweise, dass Korrelationsschätzungen und Trendwahrnehmungen von Betrachtern von deren Vorerwartungen beeinflusst werden.

Die drei Prozessebenen der kognitiven Verarbeitung von Diagrammen sind inkrementell und interaktiv. Nacheinander enkodieren die Betrachter einen Teil des visuellen Musters; identifizieren, welche quantitative Tatsache oder Funktion darin enthalten ist; und stellen eine Beziehung zu den Inhalten her, auf die die Grafik Bezug nimmt.

In Verbindung mit Forschungsarbeiten zum multimedialen Lernen wurde immer wieder hervor gehoben, dass Diagramme ein Beispiel für so genannte *multiple Repräsentationen* (zu multiplen Repräsentationen s. Van Someren, Reimann, Boshuizen & De Jong, 1998; Ainsworth, 1999; Ainsworth, Bibby & Wood, 2002; Ainsworth & Loizou, 2003) sind, da sich jedes Diagramm aus einer Mischung von Text und Bildern bzw. Grafiken zusammensetzt. Beim Verstehen eines Diagramms muss der Leser bzw. Betrachter textliche (Kennzeichnungen, Legende, Beschriftungen, Überschriften, Zahlen) und grafische (z.B. Linien, Kurven, Säulen) Informationen interpretieren und in kohärenter Weise aufeinander abstimmen, um eine mentale Repräsentation eines Diagramms erstellen zu können (De Jong

et al., 1998). Der Text und die Grafik, aus denen sich ein Diagramm zusammensetzt, können (i) einander ergänzen, indem unterschiedliche Informationen bereit gestellt und unterschiedliche kognitive Prozesse in Gang gesetzt werden; sie können (ii) den Interpretationsspielraum eingrenzen und (iii) zu einem tieferen Verständnis führen (Ainsworth, 1999). In Verbindung mit ihrer Ergänzungsfunktion können sich multiple Repräsentationen beispielsweise dann als nützlich erweisen, (i) wenn unterschiedliche Lerner Präferenzen für unterschiedliche Präsentationsformen haben, (ii) wenn der Lernende unterschiedliche (Teil-) Aufgaben auszuführen hat, (iii) wenn der Einsatz von mehr als einer Strategie die Leistung verbessert und (iv) wenn Präsentationsform und die Auswahl von unterschiedlich elaborierten Strategien miteinander interagieren.

#### **6.4 Empirische Studien zum Vergleich zwischen Diagrammen, Tabellen und Texten**

Nach Jacobs (1994) sind narrative Fließtexte Diagrammen und Tabellen vorzuziehen, wenn nur wenige Daten mitgeteilt werden, deren Beziehungen untereinander entweder recht einfach oder von untergeordnetem Interesse sind. Bei größeren Datenmengen kann eine übersichtlich gestaltete Tabelle die Orientierung innerhalb der Daten vereinfachen. Diagramme sind dann sinnvoll, wenn die Beziehungen der dargestellten Variablen unmittelbarer und mit einem geringeren kognitiven Aufwand erfassbar sind als ohne Grafik (Jacobs, 1994).

Schon 1927 hat Washburne den Einfluss verschiedener Präsentationsformen wie Tabelle, Säulendiagramm, Liniendiagramm und narrativer Texte auf das Erlernen (Wiedergabeleistung) quantitativer Informationen verglichen und kam zu dem Schluss, dass keine Präsentationsform für alle Aufgaben, Zwecke und Begleitumstände gleichermaßen gut geeignet ist. Im Einzelnen konnte Washburne (1927) zeigen, dass Säulendiagramme für eine Präsentation statischer Vergleichsaufgaben ideal sind, während sich Liniendiagramme am meisten bei dynamischen Vergleichen auszeichnen konnten. Für das Ablesen von bestimmten Beträgen eignen sich am ehesten Tabellen mit gerundeten Zahlenwerten. Für nicht empfehlenswert hielt es Washburne (1927), numerische Daten in Textform zu präsentieren. Meyer (1997) hat in seiner Replikation der Untersuchung von Washburne (1927) die Gültigkeit dieser Ergebnisse jedoch eingeschränkt, seine Daten stützen lediglich die Aussage, dass Tabellen beim Erinnern exakter Zahlenwerte am besten abschneiden.

In jüngerer Zeit wurden beim Wissenserwerb und Problemlösen insbesondere *Diagramme* und *Texte* miteinander verglichen. Lohse (1997) vermutet, dass Diagramme und Grafiken im Vergleich zu Texten den Erwerb und die Verarbeitung von Informationen effizienter machen,

weil Informationen, die aus Diagrammen oder Grafiken entnommen werden, leichter wahrgenommen und besser kognitiv verarbeitet werden können. Bei der Verarbeitung von Informationen aus Diagrammen, Grafiken und Texten kommt dem Arbeitsgedächtnis eine entscheidende Rolle zu (Kosslyn, 1989; Lohse, 1997; Pinker, 1990; Wickens & Carswell, 1995). Larkin & Simon (1987) argumentieren beispielsweise, dass Diagramme im Vergleich zu Texten die Anforderungen an das Arbeitsgedächtnis reduzieren, indem sie unmittelbar diejenigen Informationen anzeigen und externalisieren, die für eine Problemlösung benötigt werden.

Scaife & Rogers (1996) haben vorgeschlagen, dass multiple externe Repräsentationen wie z.B. Diagramme zu einem effizienteren Wissenserwerb und zu verbesserten Leistungen beim Problemlösen führen als Texte, (i) weil der kognitive Aufwand durch die Vereinfachung von Such- und Vergleichsprozessen auf der Ebene der Wahrnehmung reduziert wird (Larkin & Simon, 1987); (ii) weil grafische Repräsentationen mehr Informationen externalisieren und dadurch den Informationsapparat entlasten (Zhang & Norman, 1994); und (iii) weil grafische Repräsentationen spezifischer, d.h. weniger vieldeutig als Texte sind, und der Interpretationsspielraum dadurch eingegrenzt und konkretisiert werden kann (Stenning & Oberlander, 1995). Ainsworth & Loizou (2003) fanden zudem heraus, dass Lernende mit Diagrammen signifikant mehr Selbst-Erklärungen generieren als Lernende mit Texten.

Überblicke zum speziellen Vergleich zwischen *Tabellen* und *Diagrammen* fallen uneinheitlich aus (Coll & Coll, 1993; DeSanctis, 1984; Jacobs, 1994, 1999; Jarvenpaa, Dickson & DeSanctis, 1985), mal schneiden Tabellen besser ab, mal verschiedene Arten von Diagrammen, mal ergeben sich keine wesentlichen Unterschiede zwischen den verschiedenen Informationsanordnungen. Insgesamt lässt sich aus den uneinheitlichen Ergebnissen lediglich ableiten, dass unterschiedliche Informationsanordnungen sich für unterschiedliche Umstände eignen.

Meyer, Shinar & Leiser (1997) haben sechs Variablenkategorien identifiziert, die nach ihrer Ansicht die relative Effizienz unterschiedlicher Präsentationsformen beeinflussen: die grafischen Eigenschaften der Informationsanordnung, die Präsentationsbedingungen, die Komplexität der Informationen, die Aufgabe (z.B. Ables-, einfache Vergleichsaufgaben und globale Einschätzungen bei Meyer et al., 1997), Anwender- oder Lernercharakteristika (z.B. Vertrautheit mit der jeweiligen Präsentationsform; Verarbeitungskapazität, Motivation, Vorwissen und Vorerwartungen) und schließlich das Kriterium für eine optimale Leistung. Die Vielzahl relevanter Einflussfaktoren und die zwischen ihnen existierenden komplexen Wechselwirkungen machen eine empirische Untersuchung aller möglichen Kombinationen

von relevanten Variablen praktisch unmöglich (Coll, Coll & Thankur, 1994; Meyer, Shinar & Leiser, 1997; Meyer, 2000).

Jacobs (1999; s.a. Coll, 1992; Coll, Coll & Thankur, 1994) argumentiert, dass es neben den Untersuchungsbedingungen wesentlich vom Verwendungszweck, der Fragestellung und der Information, die man einer Präsentation entnehmen will, abhängt, welche Präsentationsform sich als nützlicher erweist. So können beispielsweise einzelne Daten mit Hilfe einer Tabelle signifikant genauer und schneller herausgefunden werden (so genanntes "*point reading*"), als mit Säulen- oder Liniendiagrammen (Carter, 1947; Coll, 1992; Meyer, 2000; Meyer, Shinar & Leiser, 1997; Wilson & Addo, 1994a). Zudem ermöglichen Tabellen bei mehreren Datenreihen eine einfachere Orientierung und Zuordnung von Kategorien zu Daten als Diagramme (Lohse, 1993b). Diagramme haben dagegen im Vergleich zu Tabellen Vorteile (für einen Überblick s. Jacobs, 1999), wenn es um Fragestellungen geht, in denen (i) mehrere Daten auf einmal miteinander verglichen oder zueinander in Beziehung gesetzt werden müssen (Coll, 1992; Spence, 1990), (ii) Daten aggregiert werden und auf der Basis dieser Aggregationen Gruppenvergleiche vorgenommen werden müssen (Jacobs, 1990; Spence & Lewandowsky, 1991), und (iii) Strukturen, Muster oder qualitative Merkmale (wie z.B. ein Trend) erfasst werden sollen (Coll, Thyagarajan & Chopra, 1991; Jacobs, 1990, 1994). Sollen allerdings relativ einfache Größenvergleiche zwischen einzelnen Datenpunkten (z.B. A größer B) vorgenommen werden, dann schneiden Tabelle und Diagramm meist ziemlich ähnlich ab (Jacobs, 1994; Spence und Lewandowsky, 1991; Wilson & Addo, 1994a).

#### **6.4.1 Die Theorie der kognitiven Passung von Vessey**

Die uneinheitlichen Ergebnisse im Hinblick auf die Frage, ob bzw. unter welchen Umständen Diagramme geeigneter als Tabellen sind oder umgekehrt, lässt sich sowohl mit dem Näheverträglichkeitsprinzip (s. Kapitel 6.3.3.; Carswell & Wickens, 1987; Wickens & Andre, 1990; Wickens & Carswell, 1995) als auch mit der *Theorie der kognitiven Passung* (englisch: *cognitive fit theory*) von Vessey (1991, 1994), einer weiteren Variante der Kompatibilitätstheorie erklären. Die Theorie der kognitiven Passung wurde im Zusammenhang mit der Erforschung des Einflusses von Präsentationsformaten im Bereich der Entscheidungsbildung entwickelt und geht wie andere Kompatibilitätstheorien, von der Annahme aus, dass die Wirksamkeit eines bestimmten Präsentationsformats von der auszuführenden Aufgabe abhängt (Benbasat, Dexter & Todd, 1986; DeSanctis, 1984; Jarvenpaa & Dickson, 1988; Tan & Benbasat, 1990; Vessey, 1991). Nach der Theorie der kognitiven Passung sind Tabellen als symbolische Informationsanordnungen besser für

symbolische Aufgaben geeignet, in denen es darauf ankommt, diskrete Werte abzulesen oder zu extrahieren, während Diagramme als räumliche Informationsanordnungen zu so genannten räumlichen Aufgaben passen, in denen Beziehungen zwischen diskreten Symbolklassen hergestellt werden sollen (Vessey, 1991). Analytische Prozesse wie z.B. Berechnungen, können demnach besser mit einer Tabelle angefertigt werden, während Wahrnehmungsprozesse wie z.B. visuelle Vergleiche besser mit Diagrammen vorgenommen werden können. Ist dagegen eine Passung zwischen Präsentationsform und Aufgabe nicht gegeben, dann muss der Aufgabenbearbeiter mehr kognitiven Aufwand betreiben, um die Information in eine passende Form für die Aufgabenbearbeitung umzuwandeln (Vessey, 1994). Ähnlich wie in der Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) geht ein gesteigerter kognitiver Aufwand mit einer geringeren Genauigkeit und einem größeren Zeitaufwand bei der Aufgabenbearbeitung einher.

Zusammengefasst scheinen narrative Fließtexte dann angebracht zu sein, wenn nur wenige Daten mitgeteilt werden, deren Beziehungen untereinander entweder recht einfach oder von untergeordnetem Interesse sind (Jacobs, 1994). Die Frage, ob Diagramme Tabellen vorzuziehen sind, lässt sich nicht einfach beantworten und scheint unter anderem vom Verwendungszweck und der Wechselwirkung zwischen Aufgabe und Präsentationsform abzuhängen. Die Überlegenheit der Tabelle zeigt sich vornehmlich dann, wenn exakte Daten abgelesen und gegebenenfalls Rechnungen mit diesen Daten vorgenommen werden müssen (Ballstaedt, 1997; Jacobs, 1994). Geht es darüber hinaus um das Sichtbarmachen von Gruppierungen, das Nahelegen von Vergleichen, die Herstellung von Beziehungen oder das Verfolgen von Trends und zeitlichen Verläufen, dann sind Diagramme sinnvoller. Ein Vorteil für Diagramme ist zudem dann erkennbar, wenn die Anzahl der zu vergleichenden Datenwerte zunimmt (Jacobs, 1994). Deutlicher als Tabellen können Diagramme die Funktion eines visuellen Arguments (Ballstaedt, 1997; Oestermeier & Hesse, 2000) ausfüllen.

## **7 Fazit und Implikationen für die empirischen Untersuchungen**

### **7.1 Zusammenfassung der wesentlichen Ergebnisse**

Zentral ist nach unserer Auffassung die Annahme, dass es nicht ausreicht, die Prozesse bei der Beurteilung der Stärke kausaler Zusammenhänge nur auf der Ebene der Kognition zu betrachten. Stattdessen erachten wir es für sinnvoll, den Prozess der kausalen Urteilsbildung als eine Interaktion von Individuum und Informationsumgebung zu begreifen (Clark, 1997; Hutchins, 1995a, b; Oestermeier & Hesse, 2000; Zhang, 1997, 2000; Zhang & Norman, 1994), wobei der Einfluss interindividueller Unterschiede nicht ausgespart werden darf. Innerhalb der Forschung zu den Informationsverarbeitungsprozessen bei der Beurteilung der Stärke von kausalen Zusammenhängen ist diese Sichtweise lange vernachlässigt worden, dort nehmen Theorien zur kausalen Urteilsbildung hauptsächlich auf kognitive Mechanismen Bezug. Der Einfluss individueller Unterschiede oder der Informationsumgebung wurde marginalisiert und blieb bei der Theorienbildung weitgehend unberücksichtigt.

In Kapitel 3 haben wir die Prozesse bei der Beurteilung der Stärke kausaler Zusammenhänge auf einer kognitiven Ebene betrachtet. Dabei ergab sich ein uneinheitliches Bild: Verschiedene Ansätze konkurrieren hier miteinander, manche Befunde können am besten durch verschiedene regel-basierte Ansätze erklärt werden, andere besser durch assoziationalistische Ansätze oder durch Theorien, welche den Einfluss des Vorwissens (mechanistische Theorien oder die Theorie der Kausalmodelle) hervorheben. Der Versuch, auf kognitiver Ebene eine allgemeingültige Theorie zu formulieren, die allen Befunden gerecht wird, kann als gescheitert angesehen werden. Stattdessen wurde von manchen Autoren (Baker, Murphy, Vallée-Tourangeau, 1996; Catena & Perales, 2002) ein evolutionärer Standpunkt eingenommen. Demnach kann die Verarbeitung von Informationen bei der Beurteilung der Stärke von Kausalzusammenhängen durch unterschiedliche Mechanismen erklärt werden. Die Frage allerdings, welcher Mechanismus unter welchen situativen oder umgebungsbezogenen Bedingungen aktiviert wird, wurde bislang nur vereinzelt gestellt und konnte nur sehr vage beantwortet werden.

Shanks (1991) ist einer der wenigen, die explizit davon ausgehen, dass der Prozess der Beurteilung der Stärke von Kausalzusammenhängen als eine Interaktion von Individuum und Informationsumgebung verstanden werden kann. In der Theorie von Shanks (1991) wird die

Auslösung unterschiedlicher kognitiver Prozesse bei der kausalen Urteilsbildung vom Situationstyp abhängig gemacht. Nach Shanks (1991) liegen der Bildung von Kausalurteilen entweder assoziative Mechanismen oder Kausalüberzeugungen zu Grunde. Assoziative Mechanismen werden in erlebten Situationen aktiviert, Kausalüberzeugungen in beschriebenen Situationen. Der Status regel-basierter Ansätze bleibt in der Theorie von Shanks (1991) unklar, in dieser Arbeit wurde die Vermutung artikuliert, dass es sich dabei um bereichsspezifische oder bereichsübergreifende Kausalüberzeugungen oder kausale Schemata handelt. Kritisiert wurde die Theorie von Shanks (1991) dafür, dass darin nicht näher spezifiziert wird, wie Individuum und Informationsumgebung auf der Ebene der Kausalkognition interagieren. Shanks (1991) Unterscheidung zwischen zwei Situationstypen ist sehr grob, der Einfluss unterschiedlicher Präsentationsformen oder des Verarbeitungsaufwandes beispielsweise spielt in seiner Theorie überhaupt keine Rolle.

Im Gegensatz zur Theorie von Shanks (1991) machen Oestermeier & Hesse (2000) detailliertere Angaben darüber, welche unterschiedlichen Präsentationsformen und Visualisierungen welche unterschiedlichen argumentspezifischen Schlussfolgerungen unterstützen können. Oestermeier & Hesse (2000) lassen es aber im Einzelnen offen, welche kognitiven Prozesse oder Mechanismen bei welchen argumentspezifischen Schlussfolgerungen ablaufen bzw. ihnen zugrunde liegen.

Weil es nicht ausreicht, die Prozesse, die bei der Beurteilung der Stärke kausaler Zusammenhänge ablaufen, nur auf der Ebene kognitiver Mechanismen zu analysieren, wurden in Kapitel 4 Überblicke zu Studien gegeben, die belegen, dass neben aufgabenbezogenen Faktoren wie z.B. die Fragestellung einer Untersuchung auch medienbezogene Faktoren wie die Darbietungsart und die Präsentationsform einen Einfluss darauf haben, wie Informationen über Ursache-Wirkungs-Zusammenhänge verarbeitet werden und welche Strategien die Probanden bei der Beurteilung der Stärke von kausalen Zusammenhängen heranziehen.

Der Einfluss medialer Faktoren auf die kognitiven Prozesse bei der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen allgemein und auf den Verarbeitungsaufwand im Besonderen wurde von theoretischer Seite lange vernachlässigt (s.a. Oestermeier & Hesse, 2000). Der Theorienmangel zum Einfluss medialer Faktoren innerhalb der Forschung zur Kausalkognition bedingt, dass der unterschiedliche Einfluss von unterschiedlichen Darbietungsarten oder Präsentationsformen wie z.B. Texten, Tabellen und Diagrammen in unterschiedlichen Variationen weder von regel-basierten noch von assoziationalistischen, noch von irgend einem anderen Ansatz im Bereich der Kausalkognition vorhergesagt werden kann.



Bisherige Untersuchungen (Arkes & Harkness, 1983; Buehner et al., 2003; Kao & Wasserman, 1993; Lober & Shanks, 2000; Ward & Jenkins, 1965) legen den Schluss nahe, dass insbesondere die Darbietungsart einen Einfluss auf die Prozesse bei der Beurteilung der Stärke von kausalen Zusammenhängen hat. So machen bei summarischer Darbietung die Probanden eher Gebrauch von komplexen regel-basierten Strategien wie der  $\Delta P$ -Regel oder der Berechnung der kausalen Stärke  $p$  nach der Power PC-Theorie, zudem fallen ihre Kontingenzurteile genauer aus als bei einer Einzelfalldarbietung. Aufgrund dieser Ergebnisse gingen einige Forschergruppen (s. z.B. Arkes & Harkness, 1983; Baker et al., 1996; Kao & Wasserman, 1993; Shaklee & Mims, 1982) davon aus, dass bei einer summarischen im Vergleich zur einzelfallweisen Darbietung die Gedächtnisbeanspruchung reduziert wird, und dass dadurch wiederum die Komplexität der Strategie oder Regel, die von den Probanden bei der Beurteilung der Stärke von kausalen Zusammenhängen herangezogen wird, steigt. Im Vergleich zum Einfluss des Faktors Darbietungsart wurden der Einfluss der Präsentationsform und der Informationsart im Bereich der Forschung zur Kausalkognition weit weniger oder gar nicht untersucht.

In Kapitel 5.1 konnte herausgearbeitet werden, dass neben den medienbezogenen Faktoren Darbietungsart und Präsentationsform auch individuelle Unterschiede einen starken Einfluss auf die kausale Urteilsbildung und die Wahl der Strategien bei der Beurteilung der Stärke kausaler Zusammenhänge haben können. Unklar blieb aber, worin sich Probanden, die regelbasiert vorgehen, von solchen unterscheiden, die nicht regelbasiert vorgehen. Weil darüber spekuliert wurde (z.B. von Shaklee & Hall, 1983), dass eventuell individuelle Unterschiede im Hinblick auf die mathematischen Fertigkeiten einen Einfluss auf die Komplexität der Strategiewahl bei der Beurteilung der Stärke von Kausalzusammenhängen haben können, halten wir es für aussichtsreich, den Einfluss der numerischen Verarbeitungskapazität auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen zu untersuchen. Vereinzelt wurde darauf hingewiesen, dass verschiedene Individuen in ein- und derselben Untersuchung jeweils von unterschiedlichen Strategien Gebrauch machen (Anderson & Sheu, 1995; Lober & Shanks, 2000; Perales & Shanks, 2003; Wasserman, 1990). Von White (1998) stammt die Beobachtung, dass bei einer übersichtlichen Darbietung von Ursache-Wirkungs-Zusammenhängen in einer Vierfeldertafel nur wenige Probanden idiosynkratische Strategien einsetzen, während bei hoher Komplexität der Datenpräsentation mehrheitlich von idiosynkratischen Strategien Gebrauch gemacht wird.

In Kapitel 5.2 wurden noch einmal näher die Auswirkungen beleuchtet, die eine zu hohe Gedächtnisbeanspruchung auf die Beurteilung der Stärke von Kausalzusammenhängen haben

kann. Zum einen kann es dazu kommen, dass (i) fehlerhafte und/oder verzerrte Kausalurteile gebildet werden (Fiedler, 2000; Kao & Wasserman, 1993; Shaklee & Mims, 1982), zum anderen dazu, dass (ii) die Probanden einfachere und ungenauere Strategien einsetzen (Kleinmuntz & Schkade, 1993; Shaklee & Mims, 1982). Neben der Gedächtnisbeanspruchung spielen vermutlich auch noch individuelle Voraussetzungen, motivationale und situative Faktoren eine Rolle bei Urteilsgenauigkeit und Strategiewahl; es kann also auch hier wieder eine dynamische Interaktion zwischen situativen (z.B. Informationsanordnung) und kognitiven Faktoren angenommen werden.

Die Arbeiten von Fiedler (2000), Kleinmuntz & Schkade (1993), Shaklee & Mims (1982), sowie Waldmann & Hagmayer (2001) sind Beispiele für Erklärungsansätze, in denen mehr oder weniger explizit von einer dynamischen Interaktion zwischen umweltbezogenen und kognitiven Faktoren ausgegangen wird. Ähnlich wie bei Kleinmuntz & Schkade (1993) sind auch nach Waldmann & Hagmayer (2001) Probanden bestrebt, den Verarbeitungsaufwand einer Lernaufgabe möglichst gering zu halten. In ihren Augen sind an der dynamischen Interaktion zwischen situativen und kognitiven Faktoren (i) das Vorwissen und die Ziele des Lernenden, (ii) die Struktur des Lerninhalts, (iii) die Art der Lernaufgabe (z.B. einzelfallweise vs. summarische Darbietung) und (iv) Schätzungen eines trade-off zwischen Strategien und Lernaufwand beteiligt. Fiedler (2000) hat sich in seinem kognitiv-ökologischen Ansatz um das Verständnis des Zustandekommens von Urteilsfehlern und Urteilsverzerrungen bemüht. Nach Fiedler (2000) entstehen Fehler nicht durch eine fehlerhafte Urteilsbildung, sondern durch eine fehlerhafte Informationssuche. Dabei hängt die Qualität der Informationssuche im hohen Maß von der Perspektive des Informationssuchenden, sowie von der Verteilung, Verfügbarkeit, Nähe und Salienz von Informationen in der Umgebung ab.

Oestermeier & Hesse (2000) haben für den Bereich der Kausalkognition die fehlende Berücksichtigung einer dynamischen Interaktion zwischen kognitiver Informationsverarbeitung und Informationsumgebung explizit kritisiert. Hinsichtlich der verschiedenen Denkmuster oder -mechanismen, die bei der kausalen Urteilsbildung zum Einsatz kommen, unterscheiden Oestermeier & Hesse (2000) unter anderem zwischen assoziationalistischen und perzeptuellen Inferenzen einerseits und reflektierten, bewusstseinsnahen Inferenzen andererseits. Letztere bilden die Grundlage für eine regelbasierte Beurteilung der Stärke von Kausalzusammenhängen, weil dabei auf der Basis von kausalen Schemata und Skripten komplexe mentale Operationen und Vergleichsprozesse ausgeführt werden. Individuen, die über keine hohe Verarbeitungs- oder Gedächtniskapazität verfügen und nicht im Einsatz komplexer regel-basierter Strategien geübt sind, benötigen für

den fehlerfreien Gebrauch regel-basierter Strategien die zusätzliche Unterstützung durch Externalisierungen und visuelle Strukturierungen des Informationsmaterials. Nach Oestermeier & Hesse (2000) sind Medien wie z.B. Tabellen, Vierfeldertafeln oder Diagramme dabei mehr als nur eine periphere Gedächtnisstütze, weil sie darüber hinaus die Salienz einzelner Informationsteile beeinflussen, eingrenzen, Vergleichsprozesse nahe legen oder Schlussfolgerungen anregen.

Die bislang im Bereich der Kausalkognition theoretisch eher vernachlässigte Position, dass bei der Analyse kognitiver Prozesse und Mechanismen von einer dynamischen Interaktion zwischen kognitiver Informationsverarbeitung und Informationsumgebung ausgegangen werden sollte, ist in anderen kognitionspsychologischen Forschungsbereichen (Gibson, 1979; Kleinmuntz & Schkade, 1993; Larkin & Simon, 1987; Sanfey & Hastie, 1998; Stenning & Oberlander, 1994, 1995; Vessey, 1991, 1994; Zhang, 1997, 2000; Zhang & Norman, 1994) und insbesondere innerhalb der angewandten Kognitionswissenschaft (Clark, 1997; Hutchins, 1995a, b; Mayer, 1997, 2001; Salomon, 1979, 1994; Zhang, Johnson, Malin & Smith, 2002) schon längst etabliert. Die in Kapitel 6.1 vorgestellte Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) stammt aus der Forschung zur Entscheidungsbildung und wurde als Erklärung dafür herangezogen, wie bei komplexen Entscheidungsprozessen und Aufgabenstellungen unterschiedliche Formen, Organisationen und Abfolgen von Information die Auswahl von Strategien zur Verarbeitung von Information beeinflussen können. Wir gehen dabei von der Annahme aus, dass die Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) als Erklärung für einen möglichen Einfluss der medienbezogenen Faktoren Darbietungsart, Präsentationsform und Informationsart auf die Strategiewahl bei der Beurteilung der Stärke kausaler Zusammenhänge in Frage kommt.

Der Ansatz von Zhang (1997, 2000; Zhang & Norman, 1994), in dem unterschiedliche Medien als Formen von externalen Repräsentationen konzeptualisiert werden, wurde in Kapitel 6.2 vorgestellt, um zu verdeutlichen, wie komplex und vielfältig die Auswirkungen unterschiedlicher Informationsumgebungen auf die kognitive Informationsverarbeitung sein können.. Dabei konnten drei zentrale Auswirkungen von externalen Repräsentationen auf die menschliche Informationsverarbeitung herausgearbeitet werden: (i) Sie unterstützen und entlasten die Informationsverarbeitung (Larkin & Simon, 1987); (ii) sie machen Strukturierungsangebote und steuern kognitives Verhalten (Larkin & Simon, 1987; Stenning & Oberlander, 1995; Zhang, 2000, 1997; Zhang & Norman, 1994); (iii) sie verlagern komplexe kognitive Operationen wie z.B. Vergleiche oder Schlussfolgerungen durch

Veranschaulichung auf die Wahrnehmungsebene und ermöglichen so einfach zu verarbeitende perzeptuelle Inferenzen (Larkin & Simon, 1987).

Kapitel 6.3 gibt einen Überblick zu besonderen Kennzeichen von Texten, Tabellen und Diagrammen und zur kognitiven Verarbeitung der von diesen Medien übermittelten Informationen. Texte, Tabellen und Diagramme wurden ausgewählt, weil sie im Bereich der Forschung zur Kausalkognition besonders häufig zum Einsatz kommen. Während Texte sich zur Vermittlung abstrakter Konzepte und Zusammenhänge eignen, unterstützen Tabellen und Diagramme die menschliche Informationsverarbeitung beispielsweise dadurch, dass sie zu visuellen Vergleichen anregen. Ein besonderer Vorzug von Tabellen liegt darin, dass die darin enthaltenen Daten genau und übersichtlich präsentiert werden können. Diagramme heben sich gegenüber Tabellen dadurch hervor, dass sie nicht sichtbare quantitative Zusammenhänge zwischen Variablen veranschaulichen können.

Wickens & Carswell (1995) haben eine Theorie aufgestellt, die erklärt, wie Grafiken und Diagramme die Anforderungen durch Informationsverarbeitungsprozesse im Arbeitsgedächtnis reduzieren. Nach dieser Theorie können mentale Operationen, die durch das Arbeitsgedächtnis ausgeführt werden, durch mehr oder weniger automatisiert ablaufende Wahrnehmungsprozesse ersetzt werden. Grafiken und Diagramme reduzieren die kognitive Belastung, indem ein Teil der kognitiven Last an das visuelle Wahrnehmungssystem abgegeben wird. Dadurch können kognitive Ressourcen für weitere Verarbeitungsprozesse frei gesetzt werden.

Stapeldiagramme, die entsprechend den Richtlinien und Akzeptanzkriterien für die Gestaltung von Diagrammen (Ballstaedt, 1997; Kosslyn, 1989; Shah & Hoeffner, 2002; Tufte, 1983, 1990) erstellt worden sind, haben nach unserer Einschätzung als Medium zur Darstellung von summarisch dargebotenen Ursache-Wirkungs-Zusammenhängen ein besonderes Potenzial, eine die Stärke von kausalen Zusammenhängen beurteilende Person zum Gebrauch komplexer regel-basierter Strategien anzuregen, weil sie den Vergleich und das Zueinander-In-Beziehung-Setzen von Erfolgs- und Basisrate entsprechend der  $\Delta P$ -Regel oder der Power PC Theorie unmittelbar nahe legen. Die Informationen zur Beurteilung der Stärke eines Kausalzusammenhangs können im Stapeldiagramm nach außen verlagert und summarisch aufbereitet (Entlastungsfunktion), sowie gruppiert und vorstrukturiert (Strukturierungsfunktion) werden. Darüber hinaus können durch eine räumliche Gegenüberstellung von Erfolgs- und Basisrate komplexe mentale Operationen (wie z.B. die Differenzbildung zwischen Erfolgs- und Basisrate bei der Berechnung des Kontingenzmaßes  $\Delta P$ , oder das Ins-Verhältnis-Setzen der Differenz von Erfolgs- und Basisrate zur Basisrate bei

der Berechnung der kausalen Stärke  $p$  nach der Power PC-Theorie) visuell veranschaulicht, unmittelbar nahe gelegt und auf die Wahrnehmungsebene verlagert werden (Ermöglichen von perzeptuellen Inferenzen). Empirische Befunde zum Vergleich zwischen Diagrammen und Tabellen (Jacobs, 1994, s.a. Kapitel 6.4) stützen die Vermutung, dass ein Diagramm einer Tabelle vorzuziehen ist, wenn Gruppierungen sichtbar gemacht, Vergleiche nahe gelegt und quantitative Beziehungen hergestellt werden sollen. Weitere Studien belegen, dass die Wirksamkeit einer bestimmten Präsentationsform von der auszuführenden Aufgabe abhängt (Benbasat, Dexter & Todd, 1986; DeSanctis, 1984; Jarvenpaa & Dickson, 1988; Tan & Benbasat, 1990; Vessey, 1991) und dass darüber hinaus die Fähigkeit zum Verständnis von Diagrammen in einem hohen Maße expertiseabhängig und an individuelle Lernvoraussetzungen gebunden zu sein scheint (Freedman & Shah, 2002; Lowe, 1993; Maichle, 1992, 1994).

## **7.2 Die vier Ebenen der Analyse kognitiver Prozesse bei der Beurteilung der Stärke kausaler Zusammenhänge**

In dieser Arbeit wird die These vertreten, dass es nicht ausreicht, die bei der Beurteilung der Stärke von kausalen Zusammenhängen ablaufenden Prozesse und Mechanismen nur auf einer kognitiven Ebene zu analysieren. Um zu einem besseren und adäquateren Verständnis dieser Prozesse und Mechanismen zu gelangen, soll hier ein Mehrebenenansatz vorgestellt werden (Beispiele für Mehrebenenansätze finden sich in zahlreichen anderen psychologischen Forschungsbereichen, unter anderem bei Bronfenbrenner, 1979; Salomon, 1979; Zhang, Johnson, Malin & Smith, 2002). Nach dem hier vorgeschlagenen Mehrebenenansatz sollte der Vorgang der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen nicht nur auf der im Bereich der Kausalkognition traditionell verankerten kognitiven Ebene beschrieben und analysiert werden, sondern auch auf einer Aufgabenebene, einer Medienebene und einer Ebene interindividueller Unterschiede. Insbesondere die Abschnitte 4 und 5 haben gezeigt, dass zahlreiche Untersuchungen Belege dafür enthalten, dass aufgaben- und medienbezogene Faktoren, sowie die individuellen Unterschiede und Voraussetzungen die Strategiewahl bei der Beurteilung der Stärke von Kausalzusammenhängen, sowie die dabei ablaufenden kognitiven Prozesse und Mechanismen beeinflussen. Nachfolgend sollen diese vier Ebenen kurz beschrieben werden:

- (i) **Aufgabenebene:** Auf der Aufgabenebene wird spezifiziert, was zu tun ist. Dabei wird einer die Aufgabe ausführenden Person unter Umständen nicht nur mitgeteilt, was sie zu tun hat, sondern auch, wie sie es zu tun hat. Aufgabenbezogene Faktoren legen das Ziel

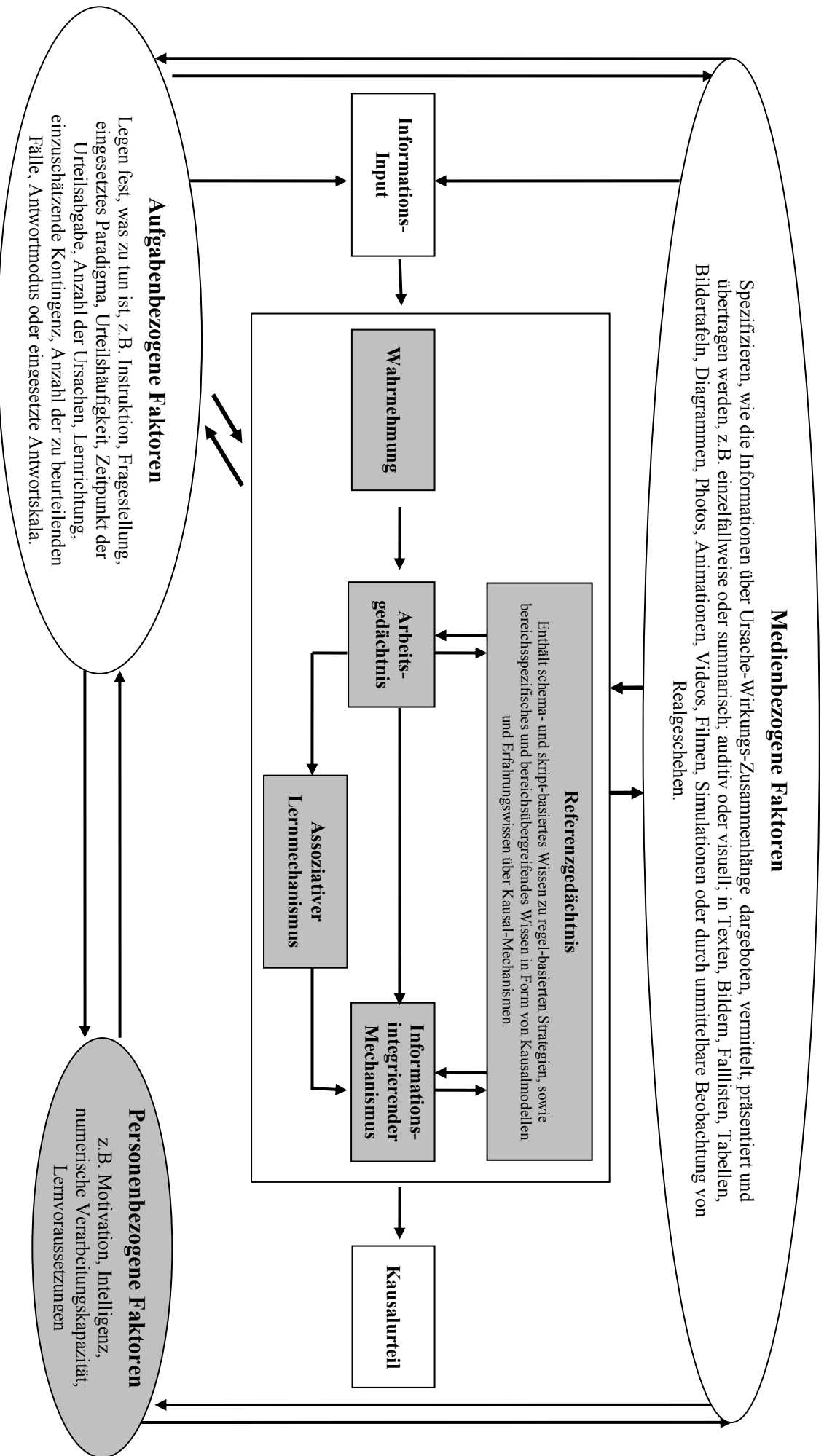
der Aufgabenbearbeitung fest und machen Angaben zur Aufgabenbearbeitung selbst, den auszuführenden Schritten zur Aufgabenbearbeitung und den zu verarbeitenden Informationen. Darüber hinaus werden Instrumente und Hilfsmittel zur Aufgabenbearbeitung zur Verfügung gestellt, wie z.B. eine Antwortskala, auf welcher die Stärke eines Kausalzusammenhangs angegeben werden soll. Faktoren, deren Einflüsse auf der Aufgabenebene variiert werden können, sind beispielsweise die Instruktion, die Fragestellung an die Probanden, das eingesetzte Paradigma, die Urteilshäufigkeit, der Zeitpunkt der Urteilsabgabe, die Anzahl der Ursachen, die Lernrichtung, die einzuschätzende Kontingenz, die Anzahl der zu beurteilenden Fälle, der Antwortmodus oder die eingesetzte Antwortskala (s.a. Kapitel 4.2, S. 74 ff.). Häufig werden in Untersuchungen aus dem Bereich der Kausalkognition alle oder ein Großteil der möglichen aufgabenbezogenen Einflussfaktoren konstant gehalten.

- (ii) **Medienebene:** Auf der Medienebene wird spezifiziert, wie die Informationen über die zu beurteilenden Ursache-Wirkungs-Zusammenhänge vermittelt werden. Die Beschaffenheit der zu vermittelnden Information selbst wird ebenfalls auf der Medienebene festgelegt. Bei der Informationsübermittlung kann auf unterschiedliche Darstellungsformen, Zeichensysteme und Technologien zurückgegriffen werden. In Verbindung mit der Darstellungsform können die zu beurteilenden Kausalinformationen summarisch oder einzelfallweise in Texten, Bildern, Falllisten, Tabellen, Bildertafeln, Diagrammen, Photos, Animationen, Videos, Simulationen oder durch unmittelbare Beobachtungen von realen Objekten, Gegenständen oder Ereignisfolgen präsentiert werden. Außerdem können die zu beurteilenden Informationen zeitlich versetzt oder zeitgleich dargeboten werden. Darüber hinaus existiert eine Wahlmöglichkeit sowohl hinsichtlich der eingesetzten Sinnesmodalität (z.B. auditiv, visuell), als auch hinsichtlich der Geräte, Gegenstände oder Maschinen, die zur Übertragung von Informationen eingesetzt werden können. Informationen über Ursache-Wirkungs-Zusammenhänge können beispielsweise filmisch, in Animationen, statisch in Bildern oder sprachlich übertragen werden (s.a. Kapitel 4.2, S. 72). Studien zum Einfluss und zu den Auswirkungen medialer Faktoren (zu den Auswirkungen medialer Faktoren und von externalen Repräsentationen im Allgemeinen s. Kapitel 6.2) sind im Bereich der Kausalkognition relativ selten, noch am häufigsten wurde der Einfluss des Faktors Darbietungsart untersucht (s. Kapitel 4.2.1).
- (iii) **Individuenebene:** Auf der Individuenebene werden die Eigenschaften und Charakteristika der Personen erfasst, welche die Stärke von Ursache-Wirkungs-

Zusammenhängen beurteilen sollen. Ganz allgemein ist dabei von Interesse, ob Unterschiede hinsichtlich Alter, Entwicklungsstand, Geschlecht, Sozialstatus, Ausbildungshintergrund, Vorwissen, kognitiven Kapazitäten und Begrenzungen, Intelligenz, Persönlichkeit, Motivation, psychisches Befinden oder auch subjektive Einstellungen und Stimmungen Auswirkungen auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen haben. In den meisten bisherigen Studien zur kausalen Urteilsbildung fehlen Angaben zu möglicherweise relevanten interindividuellen Unterschieden, untersucht wurden bislang lediglich die Einflüsse von Geschlecht, Lebensalter und Depressivität (s. Kapitel 5.1). Worauf die in diesen Studien gefundenen Unterschiede letztendlich beruhen, ist unklar.

- (iv) **Ebene der Kognition:** Der Prozess der kausalen Urteilsbildung scheint sich aus einem netzwerkartigen Zusammenspiel verschiedener kognitiver Teilprozesse zu ergeben. An Hand der Forschungslage (Kapitel 3 und 5.2) ist davon auszugehen, dass dabei neben assoziativen Lernmechanismen und/oder kausalen Schemata und Skripts, in denen Wissen über regel-basierte Ansätze repräsentiert ist, Wahrnehmungs- und Gedächtnisprozesse (zur Aufnahme und Speicherung von Kausalinformationen), sowie Vorwissen in Form von bereichsübergreifenden Kausalmodellen und –überzeugungen oder von bereichsspezifischem Wissen über kausale Mechanismen und Abläufe beteiligt sind. Personenbezogene Faktoren, sowie unterschiedliche Auswirkungen von aufgaben- und medienbezogenen Faktoren beeinflussen oder moderieren das Zusammenspiel dieser kognitiven Teilprozesse.

Das auf Seite 132 abgebildete Schaubild (Abbildung 5) veranschaulicht das Zusammenspiel von Kognition, sowie medien-, aufgaben- und personenbezogenen Faktoren bei der Beurteilung der Stärke von kausalen Zusammenhängen. Im Grunde zeigt das Schaubild in Abbildung 5 das Zusammenspiel zwischen Individuum (kognitive Prozesse und personenbezogene Eigenschaften in den grauen Feldern) und Umgebung (aufgaben- und medienbezogene Faktoren in den weißen Feldern). Auf kognitiver Ebene wird der Informations-Input, bestehend aus Informationen über Ursache-Wirkungs-Zusammenhänge, vom Organismus wahrgenommen und über die Schnittstelle Arbeitsgedächtnis entweder direkt oder durch einen assoziativen Lernmechanismus an einen informationsintegrierenden Mechanismus weitergeleitet und dort verarbeitet. Die Verarbeitung der Informationen über Ursache-Wirkungs-Zusammenhänge durch den informationsintegrierenden Mechanismus beinhaltet im Wesentlichen das Ordnen (Gruppieren) und Aggregieren von Einzelinformationen und schließlich die Urteilsbildung. Die durch das Arbeitsgedächtnis und



**Abbildung 5:** Zusammenspiel von Kognition, sowie medien-, aufgaben- und personenbezogenen Faktoren bei der Beurteilung der Stärke von Kausalzusammenhängen.



den informationsintegrierenden Mechanismus zu verarbeitenden Informationen werden einerseits im Gedächtnis gespeichert, andererseits wird die Arbeit des informationsintegrierenden Mechanismus durch bereits vorhandenes Wissen aus verschiedenen Instanzen des Referenzgedächtnisses wie z.B. kausalen Schemata, Skripts oder Modellen mit beeinflusst, strukturiert und gesteuert. Solche Schemata oder Skripts können beispielsweise Wissen über eine bestimmte regel-basierte Strategie und deren Anwendung beinhalten. Endprodukt der kognitiven Verarbeitung von Informationen über Ursache-Wirkungs-Zusammenhänge ist ein Urteil über die Stärke des Kausalzusammenhangs.

Die Weise der kognitiven Verarbeitung und Beurteilung von Informationen über die Stärke von Ursache-Wirkungs-Zusammenhängen ist an personenbezogene Faktoren und individuelle Voraussetzungen gekoppelt und steht darüber hinaus in Wechselwirkung mit aufgaben- und medienbezogenen Faktoren. Aufgaben- und medienbezogene Faktoren beeinflussen sich ebenfalls gegenseitig und legen die Beschaffenheit, Strukturiertheit und Vermittlung des Informationsinputs fest. Aufgaben- und medienbezogene Faktoren können den Informationsinput so modifizieren, dass unterschiedliche Informationsverarbeitungsprozesse in Gang gesetzt werden.

Personenbezogene Faktoren haben keinen unmittelbaren Einfluss auf die Beschaffenheit des Informationsinputs. Sie moderieren aber dessen Verarbeitung und stehen darüber hinaus in einer Wechselbeziehung mit aufgaben- und medienbezogenen Faktoren. So kommen unterschiedliche Individuen mit bestimmten Aufgaben unterschiedlich leicht oder schnell zurecht und bearbeiten sie unterschiedlich motiviert. Mit unterschiedlichen Medien sind unterschiedliche Individuen unterschiedlich gut vertraut und entnehmen die Informationen aus ihnen mit unterschiedlicher Geschwindigkeit, Genauigkeit und Interesse.

### 7.3 Unter welchen Bedingungen kommen regel-basierte Strategien mit höherer Wahrscheinlichkeit zum Einsatz?

Der in Abbildung 5 aufgezeigte theoretische Rahmen für ein umfassenderes Verständnis der Prozesse bei der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen kann als Ausgangs- und Anregungspunkt für zukünftige empirische Untersuchungen herangezogen werden. Da die Suche nach einer allumfassenden Theorie zur Erklärung der Prozesse bei der Beurteilung der Stärke von kausalen Zusammenhängen als gescheitert angesehen werden kann (Baker et al., 1996; De Houwer & Beckers, 2002), halten wir es nicht für sinnvoll, die verschiedenen Ansätze experimentell gegeneinander auszuspielen. Viel versprechender könnte es sein, Umgebungsbedingungen und individuelle Voraussetzungen zu identifizieren, unter denen ein bestimmter kognitiver Mechanismus oder eine bestimmte Strategie zur Beurteilung der Stärke eines Kausalzusammenhangs ausgewählt wird. Die Fragestellung sollte also nicht lauten, ob der grundlegende Mechanismus bei der Beurteilung der Stärke von Kausalzusammenhängen regel-basiert oder assoziativ ist (Allan, 1993; Baker et al., 1996), sondern eher unter welchen situativen Umständen werden bei welchen Probanden welche Informationsverarbeitungsprozesse in Gang gesetzt oder ausgelöst.

Im experimentellen Teil der vorliegenden Arbeit wird die Frage verfolgt, ob und warum Variationen der medialen Faktoren Darbietungsart, Präsentationsform und Informationsart einen Einfluss darauf haben, ob und wie häufig Probanden regel-basiert vorgehen oder nicht. Die in Kapitel 3 bis 5 vorgestellten empirischen Untersuchungen dienen auch dazu, Bedingungen zu spezifizieren, unter denen der Einsatz regel-basierter Strategien wahrscheinlicher wird.

Auf der *Aufgabenebene* erscheint der Einsatz regel-basierter Strategien wahrscheinlicher zu sein, wenn die Aufgabenstellung eine geringe Komplexität aufweist (Baker et al., 1996; De Houwer & Becker, 2002). Die Komplexität einer Aufgabe ist beispielsweise dann gering, wenn die Probanden die Wirksamkeit von nur einer Ursache im Hinblick auf nur einen Effekt beurteilen müssen (monokausaler Zusammenhang). Von einigen Autoren wird hervorgehoben, dass die vorgegebene Instruktion (Beyth-Marom, 1982; Buehner, Cheng, & Clifford, 2003; Buehner & Cheng, 1997; Collins & Shanks, 2006; Crocker, 1981, 1982; Perales & Shanks, in press; Tenenbaum & Griffiths, 2001; White, 2003b) und die eingesetzte Antwortskala (Allan, 1993) zur Beurteilung des Kausalzusammenhangs einen Einfluss darauf haben, wie und mit Hilfe welcher Strategie ein Kausalzusammenhang beurteilt wird. Regel-basierte Strategien scheinen vor allem dann zum Einsatz zu gelangen, wenn die Probanden

danach gefragt werden, wie stark die fragliche Ursache die Wirkung verursacht (Buehner, Cheng & Clifford, 2003; Collins & Shanks, 2006, Perales & Shanks, in press) und wenn sie ihre Antwort auf einer Skala von 0 bis 100 eintragen müssen (z.B. Buehner, Cheng & Clifford, 2003; Ward & Jenkins, 1965). Buehner, Cheng & Clifford (2003) sehen allerdings in einer solchen Fragestellung eine Mehrdeutigkeit, da die urteilenden Personen die Frage auf zweifache Weise interpretieren können: (i) Welchen Unterschied macht die fragliche Ursache in einer aktuellen Lernumgebung, in welcher alternative Ursachen die Wirkung schon mit einem gewissen Prozentsatz hervorgerufen haben? Oder (ii) Welchen Unterschied macht die fragliche Ursache, wenn alternative Ursachen die Wirkung niemals hervorbringen? Während Interpretation (i) nach Buehner, Cheng & Clifford eher zum Einsatz der  $\Delta P$ -Regel führt, gründen die Probanden ihr Urteil nach Interpretation (ii) eher auf einer Berechnung der kausalen Stärke nach der Power PC-Theorie. Um die Vorhersagen der von ihnen propagierten Power PC-Theorie besser nachweisen zu können, legten Buehner, Cheng & Clifford (2003, Ex 2, 3 und 4) ihren Probanden eine eindeutige kontrafaktische Fragestellung vor, indem sie von den Probanden wissen wollten, in wie vielen von 100 Fällen, in denen eine Wirkung ohne Ursache ausbleibt, die Wirkung eintreten würde, wenn die fragliche Ursache eingeführt werden würde. Die meisten Probanden, die so befragt wurden, legten ihren Kausalurteilen in der Tat eine Berechnung oder Ermittlung der kausalen Stärke nach der Power PC-Theorie zu Grunde.

*Mediale Bedingungen*, die in unseren Augen den Einsatz einer regel-basierten Strategie wahrscheinlicher machen, sind für uns eine summarische Darbietungsart, sowie Präsentationsformen, die (i) die Informationsverarbeitung (Arbeitsgedächtnis) entlasten und strukturieren; die (ii) Arbeitsschritte, die beim Einsatz einer regel-basierten Strategie erforderlich sind, externalisieren; und (iii) die perzeptuelle Inferenzen ermöglichen, indem die Informationsverarbeitung von der Gedächtnis- oder Informationsintegrationsebene auf die Wahrnehmungsebene vor verlagert wird.

Die Studie von Buehner, Cheng & Clifford (2003) belegt nicht nur eindrucksvoll, wie Faktoren auf der Aufgabenebene - hier die verwendete Antwortskala, sowie die Instruktion oder die Wortwahl der Frage nach einem Kausalzusammenhang – den Einsatz einer regel-basierten Strategiewahl oder Beurteilung der Stärke von Kausalzusammenhängen begünstigen können, sondern es lassen sich darin auch *mediale Bedingungen* finden, die die Probanden zum Gebrauch regel-basierter Strategien anregen. Nach unserer Einschätzung bewirkten die von Buehner, Cheng & Clifford (2003) eingesetzten Bildertafeln (s. Abbildung 4, S.78) nicht nur eine kognitive Entlastung, sondern sie strukturieren und steuern kognitives

Verhalten und ermöglichen darüber hinaus einfach zu verarbeitende Inferenzen auf der Wahrnehmungsebene, indem sie zu einem Vergleich oder einer Differenzbildung zwischen Erfolgs- und Basisrate anregen oder diese unmittelbar nahe legen. Wie von uns in Kapitel 6.3, 6.4 und 7.1 heraus gearbeitet wurde, dürften tabellarische Vierfeldertafeln und insbesondere Stapeldiagramme ebenfalls ein hohes Potenzial beinhalten, Probanden den Einsatz regel-basierter Strategien nahe zu legen.

Welche *personenbezogenen Faktoren* den Einsatz regel-basierter Strategien begünstigen, bleibt in den allermeisten Studien unklar, da in der Regel keine näheren Angaben zu den Probanden gemacht werden. Häufig handelt es sich bei den Teilnehmern um Studenten, wobei die Gruppe der Psychologie-Studenten, die in den ersten Semestern statistisch geschult werden, am häufigsten vertreten ist. Untersuchungen wie diejenige von Smedslund (1963), in denen die Experimente mit Krankenschwesternschülerinnen durchgeführt wurden, bilden eine seltene Ausnahme. Wir vermuten, dass häufiger komplexe regel-basierte Strategien eingesetzt werden, wenn die Probanden über eine hohe numerische Verarbeitungskapazität verfügen. In Kapitel 3.4 wurde die Vermutung geäußert, dass der Einsatz regel-basierter Strategien auf dem Abruf von erworbenen kausalen Schemata oder Skripts beruht. Darüber hinaus kann nicht ausgeschlossen werden, dass einzelne Probanden aufgrund ihres Vorwissens, ihrer Lerngeschichte oder aufgrund ihrer rechnerischen und kognitiven Voraussetzungen dazu in der Lage sind, komplexe regel-basierte Strategien wie die  $\Delta P$ -Regel oder die Bestimmung der kausalen Stärke  $p$  nach der Power PC Theorie bei der Beurteilung der Stärke von kausalen Zusammenhängen neu zu generieren. Da Studenten über eine höhere Schulbildung verfügen, kann davon ausgegangen werden, dass sie im Vergleich zu anderen Bevölkerungsgruppen eher über das entsprechende Wissen oder die Kapazität verfügen, komplexe regel-basierte Strategien aus dem Gedächtnis abzurufen oder neu zu generieren.

Auf *kognitiver Ebene* sind bei einer regel-basierten Beurteilung der Stärke von Kausalzusammenhängen im Wesentlichen Wahrnehmungs-, Gedächtnis- und höhere Bewusstseinsprozesse (Vergleiche und Berechnungen anstellen, Schlussfolgerungen ziehen) beteiligt (s. Schaubild auf Seite 129). Der Abruf, die Generierung oder der Einsatz komplexer regel-basierter Strategien wird insbesondere dann erleichtert, wenn die Beanspruchung des Arbeitsgedächtnisses gering ist und wenn es gelingt, höhere Bewusstseinsprozesse auf die Wahrnehmungsebene zu verlagern (Lohse, 1997; Wickens & Carswell, 1995). In Kapitel 4.1 und 5.2 haben wir gesehen, dass die Beanspruchung des Gedächtnisses von der Art der Darbietung (Arkes & Harkness, 1983; Baker et al., 1996; Kao & Wasserman, 1993; Shaklee & Mims, 1982) und der Präsentation der Daten (Baddeley & Logie, 1999; Goolkasian &

Foos, 2002; Smith & Magee, 1980) abhängt. Das Ausmaß der Gedächtnisbeanspruchung wird nicht nur durch mediale, sondern auch noch durch aufgaben- und personenbezogene Faktoren beeinflusst. Bei einer passenden und wenig komplexen Aufgabenstellung, sowie einer summarischen Darbietung von Ursache-Wirkungs-Informationen in einem übersichtlich gestalteten und relevante Vergleichs- und Schlussfolgerungsprozesse unterstützenden Medium dürfte insbesondere bei Individuen, die über eine hohe Gedächtnis- und numerische Verarbeitungskapazität verfügen, die Gedächtnisbeanspruchung gering ausfallen. Basiert die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen auf dem Abruf einer regelbasierten Strategie aus dem Gedächtnis, dann kann dies natürlich nur unter der Voraussetzung geschehen, dass die entsprechende Strategie im Gedächtnis abgespeichert oder das dafür notwendige Wissen mental repräsentiert ist bzw. zur Verfügung gestellt werden kann. Wenn die entsprechende Strategie nicht im Gedächtnis repräsentiert ist, dann muss sie neu generiert werden. Ob das geschehen kann, hängt davon ab, ob neben den entsprechenden Lernvoraussetzungen (Verständnis der dargestellten Kausalzusammenhänge, Aufgabenverständnis, Kenntnisse im Umgang mit den vorgegebenen Medien und einer der Aufgabe angemessenen Informationsentnahme aus ihnen), auch kognitive (Wahrnehmungs-, Gedächtnis- und Schlussfolgerungskapazität) und motivationale Voraussetzungen vorliegen.

## 7.4 Ausblick auf Fragestellungen und Untersuchungsziele

Im Rahmen der vorliegenden Arbeit sollten in drei Experimenten die folgenden Fragestellungen untersucht werden:

1. Welchen Einfluss haben die Faktoren Präsentationsform, Informationsart und Darbietungsart auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen und auf die Strategiewahl der Probanden?
2. Unter welchen medialen Rahmenbedingungen gehen Probanden exakt regel-basiert vor?
3. Unterscheiden sich Individuen hinsichtlich Urteilsverhalten und Strategiewahl unter verschiedenen medialen Bedingungen voneinander?
4. Welchen Einfluss haben unterschiedliche individuelle Voraussetzungen im Hinblick auf die numerische Verarbeitungskapazität auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen?
5. In welchem Ausmaß lassen sich unterschiedliche Auswirkungen des Faktors Darbietungsart auf eine unterschiedliche Gedächtnisbeanspruchung zurückführen?

Der Einfluss aufgabenbezogener Faktoren sollte im Rahmen dieser Arbeit nicht untersucht werden, deshalb wurde der Einfluss von Faktoren wie Instruktion, Fragestellung, Antwortskala oder Anzahl der Ursachen konstant gehalten. In allen drei Experimenten sollten die Probanden auf einer Skala von 0 bis 100 die Stärke von monokausalen, probabilistischen Ursache-Wirkungs-Zusammenhängen beurteilen. Ganz allgemein geht es in dieser Arbeit darum, medienbezogene Einflüsse zu identifizieren, die den Gebrauch von komplexen regel-basierten Strategien wie der  $\Delta P$ -Regel oder der Bestimmung der kausalen Stärke  $p$  nach der Power PC Theorie wahrscheinlicher machen. Da individuelle Unterschiede einen nicht zu vernachlässigenden Einfluss auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen haben, haben wir eine individualpsychologische Herangehensweise gewählt. Darunter verstehen wir, dass jedes einzelne Urteil oder Wirksamkeitsrating über die Stärke eines Kausalzusammenhangs von uns daraufhin untersucht wurde, ob es mit einer regel-basierten Strategie exakt übereinstimmt oder nicht.

Teil 2

## **Empirische Untersuchungen**





## 8 Experiment 1: Einfluss von Präsentationsform und Informationsart

In Experiment 1 sollte untersucht werden, welchen Einfluss unterschiedliche Präsentationsformen und Informationsarten auf die Beurteilung der Stärke kausaler Zusammenhänge und die Einschätzung der Wirksamkeit einer Ursache haben. Im Besonderen sollte der Einfluss dieser Faktoren auf die *Strategiewahl* und die *Anzahl von Lösungen* pro Versuchsperson untersucht werden, *die exakt mit den Vorhersagen der  $\Delta P$ -Regel oder der Power PC-Theorie übereinstimmen*. Ausgangspunkt dieser besonderen Fragestellung ist die These, dass der Gedächtnisabruf oder die mentale Konstruktion regel-basierter Strategien, sowie deren Anwendung durch eine unterschiedliche Präsentation der für die Beurteilung relevanten kausalen Informationen entweder vereinfacht oder erschwert werden kann. Um den Einfluss der unabhängigen Variablen auf die mentale Beanspruchung während der Aufgabenbearbeitung einschätzen zu können, wurde neben der *Anzahl regel-konformer Lösungen* auch noch die *Zeit* gemessen, die die Probanden zur Aufgabenbearbeitung benötigten. Durch das Erfassen von *Bearbeitungszeit* und *Anzahl regel-konformer Lösungen* eröffnet sich darüber hinaus die Möglichkeit, die Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) in Verbindung mit der Beurteilung der Stärke von Ursache-Wirkungszusammenhängen zu operationalisieren.

Der Einfluss des Faktors *Darbietungsart* wurde in Experiment 1 konstant gehalten, die Darbietung der zu beurteilenden Ursache-Wirkung-Zusammenhänge erfolgte ausschließlich in summarischer Form. In Kapitel 4.2.1 konnte herausgearbeitet werden, dass eine Reihe von Untersuchungen (Arkes & Harkness, 1983; Buehner et al., 2003; Kao & Wasserman, 1993; Lober & Shanks, 2000; Ward & Jenkins, 1965) dafür sprechen, dass die Probanden bei summarischer Darbietung eher regel-basierte Strategien einsetzen und genauere Kontingenzurteile bilden als bei einzelfallweiser Darbietung. Dennoch gehen wir von der Annahme aus, dass die Faktoren *Präsentationsform* und *Informationsart* auch bei einer einheitlichen summarischen Informationsdarbietung unterschiedliche Auswirkungen auf die Beurteilung der Stärke von Ursache-Wirkung-Zusammenhängen und die dabei zum Einsatz kommenden Strategien haben können (s. Kapitel 4.2.2 und 4.2.3). Auf eine einzelfallweise Informationsdarbietung wurde in Experiment 1 verzichtet, da unter solchen Umständen die Komplexität der Aufgabenstellung und die Gedächtnisbeanspruchung ansteigen und einen

Einsatz regel-basierter Strategien unwahrscheinlicher machen. Dadurch könnten mögliche Auswirkungen der Faktoren *Präsentationsform* und *Informationsart* überlagert werden.

Konkret wurden den Probanden aus Experiment 1 die summarisch dargebotenen Kausalinformationen entweder als Tabelle mit Häufigkeiten, als Tabelle mit Wahrscheinlichkeiten, als Stapeldiagramm mit Häufigkeiten oder als Stapeldiagramm mit Wahrscheinlichkeiten präsentiert. Diesen vier Experimentalbedingungen, die sich aus einer Kreuzung der zu untersuchenden Faktoren *Präsentationsform* und *Informationsart* ergeben hat, wurde zusätzlich als eine weitere summarische Kontrollbedingung eine Protokollliste gegenübergestellt. Es sollte herausgefunden werden, welchen Einfluss die verschiedenen Untersuchungsbedingungen auf die verschiedenen abhängigen Variablen wie zum Beispiel die Strategiewahl oder die mentale Beanspruchung bei der Aufgabenbearbeitung haben. Nicht verwendet wurde das von Buehner, Cheng und Clifford (Experiment 2, 2003) entwickelte Stimulusmaterial (siehe Abbildung 4). Obwohl sich die übersichtlich gestalteten Bildertafeln in Experiment 2 von Buehner, Cheng & Clifford (2003) als überaus geeignet erwiesen, die Probanden zum Gebrauch regel-basierter Strategien anzuregen, haben wir auf ihren Einsatz verzichtet, weil damit keine Variation des Faktors *Informationsart* möglich ist. Zudem glauben wir, dass die Bildertafeln von Buehner, Cheng & Clifford (2003) eine geringere ökologische Validität aufweisen, als die von uns in Experiment 1 eingesetzten Tabellen und Stapeldiagramme.

Neben dem Einfluss unterschiedlicher summarischer Präsentationsformen und Informationsarten auf Strategiewahl und mentale Beanspruchung bei der Aufgabenbearbeitung sollte auf der Individuenebene untersucht werden, ob unterschiedliche mathematische Fertigkeiten der Probanden einen Einfluss auf den Einsatz regel-basierter Strategien oder die Urteilsgenauigkeit haben. Um diese Annahme zu testen, wurden aus einer studentischen Population zwei gleichgroße Probandengruppen gebildet, die sich hinsichtlich ihrer Studiaausrichtung unterschieden. In Anlehnung an Shaklee & Hall (1983) wurde vermutet, dass Probanden aus mathematisch-naturwissenschaftlich ausgerichteten Studiengängen eher regel-basierte oder komplexere Strategien generieren oder aus dem Gedächtnis abrufen als Probanden aus geisteswissenschaftlichen Studiengängen.

## 8.1 Untersuchungsdesign

### Versuchsplan und unabhängige Variablen

Es sollte der Einfluss der beiden zweistufigen Between-Faktoren *Präsentationsform* und *Informationsart* untersucht werden. Hinsichtlich der *Präsentationsform* wurden tabellarische Vierfeldertafeln mit Stapeldiagrammen verglichen, hinsichtlich der *Informationsart* Wahrscheinlichkeiten mit Häufigkeiten. Aus der Kreuzung der beiden Faktoren *Präsentationsform* und *Informationsart* ergibt sich ein 2 x 2 - Design mit den vier Bedingungen *Tabelle mit Häufigkeiten*, *Tabelle mit Wahrscheinlichkeiten*, *Stapeldiagramm mit Häufigkeiten* und *Stapeldiagramm mit Wahrscheinlichkeiten*, welche in Abbildung 6 gezeigt werden. Bei der ebenfalls in Abbildung 6 abgebildeten Kontrollbedingung *Protokollliste* handelt es sich um eine einfache Aneinanderreihung einzelner Beobachtungen bezüglich des Regenverhaltens von Wolken nach Impfung oder ausbleibender Impfung.

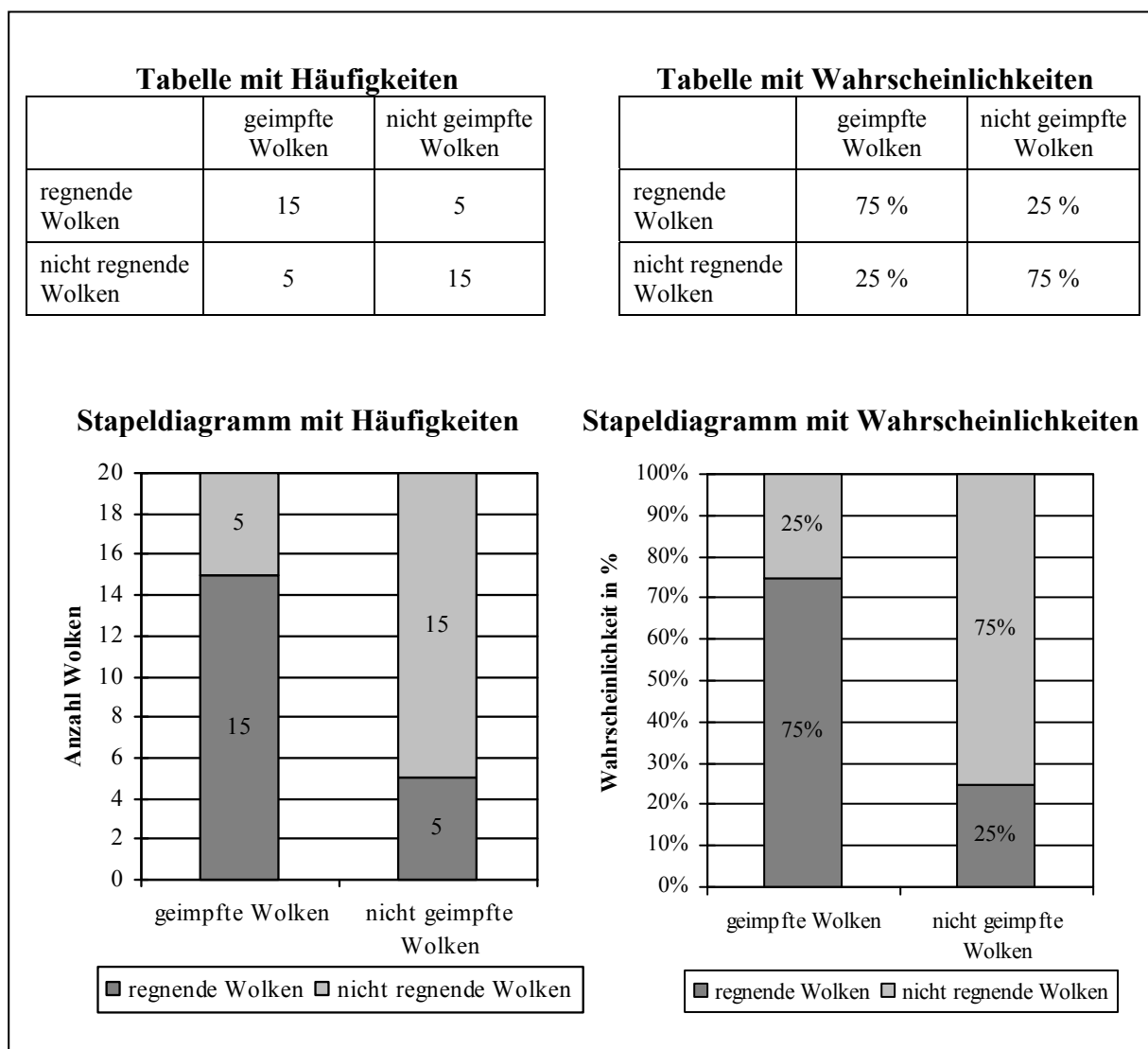


Abbildung 6: Experiment 1: 2 x 2-Design

### Abhängige Variablen

Eine wesentliche Fragestellung dieser Untersuchung war es, herauszufinden unter welchen Bedingungen Versuchspersonen von regel-basierten Strategien wie der  $\Delta P$ -Regel oder einer Berechnung der kausalen Stärke entsprechend den Vorgaben der Power PC Theorie Gebrauch machen.

#### Abhängige Variablen zur Messung der Strategiewahl:

*Anzahl regel-konformer Lösungen pro Versuchsperson:* Für jeden Probanden wurde festgehalten, wie viele regel-konforme Lösungen hervorgebracht wurden. Da jeder Proband vier Aufgaben mit unterschiedlichen Kontingenzen zu lösen hatte, konnten maximal vier regel-konforme Lösungen erreicht werden. Um herauszufinden, ob ein Proband regel-konform vorgegangen ist, wurden die Wirksamkeitsratings für jede einzelne Aufgabe daraufhin untersucht, ob sie exakt mit den

Vorhersagen der  $\Delta P$ -Regel oder der Power PC-Theorie übereinstimmten. Zudem wurden die Probanden unmittelbar im Anschluss an die Aufgabenbearbeitung zu ihrer Vorgehensweise bei der Bestimmung der Wirksamkeiten befragt (s. Anhang A1.3). Durch die Nachbefragung sollte die Kategorisierung einer Lösung als regel-konform ausgeschlossen werden, wenn der Teilnehmer zufällig, durch Raten oder mittels einer anderen Strategie als der  $\Delta P$ -Regel oder der Bestimmung der kausalen Stärke entsprechend der Power PC Theorie auf eine regel-konforme Lösung gekommen war.

*Strategiewahl:* Für jede Person wurde anhand der Wirksamkeitsratings und der Nachbefragung ermittelt, welche Strategie oder welche Strategien sie bei der Aufgabenbearbeitung eingesetzt hat. Alle Probanden wurden dahingehend kategorisiert, ob es sich um Anwender einer durchgängigen Strategie handelte oder um Strategienwechsler. Weiter wurde kategorisiert, von welcher Strategie genau ein Proband Gebrauch gemacht hat

Wolke	Geimpft ?	Zustand der Wolke am Ende des Beobachtungszeitraum
1	Ja	Regen
2	Nein	Kein Regen
3	Ja	Kein Regen
4	Nein	Regen
5	Ja	Regen
6	Nein	Kein Regen
7	Ja	Regen
8	Nein	Kein Regen
9	Ja	Kein Regen
10	Nein	Regen
11	Ja	Regen
12	Nein	Kein Regen
13	Ja	Regen
14	Nein	Kein Regen
15	Ja	Regen
16	Nein	Kein Regen
17	Ja	Regen
18	Nein	Kein Regen
19	Ja	Regen
20	Nein	Kein Regen
21	Ja	Kein Regen
22	Nein	Kein Regen
23	Ja	Regen
24	Nein	Regen
25	Ja	Regen
26	Nein	Kein Regen
27	Ja	Regen
28	Nein	Kein Regen
29	Ja	Regen
30	Nein	Regen
31	Ja	Kein Regen
32	Nein	Kein Regen
33	Ja	Regen
34	Nein	Kein Regen
35	Ja	Regen
36	Nein	Regen
37	Ja	Kein Regen
38	Nein	Kein Regen
39	Ja	Regen
40	Nein	Kein Regen

**Abbildung 7:** Experiment 1: Kontrollbedingung Protokollliste

(z.B. Power PC-Theorie,  $\Delta P$ -Regel, ab-Heuristik, unklare/intuitive Heuristik). Handelte es sich um Strategiewechsler, so wurde kategorisiert, um welche Art von Strategiewechsler es sich dabei handelte (z.B. unklar zu  $\Delta P$ -Regel). Probanden, die entweder die  $\Delta P$ -Regel oder die Power PC-Theorie eingesetzt haben, wurden auch zur übergeordneten Kategorie von Anwendern einer regel-basierten Strategie zusammengefasst. Während sich der Begriff "regel-basiert" auf eine oder mehrere von einer Versuchsperson gewählte Strategien bezieht, nimmt der Begriff "regel-konform" Bezug auf die Anzahl von Lösungen pro Versuchsperson, die exakt mit den Vorhersagen der  $\Delta P$ -Regel oder der Power PC-Theorie übereinstimmen. In Verbindung mit der Strategiewahl wurden qualitative Daten erhoben, die nicht varianzanalytisch, sondern nur mittels Häufigkeitsanalysen ausgewertet werden konnten.

Indem für jeden Probanden möglichst genau erfasst wurde, von welcher oder welchen Strategie/n er bei der Beurteilung der Stärke von Kausalzusammenhängen Gebrauch gemacht hat, sollten Aufschlüsse darüber gewonnen werden, welcher kognitive Verarbeitungsmechanismus oder welche kognitiven Verarbeitungsmechanismen in Abhängigkeit von den unterschiedlichen Versuchsbedingungen aktiviert wird oder werden.

#### Abhängige Variablen zur Messung der mentalen Beanspruchung bei der Aufgabenbearbeitung

*Mittlere Komplexität der Lösungen:* Die rechnerische Komplexität der gefundenen Lösungen misst die mentale Beanspruchung bei der Aufgabenbearbeitung nicht direkt. Sie kann aber als ein Indikator für die mentale Beanspruchung während der Aufgabenbearbeitung angesehen werden. Wenn die mentale Beanspruchung während der Aufgabenbearbeitung hoch ist, dann ist in Verbindung mit der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) zu erwarten, dass die produzierten Lösungen rechnerisch einfacher werden. Mit zunehmender Komplexität der Aufgabe sinkt also die rechnerische Komplexität der gefundenen Lösung.

Die rechnerische Komplexität der gefundenen Lösung wurde ermittelt, indem die Anzahl vorkommender Rechenzeichen (Zahlen und Operatoren, keine Klammern) für die von den Probanden eingesetzten Gleichungen addiert und pro Versuchsperson gemittelt wurden. Die Bestimmung der mittleren rechnerischen Komplexität einer gefundenen Lösung soll an einem fiktiven Beispiel veranschaulicht werden. Angenommen die Versuchsperson X verwendet bei der ersten Aufgabe keine Strategie (intuitives Schätzen), bei der zweiten Aufgabe die ab-Regel, bei der dritten Aufgabe die  $\Delta P$ -Regel und bestimmt schließlich bei der vierten Aufgabe die kausale Stärke einer Ursache nach der Power PC-Theorie, so beträgt die mittlere rechnerische Komplexität der gefundenen Lösung 6.75, wie die folgende Rechnung zeigt:

Aufgabe 1:	Komplexität = 0,	da keine Formel eingesetzt wurde.
Aufgabe 2:	Komplexität = 4,	da sich die Formel $a / (a + b)$ zur Bestimmung der Erfolgsrate aus zwei unterschiedlichen Variablenwerten (a, b) und zwei Operatoren (/ , +) zusammensetzt.
Aufgabe 3:	Komplexität = 9,	da sich die Formel $a / (a + b) - c / (c + d)$ zur Bestimmung von $\Delta P$ aus vier unterschiedlichen Variablenwerten (a, b, c, d) und fünf Operatoren (/ , + , - , / , +) zusammensetzt.
Aufgabe 4:	Komplexität = 14,	da sich die Formel $(a / (a + b) - c / (c + d)) / (1 - c / (c + d))$ zur Bestimmung der kausalen Stärke p aus vier unterschiedlichen Variablenwerten (a, b, c, d), einer Konstante (1) und 9 Operatoren (/ , + , - , / , + , / , - , / , +) zusammensetzt.
Gesamt:	Komplexität = 27,	woraus sich eine gemittelte rechnerische Komplexität der gefundenen Lösungen von 6.75 ( $27 / 4$ ) für die Versuchsperson X ergibt.

Formal müssten alle errechneten Werte durch eine Multiplikation mit 100 in Prozentwerte umgewandelt werden, damit sie sinnvoll auf die von 0 bis 100 reichende Antwortskala eingetragen werden können. Da eine Umwandlung in Prozentwerte bei allen rechnerischen Bestimmungen der Wirksamkeit zum Einsatz käme und insofern eine Konstante darstellt, haben wir auf sie bei der Ermittlung der mittleren rechnerischen Komplexität einer gefundenen Lösung verzichtet.

Anders als bei der abhängigen Variable *Anzahl regel-konformer Lösungen*, wo pro Versuchsperson aufaddiert wurde, bei wie vielen Aufgaben sie entweder von der  $\Delta P$ -Regel oder von der Power PC-Theorie Gebrauch gemacht hat, fällt bei der abhängigen Variablen *mittlere rechnerische Komplexität der Lösungen* der Einsatz unterschiedlicher regel-basierter Strategien unterschiedlich ins Gewicht. Zudem findet bei der Bestimmung der mittleren Komplexität auch der Einsatz anderer Strategien und Heuristiken (wie zum Beispiel der ab-Regel) als  $\Delta P$  oder der Power PC-Theorie Berücksichtigung.

Unsere Untersuchungen haben ergeben, dass die beiden abhängigen Variablen *Anzahl regel-konformer Lösungen* und *mittlere Komplexität der Lösungen* in allen drei Experimenten hochsignifikant miteinander korrelieren (Experiment 1  $r = .86$ ; Experiment 2  $r = .91$ ; Experiment 3  $r = .91$ ;  $p$  jeweils  $< .001$ ). Die Erklärung hierfür ist einfach, da regel-konforme *Lösungen* den Einsatz hochkomplexer mathematischer Strategien erforderlich machen. Weil für die beiden abhängigen Variablen *Anzahl regel-konformer Lösungen* und *mittlere Komplexität der Lösungen* (i) gleiche Vorhersagen gemacht wurden, (ii) annähernd identische Ergebnisse bei den durchgeführten Varianz- und Kovarianzanalysen zustande kamen und (iii) beide Variablen hoch signifikant miteinander korrelierten wurde in der Darstellung der Ergebnisse auf die abhängige Variable *mittlere Komplexität* verzichtet.

In theoretischer Hinsicht sind die hohen Korrelationen der beiden abhängigen Variablen *Anzahl regel-konformer Lösungen* und *mittlere Komplexität der Lösungen* insofern interessant, weil dadurch die beiden abhängigen Variablen *Anzahl regel-konformer Lösungen* und *mittlere Komplexität der Lösungen* als Maße für die mentale Beanspruchung während der Aufgabenbearbeitung angesehen werden können. Wir gehen dabei von der Annahme aus, dass beide Variablen Rückschlüsse auf kognitive Prozesse in Verbindung mit der Urteilsbildung und der Integration von Information aus dem Arbeits- und Referenzgedächtnis zulassen. Wie eine hohe rechnerische Komplexität der Lösung, so weist auch eine hohe Anzahl regel-konformer Lösungen auf eine geringe mentale Beanspruchung während der Aufgabenbearbeitung hin.

Legt man die Prämisse zu Grunde, dass Prozesse in Verbindung mit der Beurteilung der Stärke kausaler Zusammenhänge tatsächlich mit Kosten-Nutzen-Abwägungen nach Kleinmuntz & Schkade (1993) einhergehen, dann kann der *Nutzen* oder die Güte einer gefundenen Lösung und der ihr zugrunde liegenden Strategie danach beurteilt werden, ob die gefundene Lösung regel-konform bzw. rechnerisch komplex aufgebaut ist und auf einer regel-basierten Strategie beruht. In Verbindung mit einer Operationalisierung der Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) bieten sich die beiden Maße *Anzahl regel-konformer Lösungen* und *mittlere Komplexität der Lösungen* also als Quantifizierungen für den Nutzen oder die Güte der gefundenen Lösung an. Wir gehen dabei von der Annahme aus, dass Probanden, die die Stärke eines Kausalzusammenhangs beurteilen sollen, darum bemüht sind, regel-basierte Strategien einzusetzen und regel-konforme Lösungen zu generieren, sofern sie über das dafür notwendige Wissen verfügen oder in der Lage sind, dieses Wissen abzurufen oder zu erwerben (individuelle Voraussetzungen). Ist der dafür erforderliche Lern- oder Arbeitsaufwand aufgrund der Konstellation von aufgaben-, medien- und personenbezogenen Faktoren zu groß (hohe Kosten infolge starker mentaler Beanspruchung), dann sinken die Anzahl regel-konformer Lösungen und die rechnerische Komplexität der gefundenen Lösung (der Nutzen); weil die Probanden entweder auf einfachere Strategien zurückgreifen oder weil sie beim Gebrauch komplexer regel-basierter Strategien zu Fehleinschätzungen und Urteilsverzerrungen gelangen.

Zur Annahme, dass die urteilenden Probanden, darum bemüht sind, regel-basierte Strategien einzusetzen und durch eine zu hohe Gedächtnisbeanspruchung daran gehindert werden, passt eine Erklärung von Buehner, Cheng & Clifford (2003) zu Abweichungen von der Power PC-Theorie: Sie führen die in ihren Experimenten zu beobachtenden Abweichungen der von ihnen favorisierten Power PC Theorie auf eine fehlerhafte Wahrnehmung von

Nullkontingenzen durch einige Probanden zurück. Das Zustandekommen dieser Fehlwahrnehmungen erklären Buehner, Cheng & Clifford (2003) unter anderem durch Gedächtnisbeschränkungen. Werden diese Gedächtnisbeschränkungen beispielsweise durch übersichtlich präsentierte und summarisch dargebotene Informationen über Ursache-Wirkungs-Zusammenhänge aufgehoben, dann stimmen die Ergebnisse von Buehner, Cheng & Clifford (2003) ziemlich genau mit den Vorhersagen der Power PC Theorie überein.

*Bearbeitungszeit*: Neben der *Anzahl regel-konformer Lösungen* und der *rechnerischen Komplexität der gefundenen Lösungen*, kann auch die Zeit, die die Probanden insgesamt zur Ermittlung der Stärke von Ursache-Wirkungs-Zusammenhängen benötigen, als ein Indikator für die mentale Beanspruchung während der Aufgabenbearbeitung angesehen werden. Wir gehen davon aus, dass sich bei einer hohen mentalen Beanspruchung längere Bearbeitungszeiten zeigen werden, als bei geringer mentaler Beanspruchung. In Verbindung mit der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) kann die *Bearbeitungszeit* als ein Maß verstanden werden, welches die *Kosten* oder den Aufwand repräsentieren, der erbracht werden muss, um zu einem Urteil über die Stärke eines Kausalzusammenhangs zu gelangen.

Wenn davon gesprochen wird, dass sowohl *Bearbeitungszeit*, als auch *mittlere Komplexität der Lösungen / Anzahl regel-konformer Lösungen* die mentale Beanspruchung während der Aufgabenbearbeitung messen, dann bedeutet das nicht, dass mit beiden Maßen dasselbe gemessen wird. Tatsächlich erfassen die Variablen *mittlere Komplexität der Lösungen / Anzahl regel-konformer Lösungen* und *Bearbeitungszeit* die mentale Beanspruchung während der Aufgabenbearbeitung nicht direkt. Sie liefern aber Hinweise auf unterschiedliche Aspekte mentaler Beanspruchung: Während die *mittlere Komplexität der Lösungen* oder auch die *Anzahl regel-konformer Lösungen* eher über die Güte, oder genauer über die Schwierigkeit oder Komplexität des am Ende des Bearbeitungsprozesses auf der Ebene der Informationsintegration abgegebenen Kausalurteils Auskunft geben, informiert die *Bearbeitungszeit* über den zeitlichen Aufwand bei der Aufgabenbearbeitung. Dabei setzt sich die zur Aufgabenbearbeitung benötigte Zeit nicht nur aus der Dauer für die Prozesse der Urteilsbildung selbst zusammen, sondern sie kann auch zeitaufwendige Handlungsroutinen mit einschließen, welche möglicherweise bei der Datenaggregation und der Datenzuordnung benötigt werden. Die *Bearbeitungszeit* umfasst also nicht nur die Dauer von kognitiven Prozessen in Verbindung mit der Urteilsbildung und der Integration gedächtnisbasierter Informationen, sondern sie schließt die Dauer peripherer Prozesse in Zusammenhang mit Informationsaufnahme und Wahrnehmung mit ein.



Abhängige Variablen zur Messung der Verarbeitungseffizienz:

In der Messung der Verarbeitungseffizienz sehen wir eine Möglichkeit, die Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) in Zusammenhang mit dem Einfluss der beiden Faktoren *Präsentationsform* und *Informationsart* auf die Strategiewahl bei der Beurteilung der Stärke von kausalen Zusammenhängen zu untersuchen. Unterschiedliche Informationsanordnungen bedingen eine unterschiedliche mentale Beanspruchung, sie beeinflussen den zeitlichen Aufwand und die Komplexität der eingesetzten Strategie und damit auch die Komplexität der generierten Lösung und die Anzahl regel-konformer Lösungen. Nach der Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) nehmen wir an, dass die Komplexität der generierten Lösung und damit auch die Anzahl von Lösungen, die im Sinne der  $\Delta P$ -Regel oder einer Berechnung der kausalen Stärke  $p$  nach der Power-PC Theorie regel-konform sind, sinken, wenn die mentale Beanspruchung oder der Aufwand für die Beurteilung der Stärke von kausalen Zusammenhängen steigen. Durch die Berechnung der Verarbeitungseffizienz können der zu erbringende Aufwand (Kosten) und der Nutzen zueinander ins Verhältnis gesetzt werden, indem die Komplexität der generierten Lösung und die Anzahl regel-konformer Lösungen (Nutzen) durch die Bearbeitungszeit (Kosten) geteilt werden. Auf diese Weise können Kosten und Nutzen miteinander verrechnet und verschiedene Präsentationsformen hinsichtlich ihrer Verarbeitungseffizienz miteinander verglichen werden. Um die Verarbeitungseffizienz, d.h. das Kosten-Nutzen Verhältnis quantifizieren zu können, wurden zwei Maße erhoben:

(i) *Anzahl regel-konformer Lösungen / Bearbeitungszeit*: Bei diesem Effizienzmaß wird für jeden Probanden die Anzahl regel-konformer Lösungen zur Bearbeitungszeit ins Verhältnis gesetzt. Die diesbezügliche Effizienz erreicht dann hohe Werte, wenn die Probanden für die Aufgabenbearbeitung wenig Zeit benötigen und dabei eine hohe Anzahl regel-konformer Lösungen hervorbringen.

(ii) *Mittlere rechnerische Komplexität der gefundenen Lösungen / Bearbeitungszeit*: Bei diesem Effizienzmaß wird für jeden Probanden die mittlere rechnerische Komplexität zur Bearbeitungszeit ins Verhältnis gesetzt. Die diesbezügliche Effizienz erreicht dann hohe Werte, wenn die Probanden für die Aufgabenbearbeitung wenig Zeit benötigen und dabei von rechnerisch möglichst komplexen Strategien Gebrauch machen. Es ist dabei nicht notwendigerweise erforderlich, dass die Probanden dabei  $\Delta P$  oder die Power-PC Theorie einsetzen, sie können - um hohe Effizienzmaße zu erreichen - auch von relativ einfachen Strategien (wie z.B. der ab-Regel) Gebrauch machen und fehlende Komplexität durch eine schnelle Bearbeitungszeit kompensieren.

Ebenso wie die beiden abhängigen Variablen *Anzahl regel-konformer Lösungen* und *mittlere Komplexität der Lösungen*, so korrelierten auch die beiden Effizienzmaße hochsignifikant miteinander (Experiment 1  $r = .92$ ; Experiment 2  $r = .95$ ; Experiment 3  $r = .94$ ;  $p$  jeweils  $< .001$ ) und führten zu annähernd identischen varianz- und kovarianzanalytischen Ergebnissen. Aus diesem Grund wurde in der Darstellung der Ergebnisse auf das zweite Effizienzmaß *mittlere Komplexität der Lösungen / Bearbeitungszeit* verzichtet.

## 8.2 Forschungshypothesen

### 8.2.1 Einfluss des Faktors Präsentationsform

#### *Einfluss der Präsentationsform auf die Anzahl regel-konformer Lösungen*

Es wird erwartet, dass Personen, die die kausalen Informationen in einem Stapeldiagramm dargeboten bekommen, signifikant häufiger regel-konforme Lösungen hervorbringen als Personen, denen die kausalen Informationen tabellarisch dargeboten werden. In einem Stapeldiagramm werden geimpfte und nicht geimpfte Wolken durch eine räumliche Gegenüberstellung gesondert gruppiert. Die mentale Gruppierung in geimpfte und ungeimpfte Wolken ist Voraussetzung dafür, dass Erfolgs- und Basisraten rechnerisch ermittelt werden können. Es ist zu vermuten, dass durch die Gegenüberstellung von geimpften und nicht geimpften Wolken eine externe Strukturierungshilfe zur individuellen Informationsverarbeitung angeboten werden kann, die bei der Konstruktion oder beim Abruf eines mentalen Modells zur Ermittlung der Wirksamkeit einer Ursache hilfreich ist. Zusätzlich wird im Gegensatz zur Vierfeldertafel im Stapeldiagramm der Unterschied zwischen regnenden und nicht regnenden Wolken durch eine unterschiedliche Farbgebung hervorgehoben. Wir vermuten, dass die Probanden Erfolgs- und Basisrate durch die Gruppierung in geimpfte und nicht geimpfte Wolken, so wie durch die unterschiedliche Farbgebung einfacher entsprechend der  $\Delta P$ -Regel oder der Power PC Theorie zueinander in Beziehung setzen und miteinander vergleichen können. Dies sollte dazu führen, dass bei der Präsentation eines Stapeldiagramms häufiger regel-konforme Lösungen hervorgebracht werden als bei der Präsentation einer tabellarischen Vierfeldertafel. In einer Vierfeldertafel wird eine Differenzbildung zwischen Erfolgs- und Basisrate trotz einer Gruppierung der summarischen Informationen entsprechend den vorgegebenen Zellkategorien nicht gesondert nahe gelegt, da die vier Zellinhalte in gleichen Abständen zueinander stehen und nicht farblich hervorgehoben werden. In Kurzschreibweise ergibt sich folgende Hypothese:

H1a: # regel-konformer *Lösungen*<sub>Diagramm</sub> > # regel-konformer *Lösungen*<sub>Tabelle</sub>

Auf der Basis der theoretischen Überlegungen zu den Auswirkungen externaler Repräsentationen (Larkin & Simon, 1987; Lohse, 1997; Wickens & Carswell, 1995; Zhang, 1997, 2000; Zhang & Norman, 1994) und auf der Basis empirischer Befunde zum Vergleich zwischen Diagrammen und Tabellen (Jacobs, 1994, 1999) kann ebenfalls angenommen werden, dass das Herstellen einer Beziehung zwischen Erfolgs- und Basisrate im Stapeldiagramm leichter gemacht wird als in einer Tabelle. Durch die räumliche Gegenüberstellung von Erfolgs- und Basisrate im Stapeldiagramm werden komplexe mentale Operationen wie die Differenzbildung zwischen Erfolgs- und Basisrate bei der Berechnung des Kontingenzmaßes  $\Delta P$  oder das Ins-Verhältnis-Setzen der Differenz von Erfolgs- und Basisrate zur Basisrate bei der Berechnung der kausalen Stärke  $p$  nach der Power PC-Theorie visuell veranschaulicht, unmittelbar nahe gelegt und partiell auf die Wahrnehmungsebene verlagert. Durch die Verlagerung auf die Wahrnehmungsebene werden die kognitive Informationsverarbeitung und insbesondere das Arbeitsgedächtnis entlastet (Fiedler, 2000; Wickens & Carswell, 1995), sowie perzeptuelle Inferenzen (Larkin & Simon, 1987) ermöglicht. Aus motivationaler Sicht führt eine Entlastung von Informationsverarbeitung und Arbeitsgedächtnis nach der Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) dazu, dass die Probanden komplexere und genauere Strategien wie zum Beispiel die  $\Delta P$ -Regel oder eine Berechnung der kausalen Stärke  $p$  nach der Power-PC Theorie einsetzen, weil sich der Aufwand für den Einsatz einer komplexen, aber genauen Strategie reduziert hat. Zudem kann erwartet werden, dass sich durch die Entlastung von Informationsverarbeitung und Arbeitsgedächtnis die Anzahl verzerrter oder fehlerhafter Kausalurteile reduziert (Shaklee & Mims, 1982; Fiedler, 2000).

#### *Einfluss der Präsentationsform auf die Bearbeitungszeit*

Es wird erwartet, dass die mittlere Bearbeitungszeit bei Personen, denen tabellarische Vierfeldertafeln präsentiert werden, höher als bei Personen ist, die mit Stapeldiagrammen arbeiten, da vermutet wird, dass den Personen, die Stapeldiagramme bearbeiten, durch die Aufbereitung der Informationen über Ursache-Wirkungszusammenhänge und durch die kognitive Entlastung mehr Verarbeitungsschritte abgenommen werden. Dadurch reduziert sich der zeitliche Aufwand (die Kosten, s.a. Kleinmuntz & Schkade, 1993) bei der Aufgabenbearbeitung. Im Vergleich zur tabellarischen Vierfeldertafel gehen wir davon aus, dass die Informationsverarbeitung durch eine Präsentation von Kausalinformationen im Stapeldiagramm entlastet und partiell auf die Wahrnehmungsebene verlagert werden kann (Fiedler, 2000; Larkin & Simon, 1987; Wickens & Carswell, 1995).

H1b: Bearbeitungszeit<sub>Diagramm</sub> < Bearbeitungszeit<sub>Tabelle</sub>

### *Einfluss der Präsentationsform auf die Verarbeitungseffizienz*

Die Verarbeitungseffizienz pro Person wird aus dem Quotienten *Anzahl regel-konformer Lösungen / Bearbeitungszeit* ermittelt. Insofern ergeben sich aus einer hohen Anzahl regel-konformer Lösungen und einer geringen Bearbeitungszeit hohe Verarbeitungseffizienzen. Aus den Hypothesen H1a und H1b kann demzufolge eine im Vergleich zu tabellarischen Vierfeldertafeln höhere mittlere Verarbeitungseffizienz bei Personen, die mit Stapeldiagrammen arbeiten, abgeleitet werden. Als theoretische Grundlage kann auch hier wieder die Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) als Erklärung herangezogen werden, die besagt, dass unterschiedliche Informationsanordnungen die Verarbeitungseffizienz beeinflussen, indem sie Auswirkungen auf die Bearbeitungszeit und die Anzahl regel-konformer Lösungen haben.

H1c: Verarbeitungseffizienz<sub>Diagramm</sub> > Verarbeitungseffizienz<sub>Tabelle</sub>

### *Einfluss der Präsentationsform auf die Strategiewahl*

Im Hinblick auf die Strategiewahl wurde jeder Proband dahingehend kategorisiert, welche Strategie/n von ihm angewendet wurden. (Power PC-Theorie,  $\Delta P$ -Regel, ab-Heuristik, unklare/intuitive Heuristik, sonstige oder Wechsler). Bei der Berechnung des Kontingenzmaßes  $\Delta P$  oder der kausalen Stärke  $p$  nach der Power PC Theorie führen die Probanden komplexe mentale Operationen aus. Diese werden nach der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) umso eher ermöglicht, je geringer der Aufwand oder die mentale Beanspruchung für die Berechnung von  $\Delta P$  oder kausaler Stärke  $p$  ist. Wir vermuten, dass Stapeldiagramme im Vergleich zu tabellarischen Vierfeldertafeln die mentale Beanspruchung bei der Verarbeitung von Informationen zur Ermittlung der Stärke eines Kausalzusammenhangs reduzieren. Es wird deshalb erwartet, dass Probanden (Pbn), die mit Stapeldiagrammen arbeiten, signifikant häufiger entweder die  $\Delta P$ -Regel oder die Power PC Theorie anwenden als Probanden, die mit tabellarischen Vierfeldertafeln arbeiten.

H1d: # Pbn mit  $\Delta P$  oder PPC<sub>Diagramm</sub> > # Pbn mit  $\Delta P$  oder PPC<sub>Tabelle</sub>

## 8.2.2 Einfluss des Faktors Informationsart

### *Einfluss der Informationsart auf die Anzahl regel-konformer Lösungen*

Es wird erwartet, dass Personen, denen die kausalen Informationen als Wahrscheinlichkeiten dargeboten werden, signifikant häufiger regel-konforme Lösungen bei der Ermittlung der Wirksamkeit einer Ursache anwenden als Personen, denen die kausalen Informationen als Häufigkeiten dargeboten werden. Verschiedene Studien (Cosmides & Tooby; 1996; Fiedler, 1988; Gigerenzer & Hoffrage, 1995) haben Hinweise darauf ergeben, dass der Faktor *Informationsart* einen Einfluss auf das Auftreten von Urteilsverzerrungen bei der Lösung statistischer Probleme haben kann. Der in Experiment 1 vermutete Vorteil von als Wahrscheinlichkeiten dargebotenen Informationen gegenüber Häufigkeiten ergibt sich daraus, dass die Probanden ihre Wirksamkeitsratings auf einer Antwortskala von 0 – 100 eintragen. Ermitteln die Probanden die Wirksamkeiten auf der Basis vorliegender Wahrscheinlichkeiten, so bleibt ihnen eine Umwandlung von Häufigkeitsangaben in Prozentwerte erspart. Liegen den Probanden Häufigkeiten vor, so stehen sie vor einer zusätzlichen Aufgabe, was die Komplexität der Aufgabe, die Wirksamkeit einer Ursache zu ermitteln, erhöhen dürfte. Ist eine Umwandlung von Häufigkeiten in Prozentwerte erforderlich, dann ist also ebenfalls mit einem Anstieg der Gedächtnisbeanspruchung zu rechnen und es kommt entweder zu fehlerhaften Kausalurteilen (Fiedler, 2000; Kao & Wasserman, 1993; Shaklee & Mims, 1982), oder es werden von den Probanden einfachere Strategien eingesetzt (Kleinmuntz & Schkade, 1993; Shaklee & Mims, 1982).

H2a: # regel-konformer Lösungen<sub>Wahrscheinlichkeiten</sub> > # regel-konformer Lösungen<sub>Häufigkeiten</sub>

Gigerenzer & Hoffrage (1995) kamen in ihrer Untersuchung zu dem Ergebnis, dass mehr Probanden Bayesianische Algorithmen anwendeten, wenn Ihnen die Ausgangsinformationen als Häufigkeiten vorgelegt wurden. Lagen die Ausgangsinformationen hingegen als in der Form von Wahrscheinlichkeiten vor, dann reduzierte sich die Anzahl der Probanden, die Bayesianische Algorithmen einsetzten. Der für unser Experiment 1 vorhergesagte Vorteil von Wahrscheinlichkeiten ergibt sich aus der beschriebenen Wechselwirkung zwischen den Einflüssen der Faktoren *Informationsart* und Antwortskala: Sollen die Probanden die ermittelten Wirksamkeiten auf einer Skala von 0 bis 100 eintragen, so müssen sie bei der Vorgabe von Wahrscheinlichkeiten nicht noch Häufigkeiten in Prozentwerte umrechnen.

### *Einfluss der Informationsart auf die Bearbeitungszeit*

Es wird erwartet, dass die Bearbeitungszeit bei Personen, denen Häufigkeiten vorgegeben werden, höher ist als bei Personen, die mit Wahrscheinlichkeiten arbeiten, da im Falle von Häufigkeiten und bei einem Verzicht auf Schätzungen zur Ermittlung eines Wertes zwischen 0 und 100 zusätzlich Prozentrechnungen durchgeführt werden müssen. Es kann also davon ausgegangen werden, dass der zeitliche Aufwand und die mentale Beanspruchung bei der Vorgabe von Wahrscheinlichkeiten geringer ausfällt als bei der Vorgabe von Häufigkeiten.

H2b  $\text{Bearbeitungszeit}_{\text{Wahrscheinlichkeiten}} < \text{Bearbeitungszeit}_{\text{Häufigkeiten}}$

#### *Einfluss der Informationsart auf die Verarbeitungseffizienz*

Entsprechend der Ableitung von Hypothese H1c, so kann auch die Hypothese H2c aus den Hypothesen H2a und H2b abgeleitet werden. Basierend auf der Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) kann erwartet werden, dass bei Personen, die Wahrscheinlichkeiten zu verarbeiten haben, eine höhere Verarbeitungseffizienz zu beobachten ist als bei Personen, die mit Häufigkeiten arbeiten.

H2c  $\text{Verarbeitungseffizienz}_{\text{Wahrscheinlichkeiten}} < \text{Verarbeitungseffizienz}_{\text{Häufigkeiten}}$

#### *Einfluss der Informationsart auf die Strategiewahl*

Es wird erwartet, dass die Zahl der Probanden, die entweder von der  $\Delta P$ -Regel oder der Power PC Theorie Gebrauch machen, unter der Wahrscheinlichkeitsbedingung höher ausfallen wird als unter der Häufigkeitsbedingung. Im Vergleich zu Häufigkeiten verringern Wahrscheinlichkeiten den Aufwand oder die mentale Beanspruchung für die Berechnung von  $\Delta P$  oder der kausalen Stärke  $p$ , der Einsatz von komplexen Rechenoperationen wird dadurch einfacher gemacht.

H2d:  $\# \text{Pbn mit } \Delta P \text{ oder PPC}_{\text{Wahrscheinlichkeiten}} > \# \text{Pbn mit } \Delta P \text{ oder PPC}_{\text{Häufigkeiten}}$

### **8.2.3 Vergleich der Experimentalbedingungen mit einer Protokollliste**

Aus der Kreuzung der beiden zweistufigen Faktoren *Präsentationsform* und *Informationsart* ergeben sich die vier Kombinationen Tabelle mit Häufigkeiten (*TH*), Tabelle mit Wahrscheinlichkeiten (*TW*), Stapeldiagramm mit Häufigkeiten (*SH*) und Stapeldiagramm mit Wahrscheinlichkeiten (*SW*). In verschiedenen Arbeiten (Van Hamme & Wassermann, 1994; Hagemayer, 2001, Ex 3; White, 2000b; White, 2003c, Ex 1, 4 und 5) wird als summarische Präsentationsform aber eine Protokollliste (kurz: *PL*) verwendet, die sich nach unserer Auffassung stark von den vier Experimentalbedingungen *TH*, *TW*, *SH* und *SW* unterscheidet.

In einer Protokollliste finden die Probanden Beobachtungen einzelner Ursache-Wirkungszusammenhänge aneinander gereiht. In den vier Experimentalbedingungen werden den Probanden im Vergleich zu einer Protokollliste das Aufaddieren und die richtige Zuordnung zu einer der vier Zellkategorien abgenommen. Zudem wird den Probanden, die mit einer Protokollliste arbeiten, im Vergleich zur Informationsdarbietung in einem Stapeldiagramm nicht nahe gelegt, wie sie die summarischen Informationen zur Ermittlung der Wirksamkeit einer Ursache zueinander in Beziehung setzen können - selbst wenn sie in der Lage waren, die einzelnen Informationen über Ursache-Wirkung-Zusammenhänge richtig aufzuaddieren und zu gruppieren. Im Vergleich zu Probanden, die die Wirksamkeiten auf einer Skala von 0 bis 100 auf der Grundlage von Wahrscheinlichkeiten ermitteln, müssen die Probanden, denen eine Protokollliste präsentiert wird, die aggregierten Summenwerten in Prozentwerte umwandeln, wenn sie von einer regel-basierten Strategie Gebrauch machen möchten.

Insgesamt ist zu erwarten, dass die Probanden bei der Präsentation einer Protokollliste im Vergleich zu allen Bedingungen, die sich aus der Kombination der Faktoren *Präsentationsform* und *Informationsart* ergeben, einer höheren mentalen Beanspruchung ausgesetzt sind und weniger Strukturierungshilfen (Larkin & Simon, 1987; Stenning & Oberlander, 1995; Zhang, 2000, 1997; Zhang & Norman, 1994) erhalten. Die bei der Darbietung von Protokolllisten höhere mentale Beanspruchung ergibt sich aus einer höheren Komplexität der Aufgabenstellung, weil den Probanden bei der Bestimmung von Wirksamkeiten nach einer regel-basierten Strategie weder Rechenschritte abgenommen werden, noch ein richtiges Zueinander-in-Beziehung-setzen und Vergleichen von Erfolgs- und Basisrate durch die grafische Anordnung der Informationen über Ursache-Wirkungszusammenhänge nahe gelegt wird. Nach Larkin & Simon (1987) kann angenommen werden, dass eine Protokollliste keine perzeptuellen Inferenzen ermöglicht, weil komplexe mentale Operationen wie z.B. Vergleiche oder Schlussfolgerungen nicht auf die Wahrnehmungsebene verlagert werden (Wickens & Carswell, 1995). Nach der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) ist anzunehmen, dass sich der höhere Aufwand für die Probanden unter der Listenbedingung in längeren Bearbeitungszeiten zeigt. Als Folge der erhöhten mentalen Beanspruchung unter der Listenbedingung setzen die Probanden weniger regel-basierte Strategien ein und produzieren im Mittel weniger regel-konforme Lösungen, weil sie entweder auf einfachere Strategien zurückgreifen (Kleinmuntz & Schkade, 1993; Shaklee & Mims, 1982), oder ihnen bei der Anwendung komplexer Strategien mehr Fehler unterlaufen (Fiedler, 2000; Kao & Wasserman, 1993; Shaklee &

Mims, 1982). Aus den längeren Bearbeitungszeiten und der geringeren Anzahl regelkonformer Lösungen kann schließlich eine geringere Verarbeitungseffizienz bei Darbietung einer Protokollliste abgeleitet werden. Für den Vergleich zwischen den vier Experimentalbedingungen *TH*, *TW*, *SH*, *SW* und der Kontrollbedingung *PL* lassen sich zusammengefasst in Kurzschreibweise folgende Hypothesen formulieren:

H3a: # regel-konformer Lösungen<sub>TH, TW, SH, SW</sub> > # regel-konformer Lösungen<sub>PL</sub>

H3b: Bearbeitungszeit<sub>TH, TW, SH, SW</sub> < Bearbeitungszeit<sub>PL</sub>

H3c: Verarbeitungseffizienz<sub>TH, TW, SH, SW</sub> > Verarbeitungseffizienz<sub>PL</sub>

H3d: # Pbn mit  $\Delta P$  oder PPC<sub>TH, TW, SH, SW</sub> > # Pbn mit  $\Delta P$  oder PPC<sub>PL</sub>

#### 8.2.4 Einfluss des Faktors Studienausrichtung

Ferner erwarten wir erhebliche individuelle Unterschiede hinsichtlich der Strategiewahl bei den einzelnen Teilnehmern. Die Heterogenität der Stichprobe stellt eine Gefährdung für die Varianzanalyse dar, die in der folgenden Untersuchung durch relativ große Zellbesetzungen ( $n = 16$ ) ausgeglichen werden soll. Eine Homogenisierung der Stichprobe wurde nur insofern vorgenommen, dass wir uns auf Studenten und angehende Studenten beschränkt haben. Eine weitere Homogenisierung, z.B. durch eine Rekrutierung von ausschließlich Mathematikstudenten, Theologen oder Psychologen wurde absichtlich nicht vorgenommen, weil dadurch die ökologische Validität der Studie und damit die Generalisierbarkeit ihrer Ergebnisse weiter eingeschränkt worden wären. Zudem ist unklar, an Hand welcher Merkmale sich eine Stichprobe sinnvoll homogenisieren ließe. Um Anhaltspunkte dafür zu bekommen, welche individuellen Voraussetzungen (Kovariaten) die Art und Weise der Beurteilung und Einschätzung von Ursache-Wirkungs-Zusammenhängen mit beeinflussen, haben wir eine schriftliche Nachbefragung mit allen Probanden durchgeführt. Diese Fragen betrafen im Wesentlichen Vorerfahrungen mit den eingesetzten Medien (Ereignislisten, Tabellen, Diagramme), sowie den Umgang mit Grundrechenarten und Prozentrechnen. Zudem wurde für alle Teilnehmer erhoben, welches Fach sie studieren bzw. welches Fach sie zu studieren beabsichtigen. Dabei wurde darauf geachtet, dass die eine Hälfte der Teilnehmer einen Studiengang mit mathematisch-naturwissenschaftlicher Studienausrichtung absolvierte oder absolvieren wollte. Für die andere Hälfte wurde darauf geachtet, dass es sich um Probanden handelte, denen im (beabsichtigten) Studium naturwissenschaftliche oder mathematische Methoden kaum oder gar nicht vermittelt werden.

In Anlehnung an Shaklee & Hall (1983) wird erwartet, dass Probanden, die Fächer mit mathematisch-naturwissenschaftlicher Ausrichtung studieren (bzw. studieren wollen) sich



leichter damit tun, ein regel-konformes mentales Modell oder Schema zur Berechnung der Wirksamkeiten entweder abzurufen oder zu erstellen, da in ihrem Gedächtnis entsprechende oder ähnliche Schemata aufgrund häufigerer Nutzung (Assoziationsstärke) leichter abrufbar sind. Demzufolge dürften diese Probanden im Vergleich zur Gruppe derjenigen, die Fächer ohne mathematisch-naturwissenschaftliche Ausrichtung studieren (bzw. studieren wollen) im Mittel häufiger regel-basierte Strategien anwenden und mehr regel-konforme Lösungen produzieren. Außerdem werden für Studierende mit mathematisch-naturwissenschaftlicher Fachausrichtung kürzere Bearbeitungszeiten und damit eine geringere Verarbeitungseffizienz angenommen. Für den Vergleich zwischen den beiden, sich hinsichtlich der Studiaausrichtung (mathematisch-naturwissenschaftlich orientiert – kurz *NW* - vs. nicht mathematisch-naturwissenschaftlich orientiert – kurz *GW* - für geisteswissenschaftlich orientiert) unterscheidenden Probandengruppen lassen sich in Kurzschreibweise folgende Hypothesen zusammenfassen:

H4a: # regel-konformer Lösungen<sub>NW</sub> > # regel-konformer Lösungen<sub>GW</sub>

H4b: Bearbeitungszeit<sub>NW</sub> < Bearbeitungszeit<sub>GW</sub>

H4c: Verarbeitungseffizienz<sub>NW</sub> > Verarbeitungseffizienz<sub>GW</sub>

H4d: # Pbn mit  $\Delta P$  oder PPC<sub>NW</sub> > # Pbn mit  $\Delta P$  oder PPC<sub>GW</sub>

### 8.3 Methodik und Vorgehensweise

#### Probanden

An dem Experiment nahmen 80 Versuchspersonen (58 Frauen und 22 Männer) teil. Das Alter variierte zwischen 19 und 52 Jahren ( $M = 23.6$ ;  $s = 6.6$ ). Voraussetzung für die Teilnahme waren gute deutsche Sprachkenntnisse. Alle Teilnehmer waren Studierende der Universität Tübingen oder angehende Studenten in Tübingen. Bei den angehenden Studenten handelte es sich um Abiturienten, die am Leibniz-Kolleg in Tübingen ein einjähriges Studium Generale absolvierten, bevor sie mit dem Studium an einer Universität beginnen wollten. Bei der einen Hälfte der Probanden handelte es sich um Teilnehmer, die einen Studiengang absolvierten oder absolvieren wollten, in denen mathematische Fächer (wie z.B. Statistik) oder naturwissenschaftliche Methoden gelehrt wurden. Die andere Hälfte der Probanden setzte sich aus Studierenden oder angehenden Studierenden geisteswissenschaftlicher Disziplinen zusammen, in denen naturwissenschaftliche oder mathematische Methoden kaum oder gar nicht vermittelt wurden. Die Zusammensetzung der Probanden hinsichtlich ihrer realen oder beabsichtigten Studienfächer und sonstiger demografischer Merkmale kann in Tabelle A1 im Anhang A1.1 nachgesehen werden. Die Versuchsteilnehmer wurden extern durch Aushänge oder Ansprache angeworben und erhielten für die Teilnahme DM 10,-. Die Zuteilung zu den Versuchsbedingungen erfolgte per Zufall. Die Probanden wurden einzeln untersucht und gleichmäßig auf die vier Experimentalbedingungen und die Kontrollbedingung verteilt.

#### Material

Alle Probanden sollten beurteilen, ob und in welchem Ausmaß ein Ereignis A (die Ursache) ein Ereignis B (die Wirkung) hervorruft. Anders als in den Untersuchungen von Van Hamme & Wasserman (1994) lagen keine offensichtlich konkurrierenden Ursachen vor. In der Instruktion, die im Anhang A1.2 eingesehen werden kann, wurde allen Probanden mitgeteilt, dass ein Chemieunternehmen an der Entwicklung neuer Substanzen arbeitet, durch welche Wolken mittels Wolkenimpfung zum Abregnen gebracht werden sollten. Das von uns gewählte Wolkenimpfparadigma wurde schon in den Studien von Ward & Jenkins (1965) eingesetzt und bietet den Vorzug, dass die Probanden in der Regel über kein spezifisches Vorwissen zu einem solchen Thema verfügen.

#### Untersuchungsdurchführung

Nach der Begrüßung durch den Versuchsleiter nahmen die Probanden an einem Tisch Platz und beantworteten auf einem Blatt Fragen zu Geschlecht, Alter und Studienfach (bzw.

Studienwunsch), danach wurde ihnen die Instruktion ausgeteilt. Der Versuchsleiter war während des gesamten Versuchs anwesend.

Über alle Versuchsbedingungen hinweg, wurde den Probanden in der Instruktion mitgeteilt, dass im Auftrag eines Chemie-Unternehmens sechs Testflüge durchgeführt worden waren und dabei jeweils sechs verschiedene Substanzen getestet wurden. Wurde den Probanden eine *Tabelle* vorgelegt, so wurden sie darüber informiert, dass die Testergebnisse für jede Substanz auf einem separaten Blatt in einer Tabelle zusammengefasst sind. Wurde den Probanden hingegen ein *Diagramm* vorgelegt, so wurde ihnen mitgeteilt, dass die Testergebnisse für jede Substanz auf einem separaten Blatt in einem Säulendiagramm zusammengefasst sind. Enthielten Tabelle oder Diagramm *Häufigkeiten*, dann wurde den Probanden weiter mitgeteilt, dass die zusammen gefassten Ergebnisse vier verschiedene Arten von Informationen enthalten würden:

1. *Anzahl* geimpfter Wolken, die im Anschluss an eine Impfung zu regnen begonnen haben.
2. *Anzahl* geimpfter Wolken, die im Anschluss an eine Impfung nicht zu regnen begonnen haben.
3. *Anzahl* nicht geimpfter Wolken, die spontan zu regnen begonnen haben.
4. *Anzahl* nicht geimpfter Wolken, die nicht zu regnen begonnen haben.

Enthielten die Tabellen oder Diagramme hingegen *Wahrscheinlichkeiten*, dann wurde in der Instruktion das Wort "Anzahl" durch *Prozentzahl* ersetzt.

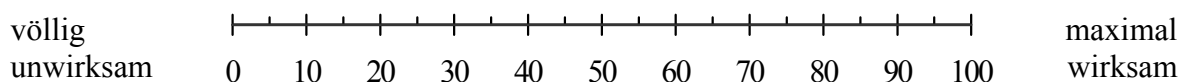
In der Kontrollbedingung *Protokollliste* wurde den Probanden erklärt, dass die Testergebnisse für jede Substanz auf einem separaten Blatt in einer Liste zusammengefasst sind. Zudem sollten auch sie an Hand der Ergebniszusammenfassung die Wirksamkeit einer jeden Substanz beurteilen. Die Liste enthielt vier verschiedene Arten von Einzelinformationen:

1. *Eine Wolke* wurde mit der Substanz geimpft und fing im Anschluss an die Impfung zu regnen an.
2. *Eine Wolke* wurde mit der Substanz geimpft und fing im Anschluss an die Impfung nicht zu regnen an.
3. *Eine Wolke* wurde nicht mit der Substanz geimpft und fing spontan an zu regnen.
4. *Eine Wolke* wurde nicht mit der Substanz geimpft und fing nicht zu regnen an.

Nach dem Lesen der Instruktion wurden die Probanden gebeten, Fragen zu stellen und die Instruktion kurz zusammenzufassen. Falls Unklarheiten bestanden, wurden diese geklärt. Dann wurden allen Probanden dieselben Daten auf sechs Aufgabenblättern mit Angaben zu sechs verschiedenen Substanzen präsentiert. Die Reihenfolge der sechs Aufgabenblätter war

zufällig. Während der eigentlichen Testphase wurden den Probanden in summarischer Form für jede Substanz vierzig Beobachtungsergebnisse von je zwanzig geimpften und nicht geimpften, ursprünglich nicht regnenden Wolken separat auf einem Blatt präsentiert. Von den sechs Substanzen waren vier Substanzen für die spätere statistische Auswertung relevant, die anderen beiden dienten als Füllmaterial. Für die vier auswertungsrelevanten Substanzen wurden vier unterschiedliche Erfolgsraten definiert: +.25; +.50; +.75 und 1.00. Die Basisrate betrug für alle diese Substanzen +.25. Würde ein Proband bei der Aufgabenbearbeitung durchgängig  $\Delta P$  (Erfolgsrate – Basisrate) berechnen und, wenn nötig, das Ergebnis durch eine Prozentumwandlung auf einer Skala von 0 bis 100 abbilden, so käme er dabei auf Werte von 0, 25, 50 und 75. Würde ein Proband seinen Berechnungen hingegen die Power PC-Theorie zu Grunde legen, so müsste er auf einer Skala von 0 bis 100 die Werte 0, 33.33, 66.67 und 100 eintragen. Für die beiden nicht in der späteren Auswertung berücksichtigten Substanzen wurden Erfolgsraten von 0 und 1.00, sowie eine Basisrate von 0 festgelegt.

Zur Einschätzung der Wirksamkeit einer Substanz stand den Probanden jeweils die in Abbildung 8 abgebildete Ankreuzskala zur Verfügung. Die Skala reichte von 0 bis 100. 0 bedeutete dabei, dass diese Substanz völlig unwirksam sei und keinerlei Einfluss auf das Regenverhalten der Wolken hatte, 100 bedeutet, dass diese Substanz maximal wirksam sei und man damit jede Wolke zum Abregnen bringen könne.



**Abbildung 8:** Experiment 1: Wirksamkeitsskala.

Falls die Probanden die Wirksamkeit einer Substanz nicht geschätzt, sondern einen genauen Wert berechnet hatten, so konnten Sie diesen Wert in ein Kästchen eintragen.

Bei der Aufgabenbearbeitung standen die Versuchspersonen nicht unter Zeitdruck. Notizen waren erlaubt, allerdings durfte zur nachträglichen Ergebniskorrektur nicht zurückgeblättert werden. Der Versuchsleiter erhob für jeden Probanden die Dauer der Aufgabenbearbeitung. Unmittelbar nach der Aufgabenbearbeitung wurden die Versuchsteilnehmer schriftlich zur Vorgehensweise beim Lösen der Aufgaben befragt (s. Anhang A1.3). Danach wurde den Probanden das Experiment erläutert und ihnen eine Abschlussbefragung (s. Anhang A1.4) ausgeteilt. Darin wurden die Probanden zu mathematischen Vorkenntnissen (Grundrechenarten und Prozentrechnen) befragt und dazu, wie vertraut ihnen die auf Beiblättern (s. Anhang A1.5) gezeigten Listen, Diagramme und Tabellen sind. Während die Versuchspersonen die Fragebögen ausfüllten, sah der Versuchsleiter die Aufgabenblätter der

Probanden durch, um sich für eine Nachbesprechung vorzubereiten. Nachdem der Versuchsleiter die Ergebnisse und die Vorgehensweise bei der Aufgabenbearbeitung besprochen und gegebenenfalls Nachfragen gestellt hatte, wurden die Probanden entlohnt und mit Dank verabschiedet.

## **8.4 Explorative Datenanalysen**

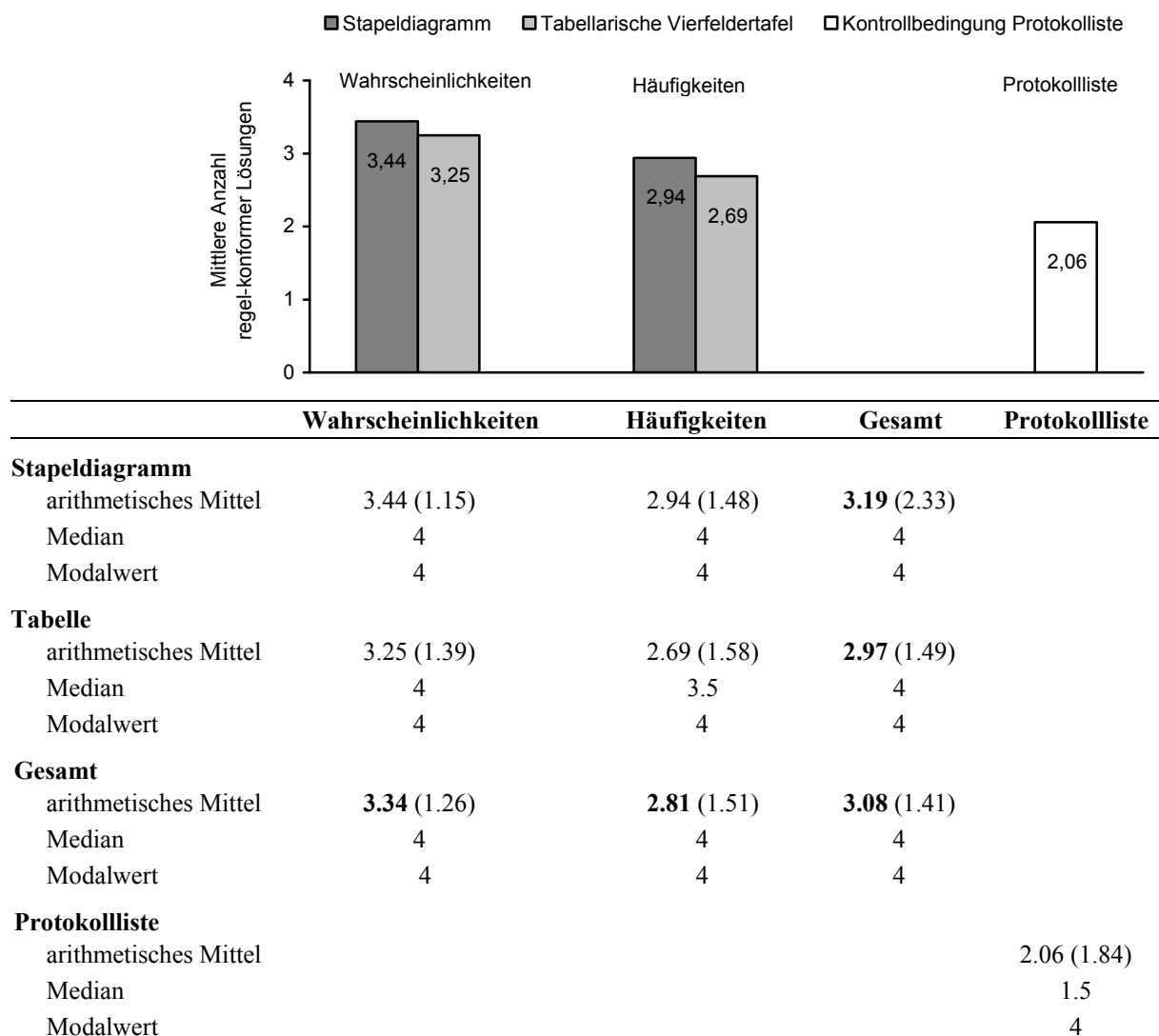
Alle statistischen Analysen wurden mit dem Programmpaket SPSS ausgewertet. Mit allen untersuchten abhängigen und unabhängigen Variablen wurden dabei zunächst explorative Datenanalysen durchgeführt, um die Voraussetzungen der Varianzanalyse zu überprüfen. Die Voraussetzung der Normalverteilung wurde durch Histogramme und statistisch mit Hilfe des Lilliefors-Tests und des Shapiro-Wilks-Tests überprüft. Um die Homogenität der Varianzen zu überprüfen, wurde der Levene-Test eingesetzt. Die dritte oft genannte Voraussetzung der Varianzanalyse (s. Bortz, 1999; Glaser, 1978), Zufallsauswahl und Zufallsverteilung der Messobjekte auf die Untersuchungsbedingungen, kann aufgrund der Versuchsdurchführung als gegeben gelten. Hinsichtlich der Voraussetzung der Normalverteilung konnten insbesondere bei den meisten Zellen der abhängigen Variable *Anzahl regel-konformer Lösungen* Verstöße festgestellt werden, die auf Deckeneffekte zurückzuführen sind, weil relativ viele Probanden eine regel-konforme Strategie bei der Aufgabenbearbeitung eingesetzt haben. Der Levene-Test erbrachte Hinweise auf heterogene Fehlervarianzen der abhängigen Variable *Anzahl regel-konformer Lösungen*. Da die Varianzanalyse aber bei ausreichender Zellgröße ( $n > 10$ ) und gleichgroßen Stichproben gegenüber Verletzungen ihrer Voraussetzungen relativ robust ist (Bortz, 1999; Glaser, 1978; Kirk, 1995), haben wir auf ihre Durchführung nicht verzichtet und zudem bei der abhängigen Variable *Anzahl regel-konformer Lösungen* den Mann-Whitney U-Test in Abhängigkeit von den beiden zweistufigen Variablen *Präsentationsform* und *Informationsart* als verteilungsfreie, wenn auch weniger aussagekräftige Alternative eingesetzt. Mit rangdatenbasierten Verfahren wie dem Mann-Whitney U-Test können keine Aussagen über Interaktionseffekte gemacht werden.

## 8.5 Einfluss der Faktoren Präsentationsform und Informationsart

### 8.5.1 Ergebnisse

#### Anzahl regel-konformer Lösungen

Abbildung 9 zeigt die Mittelwerte für die Anzahl regel-konformer Lösungen in Abhängigkeit von den beiden Faktoren *Präsentationsform* und *Informationsart* im Vergleich zum Mittelwert für die Kontrollbedingung *Protokollliste*. Die angehängte Tabelle enthält neben dem arithmetischen Mittel noch Median und Modalwert.



**Abbildung 9:** Experiment 1: Mittlere Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) in Abhängigkeit von *Präsentationsform* und *Informationsart*.

Eine Varianzanalyse unter Ausschluss der Kontrollbedingung *PL* mit den Faktoren *Präsentationsform* und *Informationsart* wurde nicht signifikant, auch die Interaktion verfehlte

das Signifikanzniveau von  $\alpha = 5\%$ ;  $F_{\text{Präsentationsform}}(1, 60) = .385$ ,  $p = .537$ ;  $F_{\text{Informationsart}}(1, 60) = 2.271$ ,  $p = .137$ ;  $F_{\text{Interaktion}}(1, 60) = .008$ ,  $p = .930$ ;  $MSE = 1.989$ . Erwähnenswert ist aber, dass der Faktor *Informationsart* in Trendnähe kam ( $p = .137$ ), was darauf hinweist, dass die Probanden bei einer Präsentation von Wahrscheinlichkeiten etwas mehr regel-konforme Lösungen produzierten als bei der Präsentation von Häufigkeiten.

Da die varianzanalytischen Voraussetzungen der Normalverteilung und der Varianzhomogenität bei der Variable *Anzahl regel-konformer Lösungen* verletzt waren, wurden die Ergebnisse in Rangdaten transformiert und zusätzlich Mann-Whitney U-Tests in Abhängigkeit von *Präsentationsform* und *Informationsart* gerechnet. Diese erbrachten keine wesentlich anderen Ergebnisse: Erneut kam der Faktor *Informationsart* in Trendnähe ( $p = .114$ ), während der Faktor *Präsentationsform* ( $p = .567$ ) das Signifikanzniveau klar verfehlte. Zu Klärung der Frage, ob sich die vier Experimentalbedingungen *SW*, *TW*, *SH* und *TH* und die Kontrollbedingung *PL* signifikant voneinander unterscheiden, wurde eine weitere Varianzanalyse durchgeführt und die *Anzahl regel-konformer Lösungen* in Abhängigkeit vom Faktor *Versuchsbedingung* untersucht. Dabei wurde für den Faktor *Versuchsbedingung* ein Trend erzielt;  $F_{\text{Versuchsbedingung}}(4, 75) = 2.037$ ,  $p = .098$ ;  $MSE = 2.270$ . Daran anschließend wurden Einzelkontraste zwischen den vier Experimentalbedingungen und der Kontrollbedingung berechnet. Orthogonale a priori Kontraste zeigten, dass die Mittelwerte für die Bedingungen *SW* und *TW* signifikant höher sind als derjenige für die Kontrollbedingung *PL* ( $p_{\text{SW-PL}} < .02$ ;  $p_{\text{TW-PL}} < .05$ ). Die Bedingungen *SH* und *TH* unterscheiden sich allerdings nicht signifikant von der Kontrollbedingung *PL*, wengleich auch hier für den Unterschied *SH-PL* Trendnähe erreicht wurde ( $p_{\text{SH-PL}} < .15$ ).

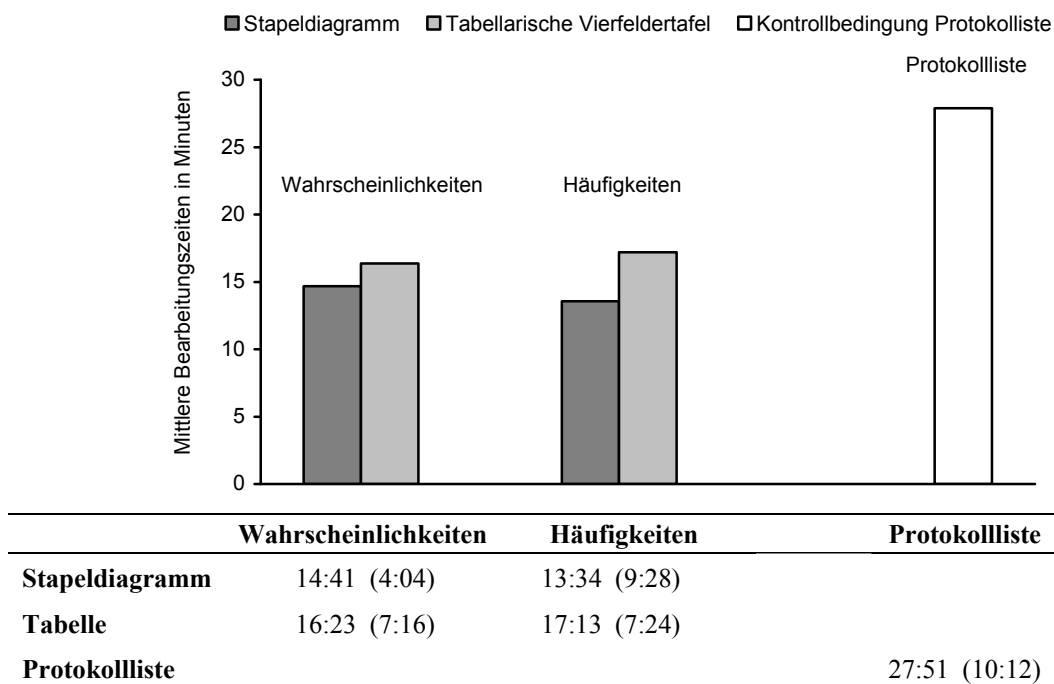
Als ein nicht orthogonales a priori Kontrastverfahren, das speziell für den Vergleich von Experimentalgruppenmittelwerten mit einem Kontrollgruppenmittelwert eingesetzt werden kann, wurde der Dunnett Multiple Comparison Test entwickelt (Dunnett, 1955; für eine nähere Beschreibung s. Kirk, 1995). Wird den Daten aus Experiment 1 eine Dunnett tDN-Statistik zu Grunde gelegt, so ergeben sich für die Kontraste *SW-PL* und *TW-PL* signifikante Unterschiede ( $p_{\text{SW-PL}} = .020$ ;  $p_{\text{TW-PL}} = .049$ ), für den Kontrast *SH-PL* ergab sich eine Trendnähe ( $p = .149$ ), während der Kontrast *TH-PL* nicht signifikant wurde ( $p = .303$ ).

Die gefundenen Ergebnisse bestätigen die zu der abhängigen Variable *Anzahl regel-konformer Lösungen* aufgestellten Hypothesen H1a und H2a nicht. In Bezug auf die postulierten Unterschiede zwischen Stapeldiagramm und tabellarischer Vierfeldertafel einerseits (H1a) und zwischen Häufigkeiten und Wahrscheinlichkeiten (H2a) müssen beide Hypothesen verworfen werden, auch wenn der Faktor *Informationsart* Trendnähe erreichte. Im Hinblick

auf den Vergleich zwischen den vier Experimentalbedingungen und der Kontrollbedingung zeigte sich, dass nur die beiden Wahrscheinlichkeitsbedingungen *SW* und *TW* der Kontrollbedingung *PL* überlegen sind, während sich die beiden Häufigkeitsbedingungen *SH* und *TH* nicht signifikant von *PL* unterscheiden. Die Hypothese H3a kann also teilweise beibehalten werden.

### Bearbeitungszeit

Abbildung 10 enthält die mittleren Bearbeitungszeiten der Probanden aus den vier Experimentalgruppen und der Kontrollbedingung.



**Abbildung 10:** Experiment 1: Mittlere Bearbeitungszeiten in Minuten : Sekunden (in Klammern Standardabweichungen) in Abhängigkeit von *Präsentationsform* und *Inhaltsart*.

Der varianzanalytische Vergleich unter Ausschluss der Kontrollbedingung *PL* mit den beiden Faktoren *Präsentationsform* und *Informationsart* brachte weder hinsichtlich der Haupteffekte noch hinsichtlich der Interaktion signifikante Ergebnisse;  $F_{\text{Präsentationsform}}(1, 60) = 2.055$ ;  $p = .157$ ,  $F_{\text{Informationsart}}(1, 60) = .007$ ,  $p = .933$ ;  $F_{\text{Interaktion}} = .273$ ,  $p = .603$ ;  $MSE = 54.926$ . Bemerkenswert ist aber, dass sich für die *Bearbeitungszeit* (s. Abbildung 10) ein etwas anderes Ergebnismuster zeigt als für die *Anzahl regel-konformer Lösungen* (s. Abbildung 9). Diesmal kam nicht der Faktor *Informationsart* in Trendnähe, sondern der Faktor *Präsentationsform* ( $p = .157$ ), was ein Hinweis darauf sein könnte, dass Stapeldiagramme von den Probanden etwas schneller bearbeitet werden als tabellarische Vierfeldertafeln, während



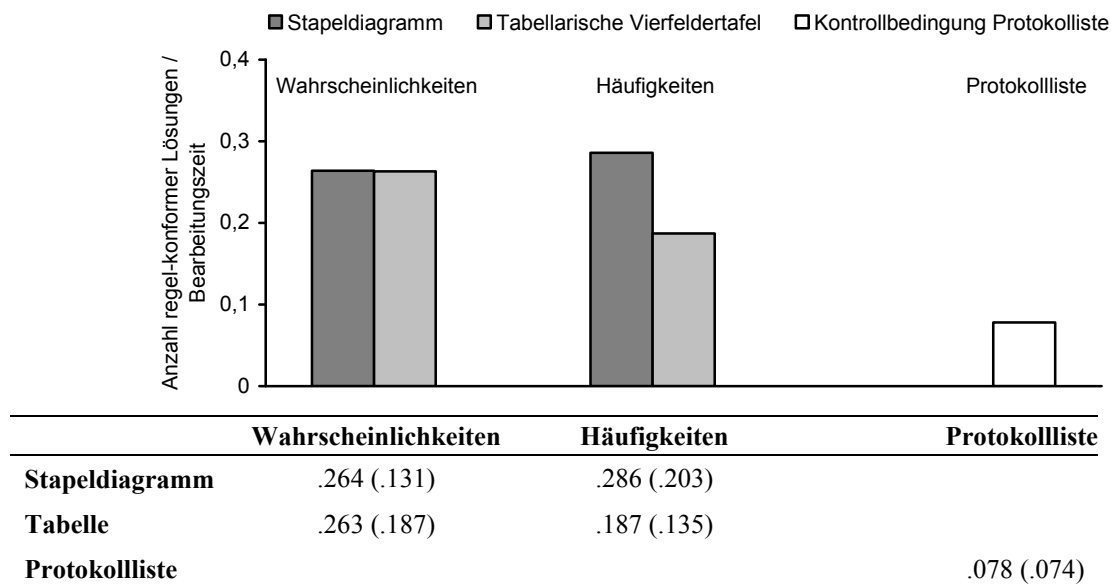
Abbildung 9 darauf hinweist, dass die Probanden bei der Darbietung von Wahrscheinlichkeiten etwas häufiger regel-konforme Lösungen produzieren.

Um herauszufinden, ob sich die vier Experimentalbedingungen signifikant von der Kontrollbedingung unterscheiden, wurde eine einfaktorielle Varianzanalyse mit dem Faktor *Versuchsbedingung* und daran anschließenden a priori Kontrasten durchgeführt. Dabei wurde der Faktor *Versuchsbedingung* hochsignifikant;  $F_{\text{Versuchsbedingung}}(4, 75) = 8.123, p < .001$ ;  $MSE = 64.738$ . Die Einzelkontraste ergaben, dass die Probanden in allen vier Experimentalbedingungen die Aufgaben hochsignifikant schneller bearbeiteten als in der Kontrollbedingung ( $p_{\text{SW-PL}} < .001$ ;  $p_{\text{TW-PL}} \leq .001$ ;  $p_{\text{SH-PL}} < .001$ ;  $p_{\text{TH-PL}} \leq .002$ ). Der nicht orthogonale Dunnett Multiple Comparison Test führte exakt zu denselben Ergebnissen.

Zusammengefasst müssen für die *Bearbeitungszeit* die Hypothesen H1b und H2b trotz der Trendnähe des Faktors *Präsentationsform* verworfen werden, während H3b beibehalten werden kann.

### Verarbeitungseffizienz

Um die *Verarbeitungseffizienz* zu berechnen, wurde die *Anzahl regel-konformer Lösungen* zur *Bearbeitungszeit* in Minuten ins Verhältnis gesetzt. Die Mittelwerte für dieses Effizienzmaß in Abhängigkeit von den vier Experimentalbedingungen und der Kontrollbedingung lassen sich aus Abbildung 11 ersehen.



**Abbildung 11:** Experiment 1: Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von den beiden Faktoren *Präsentationsform* und *Inhaltsart*.

Im varianzanalytischen Vergleich mit den beiden Faktoren *Präsentationsform* und *Informationsart* und unter Ausschluss der Kontrollbedingung *PL* waren weder Haupteffekte

noch Interaktion signifikant;  $F_{\text{Präsentationsform}}(1, 60) = 1.424, p = .238$ ;  $F_{\text{Informationsart}}(1, 60) = .410, p = .524$ ;  $F_{\text{Interaktion}} = 1.38; p = .245$ ;  $MSE = .0279$ . In der einfaktoriellen Varianzanalyse mit dem Faktor *Versuchsbedingung* und unter Hinzunahme der Kontrollbedingung *PL* wurde *Versuchsbedingung* hochsignifikant;  $F_{\text{Versuchsbedingung}}(4, 75) = 4.977, p = .001$ ;  $MSE = .023$ . Die im Anschluss daran durchgeführten Einzelkontraste weisen darauf hin, dass die Probanden in den vier Experimentalbedingungen die Aufgaben signifikant effizienter bearbeiteten als in der Kontrollbedingung ( $p_{\text{SW-PL}} < .001$ ;  $p_{\text{TW-PL}} < .002$ ;  $p_{\text{SH-PL}} < .001$ ;  $p_{\text{TH-PL}} < .05$ ). Im zweiseitigen Dunnett Multiple Comparison Test wurde für den Vergleich TH – PL ein Trend erzielt ( $p_{\text{TH-PL}} = .075$ ), alle anderen Kontraste waren hochsignifikant ( $p_{\text{SW-PL}} < .002$ ;  $p_{\text{TW-PL}} < .002$ ;  $p_{\text{SH-PL}} < .001$ ).

Die empirischen Ergebnisse für das Effizienzmaß *Anzahl regelkonformer Lösungen / Bearbeitungszeit* können aus den Ergebnissen in den Abbildungen 9 und 10 abgeleitet werden und bestätigen die aufgestellten Hypothesen erneut nur teilweise. Während H1c und H2c verworfen werden müssen, kann H3c beibehalten werden.

### 8.5.2 Diskussion

Insgesamt bleibt festzuhalten, dass in Experiment 1 weder die beiden Hauptfaktoren *Präsentationsform* und *Informationsart*, noch die Interaktion *Präsentationsform*  $\times$  *Informationsart* einen signifikanten Einfluss auf die abhängigen Variablen *Anzahl regelkonformer Lösungen*, *Bearbeitungszeit* und *Verarbeitungseffizienz* hatten. Obwohl die Mittelwertunterschiede meist in die vorhergesagte Richtung gingen (Stapeldiagramme besser als tabellarische Vierfeldertafeln; Wahrscheinlichkeiten besser als Häufigkeiten), mussten die Hypothesen H1a bis H1c und H2a bis H2c durchgängig verworfen werden.

Etwas anders liegt der Fall bei den Hypothesen H3a bis H3c. Im Zusammenhang mit der *Verarbeitungseffizienz* und der *Bearbeitungszeit* waren die vier Experimentalbedingungen *SW*, *TW*, *SH* und *TH* der Kontrollbedingung *PL* durchgängig überlegen. In Verbindung mit der *Anzahl regelkonformer Lösungen* schnitten *SW* und *TW* signifikant besser als *PL* ab, der Unterschied zwischen *SH* und *PL* erreichte Trendnähe.

#### Vergleich der vier Experimentalbedingungen SH, SW, TH und TW untereinander

Die Ergebnisse aus Experiment 1 werfen die Frage auf, warum die Unterschiede zwischen den vier Experimentalbedingungen das Signifikanzniveau von  $\alpha = 5\%$  verfehlten:

*Anzahl regelkonformer Lösungen*: In Verbindung mit der abhängigen Variable *Anzahl regelkonformer Lösungen* könnten Deckeneffekte ein Signifikantwerden der Faktoren *Präsentationsform* und *Informationsart* verhindert haben, d.h. es gab zu viele Probanden, die

in allen vier Experimentalbedingungen durchgängig regel-konforme Lösungen hervorgebracht haben. Das Auftreten von Deckeneffekten bei der Variable *Anzahl regel-konformer Lösungen* hätte zudem Auswirkungen auf die Verarbeitungseffizienz, nicht aber auf die Bearbeitungszeit. Deckeneffekte können entweder auftreten, wenn die Schwierigkeit der zu bearbeitenden Aufgabe zu gering ist und/oder wenn die die Aufgabe bearbeitenden Individuen zu gut sind. Die geringe Aufgabenschwierigkeit lässt sich nach unserer Auffassung dadurch erklären, dass nur die Stärke einer Ursache beurteilt werden sollte und dass den Probanden die Informationen über Ursache-Wirkungs-Zusammenhänge durchgängig in einer summarischen Form dargeboten wurden. Eine summarische Darbietungsart reduziert die Komplexität der Aufgabe und damit die Anforderungen für das Gedächtnis (Shaklee & Mims, 1982). Durch eine reduzierte Gedächtnisbeanspruchung werden die Probanden eher in die Lage versetzt, bei der Beurteilung der Stärke von kausalen Zusammenhängen komplexe regel-basierte Strategien heranzuziehen und regel-konforme Lösungen zu produzieren.

Die Schwierigkeit einer Aufgabe kann nicht gänzlich unabhängig von den Individuen beurteilt werden, die sie zu bearbeiten haben. In Kapitel 8.7 wird näher auf Charakteristika der Stichprobe in Experiment 1 eingegangen.

Die in Experiment 1 als Folge von geringer Aufgabenkomplexität und/oder Stichprobenselektion vermuteten Deckeneffekte haben möglicherweise einen sichtbaren Einfluss der Faktoren *Präsentationsform* und *Informationsart* auf die *Anzahl regel-konformer Lösungen* verhindert. Eine andere Interpretation der Ergebnisse aus Experiment 1 geht nicht davon aus, dass Deckeneffekte den möglichen Einfluss des Faktoren *Präsentationsform* und *Informationsart* verschleiern, sondern dass die Unterschiede zwischen den Präsentationsformen tabellarische Vierfeldertafel und Stapeldiagramm einerseits und den beiden Informationsarten Häufigkeiten und Wahrscheinlichkeiten andererseits im Zusammenhang mit der gestellten Aufgabe zu schwach oder zu wenig bedeutsam sind.

In Verbindung mit dem Einfluss der beiden Präsentationsformen Tabelle und Diagramm ist aus früheren Untersuchungen (s. Coll, 1992; Jacobs, 1994; Meyer, 2000; Meyer, Shinar & Leiser, 1997; Wilson & Addo, 1994a) bekannt, dass es entscheidend von der gestellten Aufgabe abhängt, welche Präsentationsform besser abschneidet (Vessey, 1991). Während Tabellen das genaue Ablesen von Datenpunkten und das Anfertigen von Berechnungen erleichtern (Coll, 1992; Meyer, 2000; Meyer, Shinar & Leiser, 1997; Wilson & Addo, 1994a), können Diagramme visuelle Vergleiche und das In-Beziehung-Setzen verschiedener Datenpunkte vereinfachen (Coll, 1992; Jacobs, 1990, 1994; Spence, 1990; Spence & Lewandowsky, 1991). Die in Experiment 1 gefundenen geringen Unterschiede zwischen

Tabellen und Stapeldiagrammen lassen sich also möglicherweise auch darauf zurückführen, dass die Beurteilung der Stärke von kausalen Zusammenhängen sowohl Ableseaufgaben als auch visuelle Vergleichsaufgaben beinhaltet, was letztlich darauf hinaus läuft, dass sich die begünstigenden Einflüsse von Tabelle und Stapeldiagramm gegenseitig neutralisiert haben könnten. Möglich ist es aber auch, dass die von den Probanden vorzunehmenden visuellen Vergleichsaufgaben einfach nicht komplex genug waren und dass sich deshalb Tabelle und Stapeldiagramm nicht signifikant unterscheiden (Jacobs, 1994; Spence und Lewandowsky, 1991; Wilson & Addo, 1994a).

Der Einfluss des Faktors *Informationsart* auf die *Anzahl regel-konformer Lösungen* ist schwer zu interpretieren. Zwar war der Einfluss des Faktors *Informationsart* nicht signifikant, doch lag er trotz der zu beobachtenden Deckeneffekte immerhin in Trendnähe. Eine höhere mittlere Anzahl regel-konformer Lösungen unter den beiden Wahrscheinlichkeitsbedingungen war von uns vorhergesagt worden, so dass es aus unserer Sicht Sinn macht, die etwas höhere mittlere Anzahl regel-konformer Lösungen unter der Wahrscheinlichkeitsbedingung auf eine im Vergleich zur Häufigkeitsbedingung geringere Gedächtnisbeanspruchung zurückzuführen.

*Bearbeitungszeit*: Im Zusammenhang mit dem Einfluss der unabhängigen Variablen auf die Bearbeitungszeit kam diesmal nicht der Faktor *Informationsart* in Trendnähe, sondern der Faktor *Präsentationsform*. Anders als bei der Anzahl regel-konformer Lösungen lassen sich die geringen Einflüsse auf die Bearbeitungszeit nicht auf Deckeneffekte zurückführen, da den Probanden keine Zeitgrenze vorgegeben wurde. Das Ausbleiben signifikanter Effekte des Faktors *Informationsart* war für uns doch überraschend und kann von uns nur durch ausgeprägte individuelle Unterschiede hinsichtlich der Bearbeitungsdauer erklärt werden, deren Einfluss denjenigen des Faktors *Informationsart* bei Weitem zu überwiegen scheint.

Auch der trendnahe Einfluss der unabhängigen Variable *Präsentationsform* war schwächer als erwartet, kann aber als Hinweis darauf verstanden werden, dass die zur Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen vorgegebenen Informationen aus einem Stapeldiagramm schneller verarbeitet werden als aus einer Tabelle. Da dieser schwache Einfluss in der vorhergesagten Richtung lag, erscheint uns die Deutung zulässig, dass im Stapeldiagramm visuelle Vergleiche unmittelbar nahe gelegt werden und so das Zueinander-in-Beziehung-setzen der vier Zellgrößen, sowie von Erfolgs- und Basisrate schneller erfolgt als in einer Tabelle. Möglicherweise lassen sich die etwas schnelleren Bearbeitungszeiten von Probanden, die mit einem Stapeldiagramm gearbeitet haben, auf eine Verlagerung von komplexen Informationsverarbeitungsprozessen auf die Wahrnehmungsebene (Fiedler, 2000;

*Wickens & Carswell, 1995*) oder auf ein Ermöglichen perzeptueller Inferenzen (*Larkin & Simon, 1987*) zurückführen.

*Verarbeitungseffizienz*: In Verbindung mit dem Einfluss von *Präsentationsform* und *Informationsart* auf die abhängige Variable *Verarbeitungseffizienz* wurden weder die beiden Hauptfaktoren, noch die Interaktion signifikant. Da bei der Berechnung der abhängigen Variable *Verarbeitungseffizienz* für jeden Probanden die Anzahl regel-konformer Lösungen zur Bearbeitungszeit ins Verhältnis gesetzt wird, lassen sich die Ergebnisse für die Verarbeitungseffizienz aus den Ergebnissen der beiden anderen abhängigen Variablen herleiten. Die Einflüsse der Faktoren *Präsentationsform* und *Informationsart* auf die Anzahl regel-konformer Lösungen und die Bearbeitungszeit waren entweder jeweils zu schwach, oder es wurde im Fall der abhängigen Variable *Anzahl regel-konformer Lösungen* durch das Auftreten von Deckeneffekten ein Sichtbarwerden des Einflusses der beiden unabhängigen Variablen verhindert.

#### *Vergleich der vier Experimentalbedingungen mit der Kontrollbedingung Protokollliste*

*Anzahl regel-konformer Lösungen*: Unter den beiden Wahrscheinlichkeitsbedingungen *SW* und *TW* erzielten die Probanden eine signifikant höhere mittlere Anzahl regel-konformer Lösungen als unter der Bedingung *PL*. Dieser Unterschied war von uns vorausgesagt worden und spricht in unseren Augen dafür, dass unter der Bedingung *PL* die Gedächtnisbeanspruchung für die Probanden höher ist als unter den beiden Wahrscheinlichkeitsbedingungen *SW* und *TW*. Die Auswirkungen der unterschiedlichen Gedächtnisbeanspruchung unter der Listenbedingung und unter den beiden Wahrscheinlichkeitsbedingungen bestehen vermutlich darin, dass es unter der Listenbedingung mehr fehlerhafte oder verzerrte Kausalurteile gebildet werden (*Fiedler, 2000*; *Kao & Wasserman, 1993*; *Shaklee & Mims, 1982*), zum anderen besteht die Möglichkeit, dass die Probanden unter der Bedingung *PL* von einfacheren Strategien Gebrauch machen (*Kleinmuntz & Schkade, 1993*; *Shaklee & Mims, 1982*) als unter den beiden Wahrscheinlichkeitsbedingungen. In Kapitel 8.6 wird die Strategiewahl der Probanden in den verschiedenen Versuchsbedingungen näher analysiert.

Unter den beiden Häufigkeitsbedingungen *SH* und *TH* haben die Probanden zwar jeweils ebenfalls eine höhere mittlere Anzahl regel-konformer Lösungen erzielt, doch waren die Unterschiede zwischen den Bedingungen *SH* und *TH* einerseits und der Bedingung *PL* andererseits nicht signifikant. Immerhin erreichte zumindest der Unterschied zwischen den Bedingungen *SH* und *PL* Trendnähe. Offen bleibt die Frage, warum die Unterschiede zwischen den beiden Häufigkeitsbedingungen und der Listenbedingung nicht wie erwartet

signifikant wurden, sondern lediglich in die vorhergesagte Richtung gingen. Im Rahmen der Diskussion zum Vergleich der vier Experimentalbedingungen *SH*, *SW*, *TH* und *TW* untereinander haben wir die Vermutung geäußert, dass Deckeneffekte ein Signifikantwerden der Faktoren *Präsentationsform* und *Informationsart* verhindert haben. Deckeneffekte als Folge von Stichprobenselektion oder Einfachheit der Aufgabe können auch bei einigen Probanden unter der Listenbedingung aufgetreten sein. Möglicherweise wurde in der Listenbedingung den Probanden das Hervorbringen regel-konformer Lösungen dadurch einfacher gemacht, dass es sich bei den Protokolllisten, die in Experiment 1 eingesetzt wurden, nicht um ungeordnete und nur minimal strukturierte Auflistungen handelte, sondern um geordnete, übersichtliche und relativ gut strukturierte Listen mit drei Spalten (s. Abbildung 7). Die tabellarisch und übersichtlich angeordneten Listenformen in Experiment 1 könnten dazu beigetragen haben, dass auch unter der Bedingung *PL* die Gedächtnisbeanspruchung zu einem gewissen, wenn auch im Vergleich zu den vier Experimentalbedingungen *SH*, *SW*, *TH* und *TW* geringeren Umfang reduziert wurde, weil unter der Listenbedingung die Prozesse von Datenzuordnung und –aggregation, eventuell auch Prozesse bei der Urteilsbildung, vereinfacht wurden.

Mit Hilfe der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) lassen sich die höhere mittlere Anzahl regel-konformer Lösungen unter den beiden Wahrscheinlichkeitsbedingungen und ansatzweise auch unter der Bedingung *SH* so erklären, dass der Aufwand oder die mentale Beanspruchung bei der Aufgabenbearbeitung im Vergleich zur Listenbedingung reduziert sind. Die Ergebnisse sprechen dafür, dass sich im Vergleich zur Listenbedingung unter den beiden Wahrscheinlichkeitsbedingungen die Schwierigkeit der kausalen Urteilsbildung reduziert und die Probanden über eine signifikant höhere kognitive Kapazität verfügen, um Informationen, die zur Bildung regel-konformer Lösungen notwendig sind, aus dem Arbeits- und Referenzgedächtnis zu integrieren.

*Bearbeitungszeit:* In allen vier Experimentalbedingungen *SH*, *SW*, *TH* und *TW* wurden signifikant schnellere mittlere Bearbeitungszeiten als unter der Kontrollbedingung *PL* erzielt. Dieses Ergebnis entspricht den Erwartungen, überraschend war nur, dass die Probanden unter der Bedingung *SH* – und nicht unter der Bedingung *SW* – am wenigsten Zeit für Informationsverarbeitung und Urteilsbildung benötigten. Unter der Bedingung *SH* kam es im Vergleich zu den anderen drei Experimentalbedingungen zu erheblichen individuellen Schwankungen, abgesehen von der Listenbedingung befanden sich sowohl die schnellste als auch die langsamste Versuchsperson in dieser Bedingung. Insgesamt 19 Probanden bearbeiteten die Aufgaben schneller als 12 Minuten, davon waren allein 11 der Bedingung *SH*

zugeordnet, die restlichen 8 Probanden entstammten den drei anderen Experimentalbedingungen *SW* (5 Probanden), *TH* (1 Proband) und *TW* (2 Probanden). Weitere Analysen ergaben, dass nur 5 der 11 schnellsten Probanden unter der Bedingung *SH* durchgängig eine regel-basierte Strategie einsetzten; von den restlichen schnellsten Probanden machten hingegen alle 8 Probanden durchgängig von regel-basierten Strategien Gebrauch. Dieser für uns nicht nachvollziehbare Befund legt die Schlussfolgerung nahe, dass unter der Bedingung *SH* einige Probanden bei der Urteilsbildung nachlässig vorgegangen sind, weil sie die Aufgaben schnell und nicht regel-konform gelöst haben.

Ansonsten entsprachen die Ergebnisse für die Bearbeitungszeiten den Erwartungen, so dass davon ausgegangen werden kann, dass der zeitliche Aufwand für die Bearbeitung der Aufgabe unter den vier Experimentalbedingungen signifikant kürzer war als unter der Listenbedingung. Die längeren Bearbeitungszeiten unter der Bedingung *PL* entstehen in unseren Augen dadurch, dass den Probanden unter dieser Bedingung weniger Aufgaben abgenommen werden, als unter den vier Experimentalbedingungen. Somit sind die Probanden unter der Listenbedingung einer höheren mentalen Beanspruchung ausgesetzt sind als unter den vier Experimentalbedingungen. Nach der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) wird der höhere Zeitaufwand für die Berechnung der Stärke eines Ursache-Wirkungs-Zusammenhangs von den Probanden unter der Listenbedingung antizipiert, so dass sie auf einfachere Strategien zurückgreifen. Alternativ kann angenommen werden, dass den Probanden unter der Listenbedingung mehr Fehler bei der Aufgabenbearbeitung unterlaufen, weil ihnen weniger Arbeitsschritte durch die Präsentationsform abgenommen werden.

*Verarbeitungseffizienz:* Die Ergebnisse für die abhängige Variable *Verarbeitungseffizienz* lassen sich aus den Ergebnissen für die abhängigen Variablen *Anzahl regel-konformer Lösungen* und *Bearbeitungszeit* herleiten. Unter den drei Experimentalbedingungen *SH*, *SW* und *TW* wurde signifikant effizienter gearbeitet als unter der Listenbedingung, für den Unterschied zwischen den Bedingung *TH* und der Kontrollbedingung *PL* ergab sich ein Trend. Auch diese Ergebnisse entsprechen im Wesentlichen den Erwartungen. Die höhere Verarbeitungseffizienz unter den vier Experimentalbedingungen legt nahe, dass es nicht gleichgültig ist, ob ein Stapeldiagramm oder eine Protokollliste als Präsentationsform eingesetzt werden.

Die in Experiment 1 gefundenen Unterschiede zwischen den vier Experimentalbedingungen und der Protokollliste enthalten aus unserer Sicht eine Teilerklärung dafür, warum es in älteren Untersuchungen, in denen eine summarische und eine einzelfallweise Darbietung miteinander verglichen wurden, zu heterogenen Ergebnissen gekommen ist. So war in einigen

Studien (Hagmayer, 2001; Van Hamme & Wasserman, 1993) kein Unterschied zwischen einer Protokollliste und einer einzelfallweisen Darbietung auszumachen, während in anderen Untersuchungen (Arkes & Harkness, 1983; Buehner et al., 2003; Lober & Shanks, 2000; Ward & Jenkins, 1965; White 2003c), in denen Tabellen, Tortendiagramme oder übersichtlich gestaltete Bildertafeln mit einer einzelfallweisen Darbietung verglichen wurden, deutliche Unterschiede in Abhängigkeit von der Darbietungsart gefunden werden konnten. Wir vermuten, dass die mentale Beanspruchung in Untersuchungen, in denen als summarische Darbietungsform eine Protokollliste mit einer einzelfallweisen Darbietung verglichen wurde, bei beiden Darbietungsarten so hoch war, dass keine nennenswerten Unterschiede auftraten. Unterschiede zwischen den verschiedenen Darbietungsarten treten dann auf, wenn die mentale Beanspruchung bei summarischer Darbietung durch eine geeignete Präsentationsform stark genug reduziert werden konnte. So haben beispielsweise die Probanden in den Untersuchungen von Ward & Jenkins (1965) und Arkes & Harkness (1983) bei summarischer Darbietung mit einer tabellarischen Vierfeldertafel häufiger von der  $\Delta P$ -Regel Gebrauch gemacht als bei einzelfallweiser Darbietung.

Im Allgemeinen wird innerhalb der Literatur zur kausalen Urteilsbildung von zwei Darbietungsarten ausgegangen: Entweder werden die Daten einzelfallweise oder sie werden summarisch dargeboten. Im Zusammenhang mit einer Protokollliste wurde in der Regel davon ausgegangen, dass die in ihr enthaltenen Informationen summarisch dargeboten werden. Betrachtet man aber die Verwendungsmöglichkeiten einer Protokollliste bei der kausalen Urteilsbildung einmal genauer, so ist es unklar, ob die in ihr präsentierten Informationen einzelfallweise oder summarisch dargeboten werden, weil eine Protokollliste keine Summenwerte enthält, sondern eine Auflistung von zeitgleich dargebotenen Einzelfällen. Während bei einer summarischen Darbietung kein assoziativer Mechanismus aktiviert werden kann (Shanks, 1991), ist es bei Darbietung von Informationen in einer Protokollliste nicht ausgeschlossen, dass die Probanden die Liste einzelfallweise durcharbeiten und ihre Eindrücke von Fall zu Fall durch einen assoziativen Mechanismus miteinander verknüpfen (Van Hamme & Wasserman, 1993; White 2003c, 2004). Alternativ zu einer assoziativen Vorgehensweise, können die Probanden, welche die Kausalinformationen in einer Protokollliste zeitgleich dargeboten bekommen, auch eigenständig Summenbildungen vornehmen, die im Unterschied zur einzelfallweisen Informationsdarbietung so lange überprüft werden können, wie den Probanden die Liste vorliegt.



Für angemessener halten wir eine Sichtweise, in der davon ausgegangen wird, dass in einer Protokollliste Informationen über Ursache-Wirkungs-Zusammenhänge sowohl summarisch als auch einzelfallweise dargeboten werden. Im Unterschied zur traditionellen einzelfallweisen Darbietung werden in einer Protokollliste die Einzelfälle nicht zeitversetzt nacheinander, sondern zeitgleich aneinandergereiht dargeboten. Weil Informationen über Ursache-Wirkungs-Zusammenhänge in Protokolllisten nicht nur summarisch oder einzelfallweise dargeboten werden, halten wir es wie White (2003c, 2004; s.a. Griffiths & Tenenbaum, 2005) für sinnvoller, die Informationsdarbietung in Protokolllisten als eine eigene Darbietungsart zu konzipieren, in der die Einzelfälle zeitgleich nacheinander aufgelistet sind und die Probanden die Möglichkeit haben, in der Liste vor- und zurückzugehen und Prüfprozesse durchzuführen. Demnach gäbe es also eine summarische Darbietung, eine Auflistung von zeitgleich dargebotenen Einzelfällen, sowie eine traditionelle Einzelfalldarbietung, bei der die einzelnen Informationen über Ursache-Wirkungs-Zusammenhänge zeitlich nacheinander dargeboten werden.

Von Shanks (1991) wurde die These vertreten, dass in Abhängigkeit vom Situationstyp unterschiedliche Verarbeitungsmechanismen aktiviert werden. In Experiment 1 waren die Probanden in allen Untersuchungsbedingungen einem beschriebenen Situationstypus ausgesetzt. Insofern muss nach Shanks (1991) davon ausgegangen werden, dass in Abhängigkeit von den beiden Faktoren *Präsentationsform* und *Informationsart* keine unterschiedlichen Verarbeitungsmechanismen aktiviert werden. Die Ergebnisse für die abhängige Variable *Anzahl regel-konformer Lösungen* legen zwar nahe, dass die meisten Probanden in den vier Experimentalbedingungen und auch einige der Probanden in der Kontrollbedingung Protokollliste von regel-basierten Strategien Gebrauch gemacht haben, doch liefern sie keinen genaueren Aufschluss darüber, welches Individuum unter welcher Versuchsbedingung von welcher regel-basierten oder nicht regel-basierten Strategie Gebrauch gemacht hat. Die Ergebnisse aus einer Reihe von Studien (Anderson & Sheu, 1995; Inhelder & Piaget, 1958; Kao & Wasserman, 1993; Lober & Shanks, 2000; Perales & Shanks, 2003; Shaklee & Elek, 1988; Shaklee & Hall, 1983; Shaklee & Mims, 1981; Shaklee & Tucker, 1980; Ward & Jenkins, 1965; Wasserman, 1990; Wasserman, Chatlosh & Neunaber, 1983; Wasserman & Shaklee, 1984; White, 1995, 1998, 2000a) legen nahe, dass sich Individuen nicht nur in unterschiedlichen, sondern selbst in ein- und derselben Untersuchungsbedingung im Hinblick auf ihre Verarbeitungsmechanismen oder ihre Strategiewahl bei der Beurteilung der Stärke von kausalen Zusammenhängen unterscheiden können.

## 8.6 Häufigkeitsanalysen zum Einfluss der Versuchsbedingungen auf die Strategiewahl

### 8.6.1 Ergebnisse

Die bisherigen Ergebnisse lassen kaum darüber Aufschlüsse zu, ob in Abhängigkeit von den unabhängigen Variablen unterschiedliche Informationsverarbeitungsmechanismen aktiviert werden oder nicht. Deshalb haben wir versucht genau zu analysieren, von welchen Probanden welche Strategien eingesetzt wurden und ob die Strategiewahl in Abhängigkeit von der jeweiligen Versuchsbedingung variiert. Weiter sollen die Analysen Aufschluss darüber geben, ob Abweichungen von regel-basierten Strategien darauf beruhen, dass von den Probanden einfachere Strategien eingesetzt wurden (Kleinmuntz & Schkade, 1993; Shaklee & Mims, 1982), oder ob es doch eher zu Urteilsfehlern gekommen ist (Fiedler, 2000; Kao & Wasserman, 1993; Shaklee & Mims, 1982).

Um die Strategien identifizieren zu können, von denen die einzelnen Probanden Gebrauch machten, wurden für jeden Probanden anhand der abgegebenen Wirksamkeitseinschätzungen, sowie der schriftlichen und mündlichen Nachbefragung ermittelt, welche Strategie(n) von ihnen eingesetzt wurden. Auf der Basis der eingesetzten Strategien wurden folgende Kategorien für Strategietypen gebildet, denen die einzelnen Probanden zugeordnet wurden:

1. Strategie unklar
2. Wechsler unklar – ab-Regel
3. ab-Regel
4. Wechsler zu einer regel-basierten Strategie
5. durchgängig regel-basierte Strategie
6. fehlerhaft eingesetzte regel-basierte Strategie

Tabelle 1 zeigt, wie sich die einzelnen Probanden insgesamt und in Abhängigkeit von den fünf Versuchsbedingungen *SW*, *TW*, *SH*, *TH* und *PL* auf die sechs Kategorien verteilen:

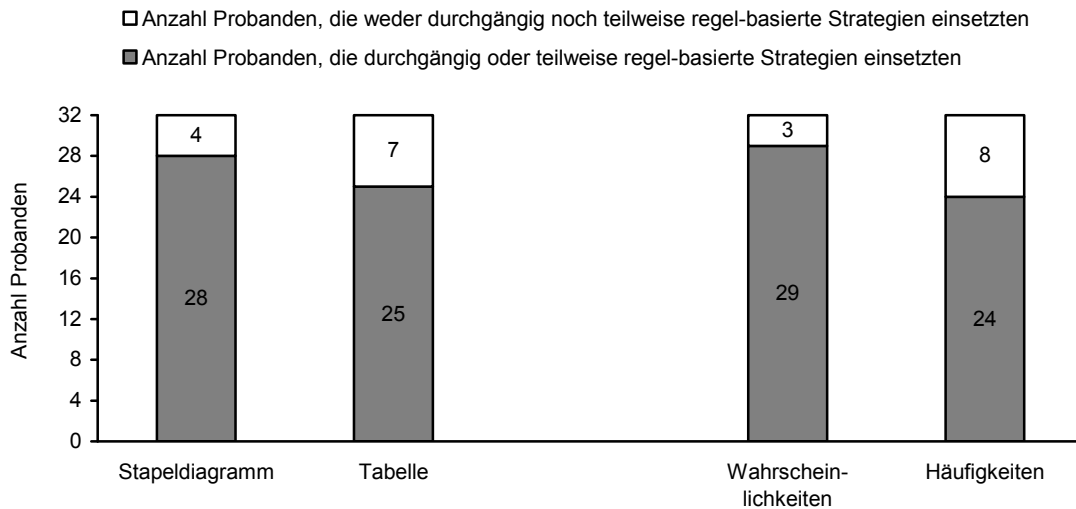
**Tabelle 1:** Experiment 1: Anzahl der Probanden in Abhängigkeit von Strategietyp und Versuchsbedingung (in Klammern Prozentangaben). Zellohäufigkeiten > 5 sind **fett** markiert.

	Strategietyp						Gesamt
	unklar	Wechsler ab-Regel	ab-Regel	Wechsler regel-basiert	regel- basiert	fehlerhaft regel-basiert	
<b>Versuchsbedingung</b>							
SW	0	1	0	2	<b>12</b>	1	16
TW	2	0	0	2	<b>11</b>	1	16
SH	0	0	3	4	<b>9</b>	0	16
TH	2	2	1	3	<b>7</b>	1	16
PL	2	2	2	3	<b>6</b>	1	16
<b>Gesamt</b>	6 (7,5 %)	5 (6,25 %)	6 (7,5 %)	14 (17,5 %)	45 (56,25 %)	4 (5 %)	80 (100 %)

Da der Anteil an Zellen mit erwarteten Häufigkeiten kleiner fünf 20 % deutlich übersteigt, wurde auf einen  $\chi^2$  – Test zur Prüfung der Annahme, ob die beiden Variablen Versuchsbedingung und Strategietyp voneinander unabhängig sind oder nicht, verzichtet (Bortz, 1999). Der Anteil an Probanden, die ausschließlich regel-basierte Strategien eingesetzt haben, ist mit 56,25 % sehr hoch, zählt man noch diejenigen Probanden hinzu, die während der Versuchsdurchführung zu einer regel-basierten Strategie übergewechselt oder eine regel-basierte Strategie fehlerhaft eingesetzt haben, so haben insgesamt 78,75 % ausschließlich oder teilweise regel-basierte Strategien eingesetzt. Die Zahlen in Tabelle 1 bestätigen die zuvor aufgestellte Vermutung von Deckeneffekten während der Versuchsdurchführung.

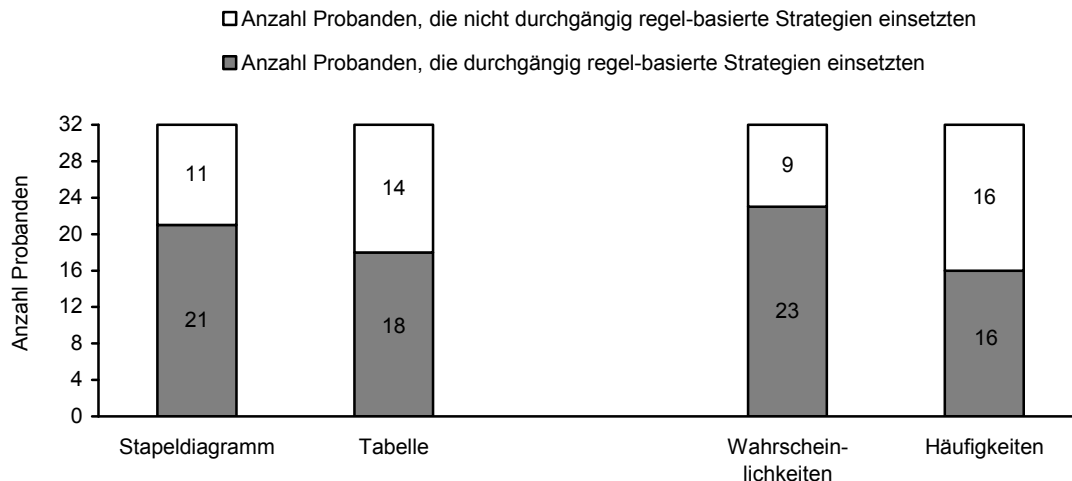
Zur Überprüfung der Hypothesen H1d und H2d wurden nun die 64 Probanden aus den vier Experimentalbedingungen *SW*, *TW*, *SH* und *TH* binär hinsichtlich des Strategietyps verteilt. Die erste Gruppe setzte sich aus den 53 Probanden zusammen, die teilweise oder durchgängig regel-basierte Strategien eingesetzt haben. Die zweite Gruppe bestand aus den restlichen 11 Probanden, die keinen Gebrauch von regel-basierten Strategien gemacht haben. Hypothese H1d sagt vorher, dass die Anzahl der Probanden, die regel-basierte Strategien einsetzen, höher ist, wenn die Probanden die Wirksamkeiten an Hand von Stapeldiagrammen ermitteln als wenn die Probanden die Wirksamkeiten an Hand von Tabellen ermitteln. Nach Hypothese H2d wird erwartet, dass die Anzahl von Probanden, die regel-basierte Strategien einsetzen höher ist, wenn die Informationen über Ursache-Wirkungs-Zusammenhänge als Wahrscheinlichkeiten präsentiert werden, als wenn sie als Häufigkeiten präsentiert werden. Mit der Hilfe eines  $\chi^2$ -Tests und des exakten Tests nach Fisher wurde untersucht, wie sich die Anzahl der Probanden in Abhängigkeit vom Strategietyp, sowie von den beiden Faktoren *Präsentationsform* und *Inhaltsart* verteilten. Abbildung 12 veranschaulicht, wie sich die

Probanden nach Strategietyp in Abhängigkeit von *Präsentationsform* und *Inhaltsart* verteilen:



**Abbildung 12:** Experiment 1: Verteilung von Probanden, die regel-basierte oder keine regel-basierten Strategien eingesetzt haben, in Abhängigkeit von *Präsentationsform* (Stapeldiagramm vs. Tabelle) und *Inhaltsart* (Wahrscheinlichkeiten vs. Häufigkeiten).

Während sich die beobachteten Häufigkeiten unabhängig von der *Präsentationsform* und dem *Strategietyp* verteilen ( $\chi^2 = .988$ ,  $p = .320$ ; einseitiger P-Wert nach Fisher's exaktem Test = .255), wurde für die Verteilung der Probanden in Abhängigkeit von *Informationsart* und *Strategietyp* immerhin ein Trend erzielt ( $\chi^2 = 2.744$ ,  $p_{\text{asymptotische Signifikanz}} = .098$ ;  $P_{\text{Fisher, einseitig}} = .092$ ). Letzteres bedeutet, dass Probanden, denen Wahrscheinlichkeiten vorgelegt wurden, tendenziell häufiger durchgängig oder teilweise regel-basierte Strategien einsetzten als Probanden, die die Wirksamkeiten an Hand von Häufigkeiten bestimmten. Sehr ähnliche Ergebnisse werden erzielt, wenn nur die 39 Probanden, die ausschließlich durchgängig von regel-basierten Strategien Gebrauch machten, mit den 25 Probanden verglichen werden, die nicht durchgängig regel-basierte Strategien einsetzten. Die sich hieraus und in Abhängigkeit von *Präsentationsform* und *Inhaltsart* ergebenden Häufigkeitsverteilungen zeigt Abbildung 13.



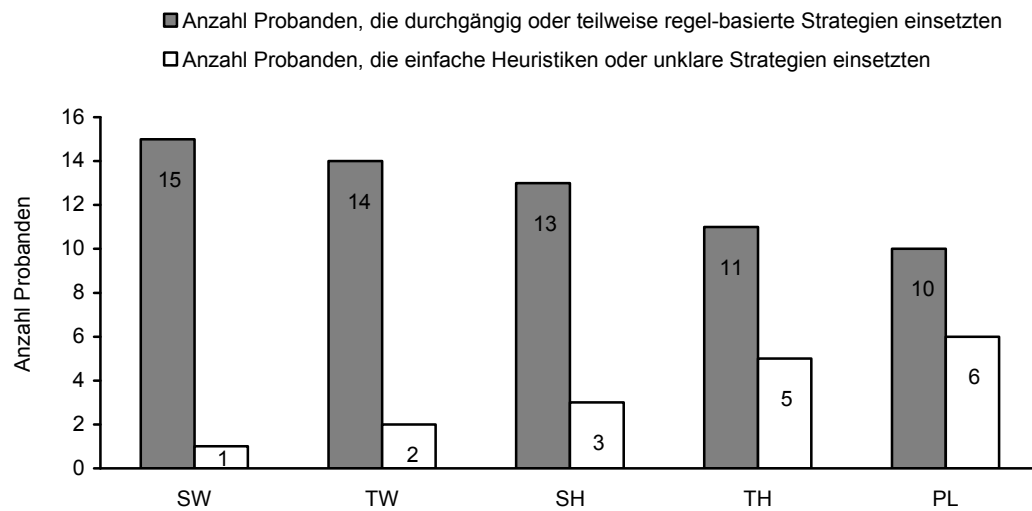
**Abbildung 13:** Experiment 1: Verteilung von Probanden, die entweder durchgängig oder nicht durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von *Inhaltsart* (Wahrscheinlichkeiten vs. Häufigkeiten) und *Präsentationsform* (Stapeldiagramm vs. Tabelle).

Erneut ergaben ein  $\chi^2$ -Test und ein exakter Test nach Fischer den Trend, dass die Anzahl der Probanden, die durchgängig von einer regel-basierten Strategie Gebrauch machten, bei Vorgabe von Wahrscheinlichkeiten höher war als bei der Präsentation von Häufigkeiten ( $\chi^2 = 3.216$ ,  $p_{\text{asymptotische Signifikanz}} = .073$ ;  $p_{\text{Fisher, einseitig}} = .062$ ). Dagegen verteilte sich die Anzahl der Probanden, die entweder durchgängig regel-basierte oder nicht durchgängig regel-basierte Strategien eingesetzt haben, unabhängig von der *Präsentationsform* ( $\chi^2 = .591$ ,  $p_{\text{asymptotische Signifikanz}} = .442$ ;  $p_{\text{Fisher, einseitig}} = .304$ ).

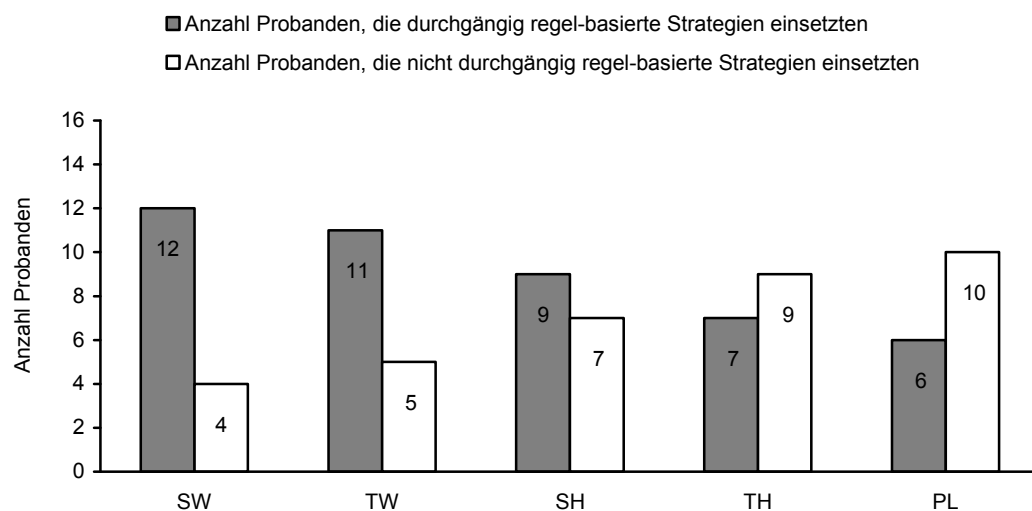
Die bislang ausgewerteten Häufigkeitsanalysen führen dazu, dass Hypothese H2d tendenziell beibehalten werden kann. So bestätigt ein Trend die Annahme, dass unter Wahrscheinlichkeitsbedingungen mehr Probanden von regel-basierten Strategien Gebrauch machen als unter Häufigkeitsbedingungen. Dagegen muss Hypothese H1d verworfen werden, da Probanden, die mit Stapeldiagrammen arbeiteten, nicht signifikant häufiger von regel-basierten Strategien Gebrauch machen als Probanden, denen tabellarische Vierfeldertafeln vorgelegt wurden.

Bei weiteren Häufigkeitsanalysen wurde die Kontrollbedingung *PL* mit einbezogen und untersucht, ob sich die Anzahl von Probanden, die regel-basierte oder keine regel-basierten Strategien eingesetzt haben unabhängig von der Versuchsbedingung verteilt. Abbildung 14 veranschaulicht den Einfluss der Versuchsbedingung auf die Verteilung der Probanden, die durchgängig oder teilweise regel-basierte Strategien eingesetzt haben oder nicht; Abbildung

15 zeigt, wie sich Probanden, die ausschließlich durchgängig regel-basierte Strategien einsetzten oder nicht, in Abhängigkeit von der Versuchsbedingung verteilten.



**Abbildung 14:** Experiment 1: Anzahl Probanden, die durchgängig oder teilweise regel-basierte Strategien eingesetzt haben und von Probanden, die einfache Heuristiken oder unklare Strategien eingesetzt haben, in Abhängigkeit von der Versuchsbedingung.



**Abbildung 15:** Experiment 1: Anzahl Probanden, die durchgängig oder nicht durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von der Versuchsbedingung.

Auch wenn beide Häufigkeitsverteilungen den Eindruck erwecken, dass die Anzahl der Probanden, die regel-basierte Strategien einsetzten, von der Bedingung SW bis zur Bedingung PL ab-, und die Anzahl der Probanden, die keine regel-basierten Strategien einsetzten, zunimmt, so wurde doch keine der durchgeführten Häufigkeitsanalysen zur Untersuchung eines Zusammenhangs zwischen Strategiewahl und Versuchsbedingung signifikant. ( $\chi^2 = 5.320$ ;  $p = .256$  für die Anzahl der Probanden, die nur durchgängig regel-basierte Strategien

einsetzten oder nicht;  $\chi^2 = 6.424$ ,  $p = .170$  für die Anzahl der Probanden, die teilweise oder durchgängig regel-basierte Strategien einsetzten oder nicht. Bei der Häufigkeitsverteilung in Abbildung 14 betrug der Anteil an Zellen mit erwarteten Häufigkeiten kleiner 5 aber 50 %).

Um herauszufinden, welche regel-basierten Strategien genau in Experiment 1 zum Einsatz kamen, wurden die insgesamt 63 Probanden, die teilweise oder durchgängig regel-basierte Strategien eingesetzt haben, weiter ausdifferenziert und daraufhin untersucht, ob sie von der  $\Delta P$ -Regel, der Power PC-Regel oder einer anderen regel-basierten Strategie Gebrauch gemacht haben. Tabelle 2 veranschaulicht, wie viele Probanden in welcher Versuchspersonengruppe welche Art von regel-basierter Strategie angewendet haben. Probanden, die von nicht regel-basierten zu einer oder mehreren regel-basierten Strategie gewechselt haben, wurden danach eingeteilt, von welcher regel-basierten Strategie sie Gebrauch gemacht haben. Im Einzelnen wurden für Tabelle 2 die folgenden Kategorien, denen die Probanden zugeordnet wurden, gebildet:

1. Delta P-Regel, kurz " $\Delta P$ -Regel"
2. Power PC-Theorie, kurz "Power PC"
3. Wechsler Delta P-Regel - Power PC Theorie, kurz " $\Delta P$  & PPC"
4. Erfolgsrate multipliziert mit dem Term "Eins minus Basisrate", kurz " $ER * (1-BR)$ "
5. Wechsler Delta P-Regel -  $ER * (1-BR)$ , kurz " $\Delta P$  &  $ER * (1-BR)$ "

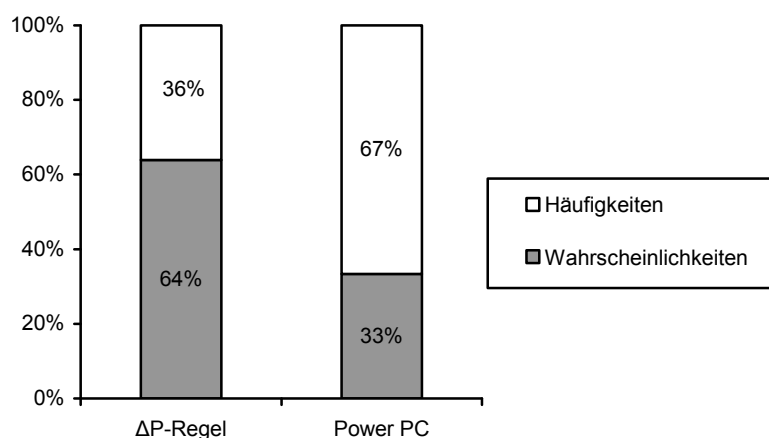
Die Kategorien 4 und 5 wurden von uns neu gebildet, da zwei Naturwissenschaftler während des Experiments von der Formel  $ER * (1 - BR)$  Gebrauch machten. Der Einsatz dieser Strategie wurde von einem mittlerweile habilitierten theoretischen Physiker mit Problemen von sowohl der  $\Delta P$ -Regel als auch der Power PC-Theorie bei extrem hohen oder niedrigen Erfolgs- und Basisraten begründet.

**Tabelle 2:** Experiment 1: Anzahl der Probanden, die durchgängig oder teilweise regel-basierte Strategien eingesetzt haben, in Abhängigkeit von der Art der regel-basierten Strategie und der Versuchsbedingung (in Klammern Prozentangaben).

	Art der regel-basierten Strategie					Gesamt
	$\Delta P$ -Regel	Power PC	$\Delta P$ & PPC	$ER * (1-BR)$	$\Delta P$ & $ER * (1-BR)$	
<b>Versuchsbedingung</b>						
SW	13	2	0	0	0	15
TW	10	2	2	0	0	14
SH	6	5	1	1	0	13
TH	7	3	1	0	0	11
PL	7	1	1	0	1	10
<b>Gesamt</b>	43 (68,3 %)	13 (20,6 %)	5 (7,9 %)	1 (1,6 %)	1 (1,6 %)	63 (100 %)

Erneut wurde auf einen  $\chi^2$  – Test verzichtet, da die Anzahl der Zellen mit erwarteten Häufigkeiten kleiner 5 die 20 % - Marke deutlich überschritten hat (Bortz, 1999). Dennoch veranschaulichen die Ergebnisse, dass die meisten Probanden, die regel-basierte Strategien angewendet haben, ganz oder teilweise von der  $\Delta P$ -Regel Gebrauch gemacht haben (68,3 %). Die Power PC-Theorie wurde von 20,6 % der Probanden ausschließlich oder teilweise eingesetzt, der Anteil der Probanden, die die Strategie ER \* (1-BR) ganz oder teilweise angewendet haben, ist mit 1,6 % als gering einzustufen.

Werden die vier Experimentalbedingungen *SW*, *TW*, *SH* und *TH* ohne die Kontrollbedingung *PL* betrachtet, so scheint die Art der beiden am häufigsten eingesetzten regel-basierten Strategien mit dem Faktor *Informationsart* zu interagieren. Diese Wechselwirkung veranschaulicht Abbildung 16: Von allen 36 Probanden, die in den vier Experimentalbedingungen allein die  $\Delta P$ -Regel als regel-basierte Strategie eingesetzt haben, haben ca. 64 % ( $n = 23$ ) mit Wahrscheinlichkeiten und nur ca. 36 % ( $n = 13$ ) mit Häufigkeiten gearbeitet. Hingegen haben von allen 12 Probanden, die in den vier Experimentalbedingungen allein die Power PC Theorie angewendet haben, nur ca. 33 % ( $n = 4$ ) mit Wahrscheinlichkeiten, aber ca. 67 % ( $n = 8$ ) mit Häufigkeiten gearbeitet. Besonders deutlich tritt die Interaktion zwischen dem Faktor *Informationsart* und der Art der beiden am häufigsten eingesetzten regel-basierten Strategien zu Tage, wenn nur die Stapeldiagramme betrachtet werden. Von allen 7 Probanden, die hier die kausale Stärke  $p$  berechnet haben, haben 5 (über 71 %) Häufigkeiten und nur 2 (knapp 29 %) Wahrscheinlichkeiten analysiert, während es von allen 19 Probanden, die  $\Delta P$  berechnet haben, nur 6 Personen waren (knapp 32 %), die mit Häufigkeiten gearbeitet haben und 13 (über 68 %), die mit Wahrscheinlichkeiten gearbeitet haben.



**Abbildung 16:** Experiment 1: Prozentanteile von Probanden, die entweder Häufigkeiten oder Wahrscheinlichkeiten dargeboten bekommen haben, an der Gesamtzahl der Nutzer von entweder der  $\Delta P$ -Regel oder der Power PC Theorie.



Um feststellen zu können, ob die in Abbildung 16 angedeutete Wechselwirkung zwischen der Informationsart und der Art der eingesetzten regel-basierten Strategie ( $\Delta P$ -Regel oder Power PC-Theorie) statistisch signifikant ist, wurden unter Ausschluss der Kontrollbedingung Protokollliste diejenigen 48 Probanden (das sind 75 % aller 64 Probanden aus den vier Experimentalbedingungen) ausgewählt, die teilweise oder ganz allein von der  $\Delta P$ -Regel oder der Power PC Theorie Gebrauch gemacht haben. Mit diesen Probanden wurden ein  $\chi^2$  – Test und ein exakter Test nach Fischer durchgeführt, die beide einen Trend ergaben, dass bei der Darbietung von Wahrscheinlichkeiten anteilmäßig mehr Probanden  $\Delta P$  berechnen, während bei der Darbietung von Häufigkeiten anteilmäßig mehr Probanden die Power PC-Theorie anwenden ( $\chi^2 = 3.414$ ,  $p_{\text{asymptotische Signifikanz}} = .065$ ;  $p_{\text{Fisher, einseitig}} = .066$ ).

Insgesamt kann festgestellt werden, dass knapp 79 % der Probanden durchgehend oder teilweise von regel-basierten Strategien Gebrauch gemacht haben. Dies bestätigt die Annahme von vorliegenden Deckeneffekten in Experiment 1. Eine weitere Ausdifferenzierung nach Art der eingesetzten regel-basierten Strategie führte zu dem Ergebnis, dass das Kontingenzmaß  $\Delta P$  mehr als dreimal häufiger berechnet wurde als die kausale Stärke  $p$ . Durchgeführte Häufigkeitsanalysen ergaben einen Trend, dass Probanden, die mit Wahrscheinlichkeiten arbeiteten, häufiger regel-basierte Strategien einsetzten als Probanden, die die Wirksamkeiten an Hand von Häufigkeiten zu ermitteln versuchten. Zudem konnte ein Trend ausgemacht werden, dass die  $\Delta P$ -Regel eher von Probanden eingesetzt wurde, denen Wahrscheinlichkeiten dargeboten wurden, während von der Power PC Theorie eher dann Gebrauch gemacht wurde, wenn Häufigkeiten dargeboten wurden.

### 8.6.2 Diskussion

Ein wesentliches Ergebnis aus den durchgeführten Häufigkeitsanalysen besteht darin, dass die Probanden in Abhängigkeit von der Versuchsbedingung keine einheitliche Strategiewahl vorgenommen haben. In allen Versuchsbedingungen gab es Probanden, die regel-basiert vorgegangen sind, und solche, die nicht regel-basiert vorgegangen sind. Unverkennbar ist aber, dass die in Abbildung 15 gezeigten Ergebnisse die Tendenz veranschaulichen, dass der Anteil der Probanden, die eine regel-basierte Strategie gewählt haben, mit zunehmender Gedächtnisbeanspruchung abnimmt. Während unter der Bedingung *SW*, wo die Probanden entsprechend unseren Vorhersagen der geringsten Gedächtnisbeanspruchung ausgesetzt waren, 75 % der Probanden durchgängig von einer regel-basierten Strategie Gebrauch gemacht haben, betrug dieser Anteil unter der Kontrollbedingung *PL*, wo die Probanden nach unserer Auffassung mental am stärksten beansprucht wurden, nur noch 37,5 %.

Besonders hoch ist der Anteil der Probanden, die durchgehend oder teilweise regel-basierte Strategien eingesetzt haben, unter den beiden Wahrscheinlichkeitsbedingungen *SW* und *TW*. Dort haben knapp 91 % der Probanden (29 von 32) zumindest teilweise und 72 % (23 von 32) durchgängig von regel-basierten Strategien Gebrauch gemacht. Unter den beiden Häufigkeitsbedingungen fielen diese Anteile geringer aus, hier haben "nur" 75 % (24 von 32) zumindest teilweise bzw. 50 % (16 von 32) der Probanden durchgängig regel-basierte Strategien eingesetzt.

Der sich aus der Auswertung der Daten in den beiden Abbildungen 12 und 13 ergebende Trend, dass Probanden unter Wahrscheinlichkeitsbedingungen häufiger regel-basiert vorgehen als unter Häufigkeitsbedingungen, bestätigt die varianzanalytischen Ergebnisse zum Einfluss des Faktors *Informationsart* auf die *Anzahl regel-konformer Lösungen*. In unseren Augen sprechen diese Ergebnisse dafür, dass sich unter den beiden Wahrscheinlichkeitsbedingungen die Komplexität der Aufgabe, die Stärke von Kausalzusammenhängen zu ermitteln, für die Probanden tendenziell reduziert hat und sie damit einer geringeren Gedächtnisbeanspruchung als unter den beiden Häufigkeitsbedingungen und der Kontrollbedingung *PL* ausgesetzt waren.

Weiter bestätigen die durchgeführten Häufigkeitsanalysen auch das varianzanalytische Ergebnis, dass Probanden, die mit Stapeldiagrammen arbeiteten, im Vergleich zu Probanden, denen tabellarische Vierfeldertafeln vorgelegt wurden, keine signifikant höhere Anzahl regel-konformer Lösungen erzielten. Der Anteil der Probanden, die durchgängig oder zumindest teilweise regel-basierte Strategien einsetzten, war unter den beiden Bedingungen *SW* und *SH* (65,6 % bzw. 87,5 %) nur unwesentlich höher, als unter den beiden Bedingungen *TW* und *TH* (56,25 % bzw. 78,1 %). Wie bereits im Zusammenhang mit der Diskussion des Einflusses des Faktors *Präsentationsform* auf die Anzahl regel-konformer Lösungen erörtert, scheint es bei der Beurteilung der Stärke von monokausalen Ursache-Wirkungs-Zusammenhängen keine Rolle zu spielen, ob die relevanten Informationen in einer tabellarischen Vierfeldertafel oder in einem Stapeldiagramm präsentiert werden.

Eine eingehende Analyse aller 63 Probanden, die teilweise oder durchgängig regel-basierte Strategien gewählt haben, zeigt, dass die meisten dieser Probanden von der  $\Delta P$ -Regel Gebrauch gemacht haben. Über alle Versuchsbedingungen hinweg, haben von allen Probanden, die regel-basiert vorgegangen sind, über 68 % die  $\Delta P$ -Regel eingesetzt, während nur knapp 21 % die Stärke des Kausalzusammenhänge nach der Power PC-Theorie ermittelten. Interessant war daran aber das Teilergebnis, dass 8 der 13 Probanden (61,5 %), die die kausale Stärke nach der Power PC Theorie zu errechnen schienen, dies unter den

beiden Häufigkeitsbedingungen taten. Unter den beiden Häufigkeitsbedingungen ging immerhin ein Drittel (8 von 24 oder 33,3 %) aller Probanden, die regel-basierte Strategien einsetzten, entsprechend der Power PC-Theorie vor und nur 54,2 % (13 von 24) nach der  $\Delta P$ -Regel. Unter den beiden Wahrscheinlichkeitsbedingungen wählten dagegen nur knapp 14 % (4 von 29 Probanden) eine Vorgehensweise, die der Power PC-Theorie entsprach, während 79 % (23 von 29 Probanden) zu einem Ergebnis gelangten, dass mit den Vorhersagen der  $\Delta P$ -Regel äquivalent war. Im Vergleich zu Wahrscheinlichkeiten scheinen Häufigkeiten die Probanden also stärker zum Einsatz einer Strategie zu animieren, deren Ergebnis mit den Vorhersagen der Power PC-Theorie identisch ist. Dieses Analyseergebnis kann, obwohl es nicht signifikant ist, als Hinweis darauf verstanden werden, dass der Faktor *Informationsart* auf der Ebene der Informationsintegration einen gewissen Einfluss darauf hat, von welcher regel-basierten Strategie Probanden, die regel-basiert vorgehen, Gebrauch machen.

Die Häufigkeitsanalysen aus Experiment 1 sollten auch darüber Aufschluss geben, ob der Strategiewahl der Probanden sowohl über alle Versuchsbedingungen hinweg als auch innerhalb einer einzigen Versuchsbedingung ein einheitlicher oder mehrere unterschiedliche kognitive Verarbeitungsmechanismen zugrunde liegen. Das Ergebnis, dass die Mehrheit unserer Probanden die Stärke der Kausalzusammenhänge in allen fünf Versuchsbedingungen sowohl entsprechend den Vorhersagen der  $\Delta P$ -Regel als auch entsprechend den Vorhersagen der Power PC Theorie ermittelt haben, bestätigt Ergebnisse aus anderen Studien (Lober & Shanks, 2000; Perales & Shanks; 2003) und spricht sowohl für die von uns angenommene hohe Bedeutung von individuellen Unterschieden (s.a. Perales & Shanks, 2003), als auch für die Annahme, dass es keinen einheitlichen Verarbeitungsmechanismus bei der Beurteilung der Stärke von Kausalzusammenhängen gibt.

Die Annahme, dass kein einheitlicher Verarbeitungsmechanismus bei der Beurteilung der Stärke von Kausalzusammenhängen existiert, steht nicht im Widerspruch zu dem in Abbildung 5 (S. 128xx) grundsätzlich skizzierten Ablauf von kognitiven Prozessen bei der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen. Sie beinhaltet auch nicht die Annahme, dass bei unseren Probanden in Experiment 1 assoziative Lernmechanismen aktiviert worden sind. Wir vermuten stattdessen, dass bei der Beurteilung der Stärke von Kausalzusammenhängen – ebenso wie bei anderen Informationsverarbeitungsprozessen auch – Informationen wahrgenommen und im Arbeitsgedächtnis gespeichert und bearbeitbar gemacht werden. Gemeinsam mit bereits vorhandenem und abrufbarem Wissen aus dem Referenzgedächtnis werden die im Arbeitsgedächtnis bereit gehaltenen Informationen über Ursache-Wirkungs-Zusammenhänge in einem informationsintegrierenden Mechanismus und

je nach Beschaffenheit des im Referenzgedächtnis aktivierten Wissens regel-basiert, heuristisch oder intuitiv-schätzend zu einem Urteil über die Stärke eines Kausalzusammenhangs weiter verarbeitet. Die Annahme unterschiedlicher Verarbeitungsmechanismen beruht nun darauf, dass sich die Probanden sowohl hinsichtlich des im Referenzgedächtnis abgespeicherten Wissens, als auch hinsichtlich personenbezogener Faktoren unterscheiden und deshalb unterschiedliche informationsintegrierende Mechanismen oder Strategien aktiviert werden. Moduliert wird dieser Vorgang durch eine individuell unterschiedlich ausgeprägte mentale Beanspruchung während der Aufgabenbearbeitung. Die tatsächliche mentale Beanspruchung eines Individuums resultiert aus dem Wechselspiel zwischen aufgaben-, medien- und personenbezogenen Faktoren. Wird beispielsweise die mentale Beanspruchung durch eine Manipulation von medien- und aufgabenbezogenen Faktoren reduziert, so führt dies dazu, dass sich der Anteil der Probanden, die eine regel-basierte Vorgehensweise einschlagen, kontinuierlich erhöht. Auf der Basis der Ergebnisse aus Experiment 1 kann davon ausgegangen werden, dass bei der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen die Informationsverarbeitung und der kognitive Apparat unter den beiden Wahrscheinlichkeitsbedingungen SW und TW am wenigsten und unter der Kontrollbedingung PL am meisten beansprucht wurde. Dementsprechend konnten die Probanden unter der Listenbedingung weniger regel-konforme Lösungen hervorbringen, weil sie mehr Urteilsfehler begingen und häufiger von einfacheren Strategien bei der Urteilsbildung Gebrauch machten. Bei den beiden Häufigkeitsbedingungen SH und TH scheint die Gedächtnisbeanspruchung etwas geringer zu sein als unter der Listenbedingung, aber etwas höher als unter den beiden Wahrscheinlichkeitsbedingungen.

Die durchgeführten Häufigkeitsanalysen geben trotz der beobachteten Deckeneffekte im begrenzten Umfang Auskunft auf die Frage, warum eine Probanden-Minderheit (21 von 80) von den Vorhersagen regel-basierter Strategien abweicht: Zum einen lassen sich die Abweichungen dadurch erklären, dass regel-basierte Strategien von 4 Probanden fehlerhaft eingesetzt wurden, zum anderen dadurch, dass 11 Probanden durchgängig oder teilweise auf eine bekannte einfachere Strategie (ab-Regel) zurückgegriffen haben. Dabei konnte anhand der Daten und Selbstauskünfte dieser von einer regel-basierten Strategie abweichenden Probanden nicht immer zweifelsfrei entschieden werden, ob die betreffenden Probanden fehlerhafte bzw. verzerrte Kausalurteile gebildet haben (Fiedler, 2000; Kao & Wasserman, 1993; Shaklee & Mims, 1982), oder ob sie aufgrund der subjektiv wahrgenommenen Aufgabenkomplexität auf andere, einfachere und ungenauere Strategien zurückgegriffen haben (Kleinmuntz & Schkade, 1993; Shaklee & Mims, 1982). Bei 6 Probanden waren die

eingesetzten Strategien gänzlich unklar, aus den eingeholten Selbstausskünften ging aber hervor, dass diese Probanden ihre Beurteilung der Stärke von Kausalzusammenhängen durchgängig auf der Basis von Schätzungen vorgenommen und ihren Schätzungen keine Berechnungen zugrunde gelegt haben.

25 der insgesamt 80 Probanden (ca. 31 %) gaben an, dass sie ihre Strategie während des Experiments gewechselt hätten. Bei 14 dieser 25 Probanden stimmten die Ergebnisse teilweise mit den Vorhersagen einer regel-basierten Strategie überein, so dass sie als "Wechsler regel-basiert" klassifiziert wurden (s.a. Tabelle 1). Von den 6 Probanden die von einer regel-basierten Strategie zu einer anderen wechselten (also durchgängig regel-basiert vorgehen), wechselten 5 von einer Ermittlung der Stärke des Kausal-Zusammenhangs entsprechend den Vorhersagen der  $\Delta P$ -Regel zu einer Ermittlung der Kausalstärke nach den Vorhersagen der Power PC Theorie (s.a. Abb. 2). Die Tatsache, dass insgesamt ein knappes Drittel der Probanden ihre Vorgehensweise bei der Beurteilung der Stärke von Kausalzusammenhängen geändert haben, kann auch als ein Hinweis darauf verstanden werden, dass diese Probanden sogar zwischen einzelnen Beurteilungsprozessen ihre Strategie adaptiv im Sinne der Kosten-Nutzen-Strategie von Kleinmuntz & Schkade (1993) anpassen oder optimieren, wobei es in einzelnen Fällen vorgekommen ist, dass die Probanden von komplexeren zu einfacheren Strategien gewechselt haben, doch meist wurde von einfachen oder unklaren Strategien zu einer komplexen regel-basierten Strategie gewechselt.

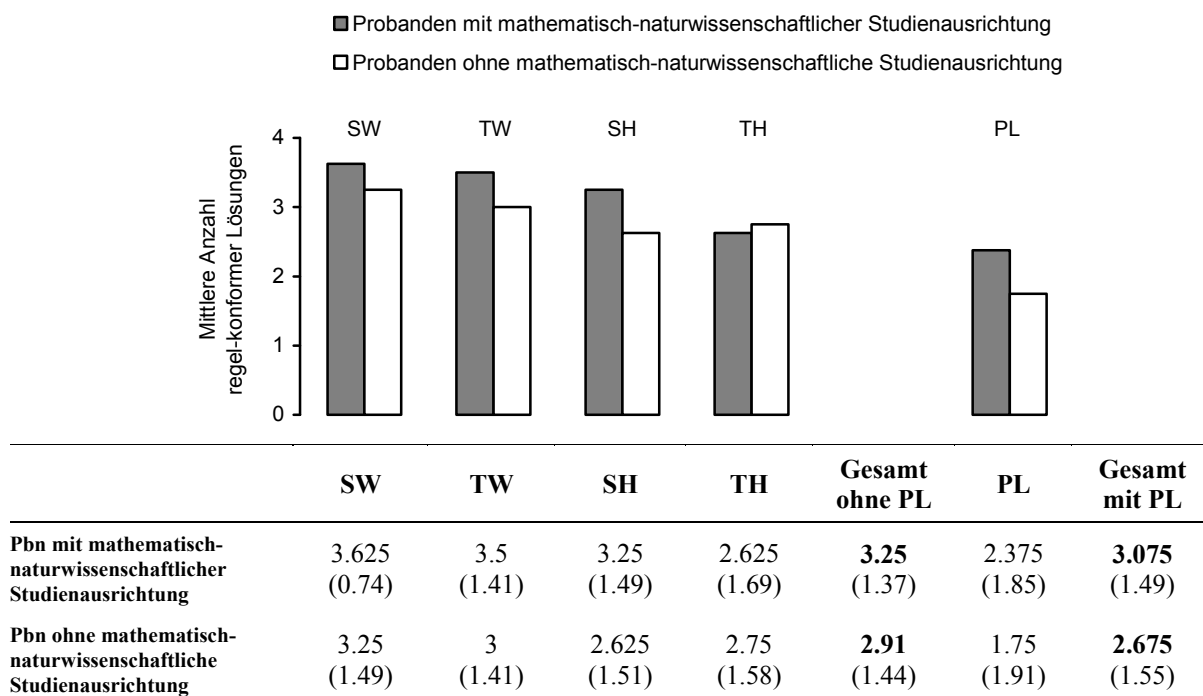
## 8.7 Einfluss der Studieneausrichtung und Analysen zu Teilstichprobeneffekten

### 8.7.1 Einfluss der Studieneausrichtung

#### 8.7.1.1 Ergebnisse

##### *Anzahl regel-konformer Lösungen*

Alle Probanden aus Experiment 1 wurden vorab befragt, welches Fach sie studieren, studiert haben oder studieren wollen. Daraufhin wurden die Probanden in eine Gruppe von Probanden mit mathematisch-naturwissenschaftlicher Studieneausrichtung und in eine Gruppe ohne mathematisch-naturwissenschaftliche Studieneausrichtung unterteilt. Beide Gruppen wurden gleichmäßig und zufällig auf die vier Experimentalbedingungen und die Kontrollbedingung verteilt. Abbildung 17 zeigt, dass die Gruppe mit mathematisch-naturwissenschaftlichen Studieninhalten mit Ausnahme von Bedingung *TH* in allen Experimentalbedingungen und der Kontrollbedingung *PL* leicht höhere Mittelwerte im Hinblick auf die *Anzahl regel-konformer Lösungen* pro Versuchsperson aufweist als die Gruppe ohne mathematisch-naturwissenschaftliche Studieninhalte.



**Abbildung 17:** Experiment 1: Mittlere Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) in Abhängigkeit von *Studieneausrichtung* und *Versuchsbedingung*.

Eine zweifaktorielle Varianzanalyse mit den Faktoren *Studienausrichtung* und *Versuchsbedingung* führte weder zu signifikanten Haupteffekten noch wurde die Interaktion *Studienausrichtung x Versuchsbedingung* signifikant;  $F_{\text{Studienausrichtung}}(1,70) = 1.353, p = .249$ ;  $F_{\text{Versuchsbedingung}}(4,70) = 1.956, p = .111$ ;  $F_{\text{Studienausrichtung} \times \text{Versuchsbedingung}}(4,70) = .164, p = .956$ ;  $MSE = 2.364$ . Zu den gleichen Ergebnissen führte eine dreifaktorielle Varianzanalyse ohne die Kontrollbedingung *PL* und mit den Faktoren *Studienausrichtung*, *Präsentationsform* und *Inhaltsart*;  $F_{\text{Studienausrichtung}}(1,56) = .912, p = .344$ ;  $F_{\text{Präsentationsform}}(1,56) = .369, p = .546$ ,  $F_{\text{Informationsart}}(1,56) = 2.178, p = .146$ ;  $F_{\text{Studienausrichtung} \times \text{Präsentationsform}}(1,56) = .188, p = .666$ ;  $F_{\text{Studienausrichtung} \times \text{Informationsart}}(1,56) = .068, p = .796$ ;  $F_{\text{Präsentationsform} \times \text{Informationsart}}(1,56) = .008, p = .931$ ;  $F_{\text{Studienausrichtung} \times \text{Präsentationsform} \times \text{Informationsart}}(1,56) = .369, p = .546$ ;  $MSE = 2.074$ .

### *Bearbeitungszeit und Verarbeitungseffizienz*

Entsprechende Varianzanalysen wurden auch mit den abhängigen Variablen *Bearbeitungszeit* und *Verarbeitungseffizienz* durchgeführt. Auch wenn die Ergebnisse auf eine leichte Überlegenheit der Gruppe mit mathematisch-naturwissenschaftlichen Studieninhalten hindeuten, so wurden die beobachteten Unterschiede ebenso wie die Interaktionen mit *Studienausrichtung* nicht signifikant. Auf eine ausführliche Darstellung dieser Ergebnisse wird aus Platzgründen verzichtet, die Mittelwerte für die *Bearbeitungszeit* und die *Verarbeitungseffizienz* in Abhängigkeit von *Studienausrichtung* und *Versuchsbedingung* finden sich in Tabelle 3.

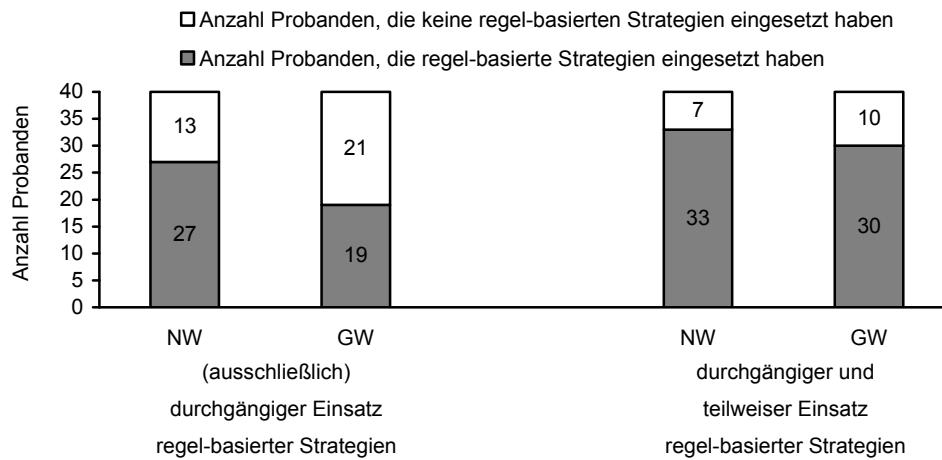
**Tabelle 3:** Experiment 1: Mittlere Bearbeitungszeiten (in Minuten : Sekunden) und mittlere Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von *Versuchsbedingung* und *Studienausrichtung* ("NW" = mathematisch-naturwissenschaftlich ausgerichtete Studienfächer; "GW" = geisteswissenschaftlich ausgerichtete Studienfächer).

	SW		TW		SH		TH		PL		Gesamt	
	NW	GW	NW	GW	NW	GW	NW	GW	NW	GW	NW	GW
<b>Bearbeitungszeit</b>	13 : 30 (5 : 02)	15 : 53 (4 : 25)	14 : 53 (8 : 22)	17 : 53 (6 : 10)	14 : 15 (7 : 33)	12 : 53 (11 : 34)	16 : 00 (6 : 19)	18 : 23 (8 : 37)	27 : 53 (9 : 17)	27 : 53 (11 : 41)	17 : 18 (8 : 54)	18 : 35 (9 : 55)
<b>Effizienz</b>	.30 (.12)	.23 (.14)	.34 (.22)	.19 (.11)	.29 (.14)	.29 (.26)	.20 (.16)	.18 (.12)	.08 (.07)	.07 (.08)	.24 (.17)	.19 (.16)

### *Strategienwahl*

Hinsichtlich der Deckeneffekte in Verbindung mit dem Einfluss des Faktors *Studienausrichtung* wurden  $\chi^2$ -Tests durchgeführt, um herauszufinden, ob die 40 Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung häufiger durchgängig oder zumindest teilweise regel-basierte Strategien eingesetzt haben als die 40 Probanden mit geisteswissenschaftlicher Studienausrichtung oder nicht. Abbildung 18 veranschaulicht, wie

viele Probanden in Abhängigkeit vom Faktor *Studienausrichtung* jeweils durchgängig oder teilweise regel-basierte Strategien eingesetzt haben oder nicht.



**Abbildung 18:** Experiment 1: Häufigkeit des durchgängigen oder zumindest teilweisen Einsatzes regel-basierter Strategien in Abhängigkeit von der *Studienausrichtung* ("NW" = mathematisch-naturwissenschaftlich; "GW" = geisteswissenschaftlich bzw. nicht mathematisch-naturwissenschaftlich).

Hinsichtlich eines durchgängigen Strategieeinsatzes (linkes Säulenpaar) verfehlte  $\chi^2 = 3,274$  knapp das Signifikanzniveau von 5 %, es ergab sich aber ein Trend ( $p = .070$ ), dass Probanden, die sich in ihrem Studium mit mathematisch-naturwissenschaftlichen Inhalten auseinandersetzen haben, etwas häufiger durchgängig von regel-basierten Strategien Gebrauch machen. Dieser Unterschiedstrend zwischen den beiden Probandengruppen verschwindet aber, wenn betrachtet wird, wie häufig regel-basierte Strategien sowohl teilweise als auch durchgängig eingesetzt wurden (rechtes Säulenpaar,  $\chi^2 = .672$ ;  $p = .412$ ). Inhaltlich bedeutet dies, dass tendenziell mehr Probanden, die mathematisch-naturwissenschaftlich ausgerichtete Fächer studierten, auf Anhieb (durchgängig) regel-basierte Strategien einsetzten, während Probanden ohne einen mathematisch-naturwissenschaftlichen Studienhintergrund dies entweder verzögert taten oder von einer regel-basierten zu einer nicht regel-basierten Strategie wechselten.

#### 8.7.1.2 Diskussion

Da bei allen abhängigen Variablen der Faktor *Studienausrichtung* und die damit verbundenen Interaktionen das Signifikanzniveau von  $\alpha = 5\%$  verfehlten, wurden die Hypothesen H4a, b, c und d verworfen. Dass der Faktor *Studienausrichtung* wider Erwarten nicht signifikant wurde, kann mit kleinen Zellgrößen ( $n = 8$ ) zusammenhängen, für wahrscheinlicher halten wir andere Gründe: Zum einen können die benannten Deckeneffekte dazu geführt haben, dass zwischen den beiden Probandengruppen (mit vs. ohne mathematische Studieninhalte) nicht ausreichend



differenziert werden konnte, da es in beiden Probandengruppen genügend Teilnehmer gab, die durchgehend oder zumindest teilweise regel-basierte Lösungen produzierten. Zum anderen könnte die Aufteilung der Probanden in solche mit mathematischen Studieninhalten und solche ohne mathematische Studieninhalte zu grob gewesen sein, da die beiden Probandengruppen auch in sich recht heterogen waren und durch die Unterteilung noch nichts darüber ausgesagt wird, in wie weit sich die beiden Probandengruppen tatsächlich hinsichtlich ihrer mathematischen Fertigkeiten oder ihrer rechnerischen Verarbeitungskapazität unterscheiden haben.

## 8.7.2 Teilstichprobeneffekte

### 8.7.2.1 Ergebnisse

Bei den Probanden aus Experiment 1 handelte es sich zum einen um 37 Leibniz-Kollegiaten und zum anderen um 43 Studenten der Universität Tübingen. Leibniz-Kollegiaten haben in der Regel gerade ihr Abitur bestanden und sich dafür entschieden, nach der Schule erst einmal ein Studium Generale zu absolvieren, bevor sie sich im Studium auf einzelne Fächer spezialisieren. Sie werden aus einer Vielzahl von Bewerbern ausgewählt, bei der Auswahl werden neben spezifischen Begabungen vor allem die Motivation und Aufgeschlossenheit gegenüber verschiedensten Fachrichtungen, sowie die Heterogenität der Gesamtgruppe berücksichtigt (letzteres war ein Motiv, Leibniz-Kollegiaten gezielt für diese Studie anzuwerben). Retrospektiv haben wir uns gefragt, ob sich die potentiell hoch motivierten Studium Generale Absolventen signifikant von den restlichen studentischen Probanden unterscheiden. Die Mittelwerte der verschiedenen Probandenstichproben (Leibniz-Kollegiaten vs. immatrikulierte Studenten) für die verschiedenen abhängigen Variablen in Abhängigkeit von den vier Experimentalbedingungen und der Kontrollgruppe finden sich in Tabelle 4.

**Tabelle 4:** Experiment 1: Mittelwerte (in Klammern Standardabweichungen) für die abhängigen Variablen *Anzahl regel-konformer Lösungen*, *Bearbeitungszeit* und *Verarbeitungseffizienz* in Abhängigkeit von der jeweiligen Teilstichprobe (Leibniz-Kollegiaten vs. immatrikulierte Studenten) und der Versuchsbedingung.

	SW	SH	TW	TH	PL	Gesamt
<b>Anzahl regel-konformer Lösungen</b>						
Kollegiaten	3.83 (0.41)	2.50 (1.51)	3.88 (0.35)	3.38 (1.06)	2.71 (1.70)	3.26 (1.01)
Studenten	3.20 (1.40)	3.38 (1.41)	2.63 (1.77)	2.00 (1.77)	1.56 (1.88)	2.56 (1.72)
<b>Bearbeitungszeit</b>						
Kollegiaten	16:00 (5:13)	8:23 (2:20)	14:23 (4:57)	17:15 (9:44)	24:52 (4:11)	16:10 (5:17)
Studenten	13:54 (4:32)	18:45 (11:11)	18:23 (8:55)	17:08 (4:46)	30:13 (12:57)	19:39 (10:23)
<b>Effizienz</b>						

Kollegiaten	.27 (.11)	.32 (.25)	.31 (.15)	.24 (.12)	.11 (.07)	.25 (.14)
Studenten	.26 (.15)	.25 (.14)	.22 (.22)	.14 (.14)	.05 (.07)	.18 (.16)
<b>Zellgrößen</b>						
Kollegiaten	6	8	8	8	7	37
Studenten	10	8	8	8	9	43

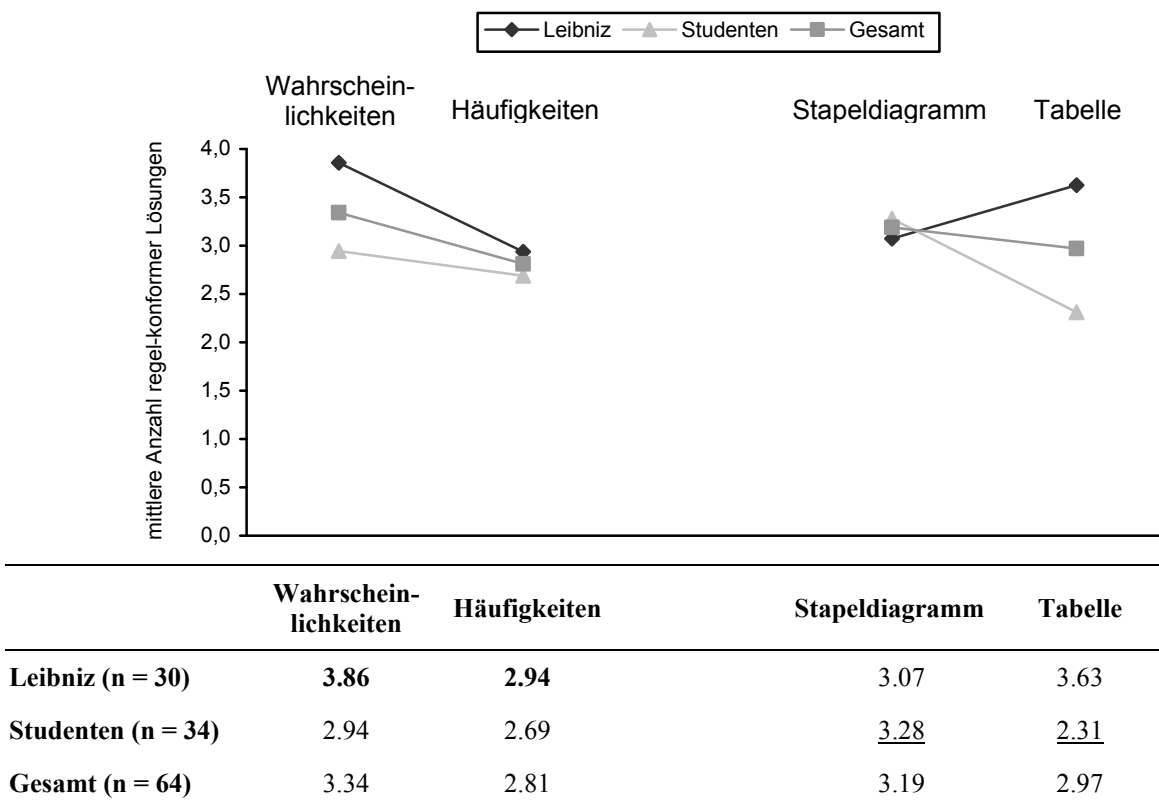
Tabelle 4 veranschaulicht, dass die Leibniz-Kollegiaten häufiger als die Studenten regelkonforme Lösungen produzierten (Ausnahme: Bedingung *SH*). Zudem arbeiteten die Leibniz-Kollegiaten insgesamt schneller (Ausnahme: Bedingungen *SW* und *TH*) und effizienter als die Studenten. Durch die ungleichen Zellgrößen waren die über die verschiedenen Versuchsbedingungen zusammengefassten Mittelwerte allerdings nur eingeschränkt durch t-tests vergleichbar. Die Mittelwerte legen aber nahe, dass die Deckeneffekte bei der abhängigen Variable *Anzahl regel-konformer Lösungen* zum Teil auf die Ergebnisse der Leibniz-Kollegiaten zurückzuführen sind. Besonders ausgeprägt waren die Deckeneffekte bei den Leibniz-Kollegiaten in den Bedingungen *SW* und *TW* (Mittelwerte bei der Anzahl regelkonformer Lösungen > 3,8 bei einem Maximalwert von 4), auch unter der Bedingung *TH* erzielten die Leibniz-Kollegiaten im Vergleich zu den Studenten höhere Werte. Bemerkenswerterweise blieben die Deckeneffekte unter der Bedingung *SH* aus, hier erzielten die Leibniz-Kollegiaten sehr geringe Werte, die Studenten die im Vergleich zu allen anderen Bedingungen höchsten Werte.

Die unterschiedlichen Mittelwerte unter den verschiedenen Versuchsbedingungen legen nahe, dass bei der abhängigen Variable *Anzahl regel-konformer Lösungen* die Faktoren *Informationsart* (Wahrscheinlichkeiten vs. Häufigkeiten) für die Leibniz-Kollegiaten und *Präsentationsform* (Stapeldiagramm vs. Tabelle) für die studentische Probandengruppe signifikant werden könnten. Zur empirischen Überprüfung dieser Vermutung haben wir getrennt für jede Probandengruppe (Leibniz-Kollegiaten und Studenten) und unter Ausschluss der Kontrollbedingung *PL* den Einfluss der beiden Faktoren *Präsentationsform* und *Informationsart* auf die *Anzahl regel-konformer Lösungen* untersucht. Die Ergebnisse aus den beiden 2 x 2 Varianzanalysen sind in Tabelle 5 abgebildet.

**Tabelle 5:** Experiment 1: Ergebnisse aus den 2 (*Präsentationsform*) x 2 (*Informationsart*) Varianzanalysen für die beiden Probandengruppen (Leibniz-Kollegiaten und Studenten) mit der abhängigen Variable *Anzahl regel-konformer Lösungen*. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen.

	<b>Faktor</b>	<b>F-Wert</b>	<b>MSE</b>	<b>P-Wert</b>
<b>Kollegiaten</b>	Präsentationsform (1, 26)	1.577	.984	.220
	Informationsart (1, 26)	6.306		<b>.019</b>
	Präsentationsform x Informationsart (1, 26)	1.303		.264
<b>Studenten</b>	Präsentationsform (1, 30)	3.187	2.512	<u>.084</u>
	Informationsart (1, 30)	.170		.683
	Präsentationsform x Informationsart (1, 30)	.536		.470

Tabelle 5 zeigt, dass je nach Probandengruppe und trotz kleinerer Stichproben unterschiedliche Haupteffekte zustande kamen. In der studentischen Teilstichprobe kam für den Einfluss der unabhängigen Variable *Präsentationsform* auf die *Anzahl regel-konformer Lösungen* ein Trend zustande. In der Teilstichprobe mit den Leibniz-Kollegiaten wurde hingegen der Faktor *Informationsart* signifikant. In der Gesamtstichprobe wurde keiner der Effekte signifikant. In Abbildung 19 werden die unterschiedlichen Haupteffekte für die beiden Teilstichproben und die Gesamtstichprobe miteinander verglichen.



**Abbildung 19:** Experiment 1: Grafische Veranschaulichung der unterschiedlichen Haupteffekte auf die Anzahl regel-konformer Lösungen für die Gesamtstichprobe und die beiden Teilstichproben (Leibniz-Kollegiaten und Studenten). In der angehängten Tabelle sind signifikante Unterschiede **fett** gedruckt, Trends wurden unterstrichen.

### 8.7.2.2 Diskussion

Die unterschiedlichen Haupteffekte für die beiden Teilstichproben werfen natürlich die Frage auf, wodurch diese Unterschiede zustande kamen. Die insgesamt höheren Werte der Leibniz-Kollegiaten könnten in einer insgesamt höheren Motivation der jüngeren Leibniz-Kollegiaten oder in einer durchschnittlich höheren numerischen Verarbeitungskapazität bei den Leibniz-Kollegiaten begründet sein, für letzteres spricht auch die Tatsache, dass das Verhältnis zwischen eher mathematisch-naturwissenschaftlich (NW) und eher geisteswissenschaftlich (GW) ausgerichteten Probanden in beiden Teilstichproben unterschiedlich war (Leibniz: 20 NW, 14 GW; Studenten: 12 NW, 18 GW). In der Untersuchung von Maichle (1992) kam heraus, dass Abiturienten mit einem stärkeren Interesse für Mathematik und Naturwissenschaften Tabellen und Diagramme besser interpretieren konnten als Abiturienten, die sich mehr für andere Fächer interessierten. In der Nachbefragung (siehe Kapitel 8.8), wo im Wesentlichen Vorerfahrungen mit den eingesetzten Medien (Ereignislisten, Tabellen, Diagramme) und der Umgang mit Grundrechenarten, Prozentrechnen und einem Tabellenkalkulationsprogramm erfasst wurden, ergaben sich allerdings keine signifikanten Unterschiede zwischen den beiden Teilstichproben.

Interessant ist es auch, ein besonderes Augenmerk auf die Bedingung SH zu richten. Hier erzielten die Studenten untypischer Weise höhere Werte (3.38) als die Leibniz-Kollegiaten (2.50). Anders als in den beiden vollständigen Teilstichproben war hier das Verhältnis zwischen eher mathematisch-naturwissenschaftlich und eher geisteswissenschaftlich ausgerichteten Probanden genau umgekehrt: Unter der Bedingung SH waren mehr mathematisch-naturwissenschaftlich ausgerichtete Probanden Studenten und mehr geisteswissenschaftlich ausgerichtete Probanden Leibniz-Kollegiaten (Leibniz: 3 NW, 5 GW; Studenten: 5 NW, 3 GW). Weitere Analysen der Nachbefragung ergaben, dass die Leibniz-Kollegiaten unter der Bedingung SH weniger häufig Diagramme eigenständig anfertigten als die Studenten unter der Bedingung SH, außerdem nahmen sich insbesondere die nicht mathematisch-naturwissenschaftlich orientierten Leibniz-Kollegiaten ungewöhnlich wenig Zeit (siehe auch Tabelle 4) für die Bearbeitung der Aufgaben. T-Tests, die nur auf die Probanden unter der Bedingung SH beschränkt waren und unter der Annahme ungleicher Varianzen durchgeführt wurden, zeigten, dass die Unterschiede zwischen den Mittelwerten der beiden Teilstichproben sowohl für die Bearbeitungsdauer ( $T_{\text{Bearbeitungsdauer}} = -2.569$ ;  $p = .035$ ) als auch für das Item "Anfertigen von Diagrammen" aus der Nachbefragung ( $T_{\text{Machen Graf}} = -2.435$ ;  $p = .031$ ) das zweiseitige Signifikanzniveau ( $\alpha = 2,5 \%$ ) nur knapp verfehlten. Der ausgesprochen kurze Zeitaufwand einer Reihe von Leibniz-Kollegiaten unter der Bedingung

SH könnte zu Lasten der Sorgfalt bei der Aufgabenbearbeitung gegangen sein, und spricht in diesem Fall gegen die sonst postulierte hohe Motivation der Leibniz-Kollegiaten. Für diese Behauptung sprechen auch die unter der Bedingung SH schwach positiven, wenn auch nicht signifikanten Rangkorrelationen (Spearman-Rho) zwischen den abhängigen Variablen *Anzahl regel-konformer Lösungen* und *Bearbeitungszeit* ( $r_s = .251$ ;  $p = .349$ ). Über alle Bedingungen hinweg waren die Rangkorrelationen zwischen der *Anzahl regel-konformer Lösungen* und der *Bearbeitungszeit* dagegen signifikant negativ ausgefallen ( $r_s = -.244$ ;  $p = .029$ ).

Insgesamt sprechen die Ergebnisse aus den Analysen der Teilstichproben dafür, dass die Wahl der Stichprobe das Auftreten oder Nicht-Auftreten von Haupteffekten der beiden Faktoren *Informationsart* und *Präsentationsform* erheblich modifizieren kann. Die in Experiment 1 vorgenommene, aber letztlich folgenlose Unterscheidung von mathematisch-naturwissenschaftlich vs. nicht mathematisch-naturwissenschaftlich ausgerichteten Probanden widerlegt diese Behauptung nicht, sie zeigt nach unserer Auffassung nur, dass eine Unterteilung nach Studienfächern allein zu grob ist. Letztlich kommt es neben der mentalen Beanspruchung durch aufgaben- und medienbezogene Faktoren auf individuelle Charakteristika an (wie z.B. Motivation, numerische Verarbeitungskapazität oder aufgabenspezifisches Vorwissen der Probanden), ob eine regel-basierte oder hochkomplexe Strategie zur Bestimmung der Wirksamkeit einer monokausalen Ursache gewählt bzw. aktiviert wird oder nicht.

## 8.8 Ergebnisse der Nachbefragung

Um Anhaltspunkte dafür zu bekommen, welche individuellen Voraussetzungen (Kovariaten) neben der Ausrichtung des Studienfachs die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen mit beeinflussen könnten, haben wir mit allen Probanden eine schriftliche Nachbefragung durchgeführt. Mit den Fragen sollten im Wesentlichen die Vorerfahrungen mit allen eingesetzten Medien (Ereignislisten, Tabellen, Diagramme), sowie zum Umgang mit Grundrechenarten, Prozentrechnen und einem Tabellenkalkulationsprogramm erfasst werden. Die Items der Nachbefragung können in Anhang A1.4 eingesehen werden.

In einer ersten Analyse konnte mittels mehrerer einfaktorieller Varianzanalysen ausgeschlossen werden, dass die Ergebnisse der Nachbefragung durch die Zuordnung in eine der Versuchsbedingungen beeinflusst worden sind; der Faktor Versuchsbedingung mit den fünf Stufen SH, SW, TH, TW und PL wurde für keines der elf in der Auswertung berücksichtigten Items aus der Nachbefragung signifikant. Dabei wurden nicht alle Items

separat ausgewertet, sondern es wurden die verschiedenen Werte zur Vertrautheit mit den in Experiment 1 eingesetzten Medien jeweils über die vier verschiedenen Lebensbereiche (Schule, Studium, Beruf, Freizeit) hinweg zusammengefasst und gemittelt.

Anschließend wurden die Ergebnisse aus den Skalen der Nachbefragung mit den Ergebnissen für die abhängigen Variablen korreliert. Da es sich bei allen Daten aus der Nachbefragung um Ordinalskalenwerte handelte, haben wir Rangkorrelationen (Kendall Tau-b) berechnet. Eine Berechnung von Regressionsgleichungen erübrigt sich durch das fehlende Intervallskalenniveau der Fragebogendaten. Tabelle 6 zeigt alle Rangkorrelationen zwischen den Werten für die abhängigen Variablen und den Werten für die Items der Nachbefragung.

**Tabelle 6:** Experiment 1: Rangkorrelationen (Kendall-Tau-b) zwischen den Werten für die abhängigen Variablen und den Werten für die Items aus der Nachbefragung. In Klammern sind die p-Werte angegeben, signifikante Korrelationen mit einem p-Wert < .05 sind **fett** gedruckt; Korrelationen mit p-Werten < .10 (Trend) wurden unterstrichen.

	Mittlere Anzahl regel-konformer Lösungen	Bearbeitungszeit	Verarbeitungseffizienz
Vertrautheit Liste	-.064 (p = .467)	<b>-.181</b> (p = .023)	.033 (p = .676)
Vertrautheit Tabelle	.083 (p = .353)	-.090 (p = .261)	.052 (p = .519)
Vertrautheit Diagramm	.062 (p = .489)	-.029 (p = .720)	-.020 (p = .808)
Häufigkeit Anfertigen Liste	-.046 (p = .626)	<b>-.202</b> (p = .019)	.091 (p = .290)
Häufigkeit Anfertigen Tabelle	.054 (p = .579)	-.096 (p = .271)	.040 (p = .647)
Häufigkeit Anfertigen Diagramm	.083 (p = .391)	-.051 (p = .565)	.054 (p = .540)
Häufigkeit Grundrechenarten	<u>.197</u> (p = .053)	-.080 (p = .387)	.088 (p = .341)
Sicherheit Grundrechenarten	.027 (p = .780)	-.059 (p = .504)	.058 (p = .510)
Häufigkeit Prozentrechnen	.018 (p = .853)	-.086 (p = .320)	-.001 (p = .990)
Sicherheit Prozentrechnen	.070 (p = .462)	-.118 (p = .172)	.095 (p = .269)
Häufigkeit Tabellenkalkulation	.087 (p = .367)	-.021 (p = .810)	.028 (p = .749)

Signifikante Korrelationen ( $p < .05$ ) ergaben sich zwischen der abhängigen Variable *Bearbeitungszeit* und den beiden Items "Vertrautheit mit einer Liste" und "Anfertigen einer Liste", ein Trend zwischen der abhängigen Variable *Anzahl regel-konformer Lösungen* und dem Item "Häufigkeit Grundrechenarten". Insgesamt fielen die Korrelationen gering aus, keine Korrelation erreichte ein Maß von mehr als .25 ( $r^2 < .10$ ). Zusätzliche Berechnungen von Produkt-Moment-Korrelationen, aus denen der Einfluss der vier Experimentalbedingungen und der Kontrollbedingung herauspartialisiert wurde, bestätigen im Wesentlichen die Ergebnisse aus Tabelle 6, lediglich die Korrelationen zwischen der

*Bearbeitungszeit* und den Items "Vertrautheit mit einer Liste" und "Anfertigen einer Liste" wurden signifikant, zudem ergaben sich Trends zwischen dem Effizienzmaß und dem Items "Anfertigen einer Liste" und "Sicherheit Prozentrechnen", sowie zwischen der *Anzahl regelkonformer Lösungen* und dem Item "Häufigkeit Grundrechenarten", wie Tabelle 7 zeigt:

**Tabelle 7:** Experiment 1: Partialkorrelationen zwischen den Werten für die abhängigen Variablen und den Werten für die Items aus der Nachbefragung, bei denen der Einfluss der vier Experimentalbedingungen und der Kontrollbedingung kontrolliert wurde. In Klammern sind die Anzahl der Freiheitsgrade und die p-Werte angegeben. Signifikante Korrelationen mit einem p-Wert < .05 sind **fett** gedruckt; Korrelationen mit p-Werten < .10 (Trend) unterstrichen und Korrelationen mit p-Werten < .15 (Trendnähe) *kursiv* gedruckt.

	Mittlere Anzahl regelkonformer Lösungen	Bearbeitungszeit	Verarbeitungseffizienz
Vertrautheit Liste	-.14 (77; p = .203)	<b>-.25</b> (77; p = .026)	.14 (77; p = .225)
Vertrautheit Tabelle	.03 (77; p = .761)	-.15 (77; p = .198)	.10 (77; p = .385)
Vertrautheit Diagramm	.06 (76; p = .586)	-.17 (76; p = .128)	.08 (76; p = .491)
Häufigkeit Anfertigen Liste	-.07 (76; p = .517)	<b>-.26</b> (76; p = .020)	<u>.21</u> (76; p = .062)
Häufigkeit Anfertigen Tabelle	.04 (75; p = .736)	-.15 (75; p = .184)	.15 (75; p = .200)
Häufigkeit Anfertigen Diagramm	-.02 (75; p = .860)	.02 (75; p = .897)	.06 (75; p = .593)
Häufigkeit Grundrechenarten	<u>.21</u> (77; p = .065)	-.18 (77; p = .118)	.12 (77; p = .285)
Sicherheit Grundrechenarten	.03 (77; p = .819)	-.11 (77; p = .350)	.08 (77; p = .495)
Häufigkeit Prozentrechnen	-.01 (77; p = .941)	-.16 (77; p = .171)	.10 (77; p = .375)
Sicherheit Prozentrechnen	.06 (77; p = .586)	-.13 (77; p = .262)	<u>.20</u> (77; p = .081)
Häufigkeit Tabellenkalkulation	.09 (77; p = .447)	-.02 (77; p = .892)	.10 (77; p = .377)

Die gefundenen Korrelationen haben lediglich hinweisenden Charakter und sind schwer zu interpretieren, einerseits weil die gefundenen Korrelationen zu gering ausfielen, andererseits weil es sich im Zusammenhang mit der Nachbefragung um Selbstauskünfte und Selbsteinschätzungen der Probanden handelte, die mit zahlreichen messtheoretischen Problemen und Urteilsfehlern behaftet sind (s. z.B. Bortz, 1984, für einen Überblick): Wer sich beispielsweise selbst für sehr sicher im Prozentrechnen hält, könnte trotzdem beim tatsächlichen Versuch, Prozentrechenaufgaben zu lösen, im Vergleich zu anderen Probanden unterdurchschnittlich abschneiden. Für die folgenden Untersuchungen haben wir deshalb beschlossen, auf Selbsteinschätzungen zu verzichten und stattdessen die numerische Verarbeitungskapazität der Probanden mittels entsprechender Items aus einem Intelligenztest zu erfassen. Ein in diesem Zusammenhang erwähnenswertes Ergebnis aus der Nachbefragung zu Experiment 1 war, dass sich Frauen im Vergleich zu Männern als signifikant weniger

sicher im Prozentrechnen einschätzten, obwohl sich die Geschlechter in allen abhängigen Variablen nicht einmal annäherungsweise signifikant unterschieden.

Einen bedeutenden Einfluss auf die Selbsteinschätzungen in der Nachbefragung hatte der Faktor *Studienausrichtung*, wie eine Reihe von durchgeführten t-Tests zeigte: So geben Probanden aus mathematisch-naturwissenschaftlich ausgerichteten Studienfächern im Vergleich zu Probanden aus eher geisteswissenschaftlich ausgerichteten Studienfächern an, dass sie häufiger Diagramme anfertigen ( $p = .002$ ), Prozentrechnen ( $p = .021$ ) und Tabellenkalkulationsprogramme wie z.B. Excel ( $p = .024$ ) einsetzen, sowie sicherer im Prozentrechnen ( $p = .007$ ) und in den Grundrechenarten ( $p = .037$ ) sind. Trends oder Trendnähe ergaben sich für den Einfluss von *Studienausrichtung* auf die Items Vertrautheit im Umgang mit Tabellen und Diagrammen, sowie Häufigkeit des Anfertigens von Tabellen. Hinsichtlich der Vertrautheit und der Häufigkeit des Anfertigens von Listen, sowie der Häufigkeit des Einsatzes von Grundrechenarten unterschieden sich die Probanden in Abhängigkeit vom Faktor *Studienausrichtung* nicht.



## 8.9 Abschließende Diskussion von Experiment 1

Die Auswertung der Ergebnisse aus Experiment 1 hat ergeben, dass der Faktor *Informationsart* die Anzahl regel-konformer Lösungen und die Strategiewahl schwach beeinflusst, nicht aber die Bearbeitungszeiten und die Verarbeitungseffizienz. Der schwache Einfluss von *Informationsart* auf die Anzahl regel-konformer Lösungen und die Strategiewahl wurde damit erklärt, dass unter den beiden Wahrscheinlichkeitsbedingungen die Beanspruchung von Informationsverarbeitung und Gedächtnis geringer als unter den beiden Häufigkeitsbedingungen ausfällt. Die reduzierte mentale Beanspruchung führt dazu, dass die Probanden insgesamt über eine höhere kognitive Kapazität verfügen, um komplexe regel-basierte Strategien selbst zu generieren oder um bereits abgespeicherte Strategien aus dem Gedächtnis abzurufen und einzusetzen. Dementsprechend zeigt sich die reduzierte mentale Beanspruchung unter den beiden Wahrscheinlichkeitsbedingungen darin, dass mehr Probanden von komplexen, regel-basierten Strategien Gebrauch gemacht haben, wodurch sich auch die Anzahl regel-konformer Lösungen pro Versuchsperson erhöht hat.

Der Einfluss des Faktors *Präsentationsform* war zwar ebenfalls schwach, aber anders geartet als derjenige des Faktors *Informationsart*. Die Unterscheidung zwischen Stapeldiagramm und tabellarischer Vierfeldertafel war wider Erwarten bedeutungslos im Hinblick auf die Strategiewahl und die Anzahl regel-konformer Lösungen, doch zeigte es sich, dass Kausalinformationen aus Diagrammen tendenziell schneller zu Kausalurteilen verarbeitet werden konnten als Informationen, die tabellarischen Vierfeldertafeln entnommen werden mussten.

Das Ergebnis, dass der Unterschied zwischen Stapeldiagramm und tabellarischer Vierfeldertafel im Hinblick auf die Strategiewahl und die Anzahl regel-konformer Lösungen nicht signifikant wurde, lässt sich zum einen auf das Auftreten von Deckeneffekten zurückführen, zum anderen besteht aber auch die Möglichkeit, dass sich die beiden Präsentationsformen tatsächlich nur unwesentlich hinsichtlich ihres Potenzials unterscheiden, den Einsatz regel-basierter Strategien nahe zu legen und die Probanden dazu zu bringen, regel-konforme Lösungen zu produzieren oder aus dem Gedächtnis abzurufen. Sowohl in Stapeldiagrammen als auch in tabellarischen Vierfeldertafeln könnten die Informationen bereits so übersichtlich und strukturiert aufbereitet gewesen sein, dass die Unterschiede zwischen beiden Präsentationsformen nicht mehr sonderlich ins Gewicht gefallen sind und es den Probanden hinreichend einfach gemacht wurde, regel-basierte Strategien einzusetzen und

regel-konforme Lösungen zu bilden. In Experiment 2 soll deshalb eine sehr übersichtlich gestaltete und visuell hoch strukturierte Präsentationsform mit einer Präsentationsform verglichen werden, in der das Potenzial oder die Strukturierungsangebote zur Induzierung des Gebrauchs regel-basierter Strategien minimiert wurde.

Im Zusammenhang mit der *Bearbeitungszeit* haben wir die etwas schnellere Informationsverarbeitung unter den beiden Diagrammbedingungen darauf zurückgeführt, dass komplexe Informationsverarbeitungsprozesse auf die Wahrnehmungsebene verlagert (Fiedler, 2000; Wickens & Carswell, 1995) und perzeptuelle Inferenzen möglich gemacht (Larkin & Simon, 1987) werden. Dies führt dazu, dass visuelle Vergleiche im Stapeldiagramm schneller und unmittelbarer nahe gelegt werden als in einer tabellarischen Vierfeldertafel.

Durch den Vergleich zwischen den vier Experimentalbedingungen und der Kontrollbedingung Protokollliste konnte gezeigt werden, dass sowohl tabellarische Vierfeldertafeln als auch Stapeldiagramme signifikant schneller und effizienter als die Protokollliste verarbeitet wurden. Ebenso waren die Probanden in den vier Experimentalbedingungen dazu in der Lage, mehr regel-basierte Strategien einzusetzen und eine höhere Anzahl regel-konformer Lösungen zu entwickeln als in der Listenbedingung, allerdings wurden diese Unterschiede nur für die Einzelkontraste zwischen den beiden Wahrscheinlichkeitsbedingungen und der Listenbedingung signifikant. Für den Unterschied zwischen den Bedingungen *SH* und *PL* wurde hinsichtlich der Anzahl regel-konformer Lösungen ein Trend erzielt. Die gefundenen Unterschiede waren von uns vorhergesagt worden. Sie sprechen in unseren Augen dafür, dass die mentale Beanspruchung unter den beiden Wahrscheinlichkeitsbedingungen und tendenziell auch unter der Häufigkeitsbedingung *SH* geringer ist als unter der Listenbedingung. Die Überlegenheit der beiden Wahrscheinlichkeitsbedingungen und der Häufigkeitsbedingung *SH* gegenüber der Listenbedingung ergibt sich auch dadurch, dass den Probanden durch die im Vergleich zur Listenbedingung übersichtlichere Informationsanordnung in den drei genannten Experimentalbedingungen mehr Strukturierungshilfen (Larkin & Simon, 1987; Stenning & Oberlander, 1995; Zhang, 2000, 1997; Zhang & Norman, 1994) angeboten werden. Zudem werden ihnen insbesondere in den beiden Diagrammbedingungen durch die Verlagerung von komplexen mentalen Operationen auf die Wahrnehmungsebene (Wickens & Carswell, 1995) perzeptuelle Inferenzen ermöglicht (Larkin & Simon, 1987).

Auch wenn die im Zusammenhang mit der *Anzahl regel-konformer Lösungen* beobachteten Unterschiede zwischen den einzelnen Versuchsbedingungen in der von uns vorhergesagten Richtung lagen, so sind sie doch schwächer ausgeprägt als erwartet. Lediglich die

Unterschiede zwischen den beiden Wahrscheinlichkeitsbedingungen einerseits und der Kontrollbedingung Protokollliste andererseits wurden signifikant. Innerhalb der vier Experimentalbedingungen waren die Einflüsse der beiden Faktoren *Präsentationsform* und *Informationsart* zu schwach, um das Signifikanzniveau von  $\alpha = 5\%$  zu überschreiten.

Die Frage, warum die Unterschiede im Zusammenhang mit der Anzahl regel-konformer Lösungen nicht höher ausgefallen sind, haben wir mit dem Auftreten von Deckeneffekten begründet. Diese Deckeneffekte sind vermutlich die Folge von Stichprobenselektion. Zu berücksichtigen ist an dieser Stelle aber auch, dass allen Probanden die Informationen über Ursache-Wirkungs-Zusammenhänge durchgängig in summarischer Form dargeboten wurden. Eine summarische Darbietung von Informationen über einfache Ursache-Wirkungs-Zusammenhänge trägt nach unserer Einschätzung ebenfalls zu den beobachteten Deckeneffekten bei, weil sie die Komplexität der zu bewältigenden Aufgabenschritte bei der Beurteilung der Wirksamkeit einer Ursache in Grenzen hält und damit mentale Beanspruchung der Probanden reduziert. Eine Reihe von Studien (Arkes & Harkness, 1983; Buehner et al., 2003; Kao & Wasserman, 1993; Lober & Shanks, 2000; Ward & Jenkins, 1965) legt nahe, dass Probanden bei einfachen und summarisch dargebotenen Ursache-Wirkungs-Zusammenhängen eher regel-konforme Lösungen produzieren als bei einer einzelfallweisen Darbietung. In Experiment 2 soll deshalb auch der Frage nachgegangen werden, ob bei einer einzelfallweisen Darbietung kausaler Informationen weniger regel-konforme Lösungen hervorgebracht werden als bei einer summarischen Informationsdarbietung.

Als Ergänzung zum Auftreten von Deckeneffekten, kann das Ausbleiben signifikanter Unterschiede im Zusammenhang mit der Anzahl regel-konformer Lösungen wie bereits erwähnt aber auch darin begründet sein, dass die Unterschiede zwischen den in den vier Experimentalbedingungen präsentierten Medien einfach nicht groß genug waren. Ein solcher Befund ist in der traditionellen Medienforschung ganz und gar nicht ungewöhnlich. Mayer (2001) referiert eine Überblicksarbeit von Clark (1994), in der argumentiert wird, dass Medien an sich keinen oder nur einen schwachen Einfluss auf das Lernen haben und es zudem nicht möglich ist, die Auswirkungen eines Mediums von den Auswirkungen einer Instruktionmethode oder einer gestellten Aufgabe zu trennen.

Die Tatsache, dass die traditionelle Medienforschung, in der jeweils ein Medium mit einem anderen verglichen wurde, enttäuschende und uneinheitliche Ergebnisse gebracht hat (Clark & Salomon, 1986; Mayer, 1997; 2001) hat dazu geführt, dass Forschung, in denen der Einfluss unterschiedlicher Medien miteinander verglichen wurde, als unproduktiv angesehen

wurde. Neben einer uneinheitlichen und wenig überzeugenden empirischen Befundlage wurden von Mayer (2001) methodologische, konzeptuelle und theoretische Gründe für die Unsinnigkeit einer traditionellen Medienvergleichsforschung vorgetragen.

In Experiment 1 wurde dennoch ein Vergleich zwischen unterschiedlichen Präsentationsformen vorgenommen, weil es dabei nicht um die Frage ging, welche der eingesetzten Präsentationsformen besser im Hinblick auf den Lernerfolg abschneidet, sondern darum, die Auswirkungen unterschiedlicher Präsentationsformen oder Informationsarten auf die bei der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen auftretenden kognitiven Verarbeitungsprozesse zu untersuchen. Dabei wurde von uns versucht, den Medienbegriff nicht als bloßes Mittel zur Informationsübermittlung zu verstehen, sondern mit zu berücksichtigen, dass der Einfluss medienbezogener Faktoren auf die menschliche Informationsverarbeitung untrennbar mit der jeweiligen Aufgabe und personenbezogenen Faktoren verknüpft ist (s.a. Abbildung 5).

Konkret kann festgehalten werden, dass die *Präsentationsform* einen geringen Einfluss auf die Wahrnehmungs- und Informationsverarbeitungsgeschwindigkeit zu haben scheint, während sich der Faktor *Informationsart* auf der Ebene von Informationsintegration und Urteilsbildung schwach auf die Strategiewahl und die Anzahl regel-konformer Lösungen auszuwirken schien. Die unterschiedlichen Auswirkungen von *Präsentationsform* und *Informationsart* bestätigen auch die bei der Vorstellung der abhängigen Variablen aufgestellte Vermutung, dass die Variablen *Bearbeitungszeit* (Wahrnehmungs- und Informationsverarbeitungsgeschwindigkeit) einerseits, sowie *Strategiewahl* und *Anzahl regel-konformer Lösungen* andererseits jeweils auf unterschiedliche Aspekte von Gedächtnisbeanspruchung Bezug nehmen.

Auch wenn der Faktor *Studienausrichtung* keinen signifikanten Einfluss hatte, so kommen wir dennoch zu dem Ergebnis, dass individuelle Unterschiede bei der Beurteilung der Stärke von Kausalzusammenhängen hoch bedeutsam sein können. Dies zeigt sich zum einen darin, dass sich bei einer getrennten Auswertung von Leibniz-Kollegiaten und Studierenden die Auswirkungen der beiden unabhängigen Variablen klar unterscheiden, zum anderen konnte bei der Analyse der Strategiewahl festgestellt werden, dass die eingesetzten Strategien auch innerhalb von ein- und derselben Versuchsbedingung uneinheitlich waren. Das relativ häufig aufgetretene Phänomen, dass die Probanden ihre Strategien wechseln, kann im Sinne der Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) als Hinweis darauf verstanden werden, dass die Probanden ihre Strategien an die subjektiv wahrgenommene Aufgabenschwierigkeit anpassen und entweder eine einfachere Strategie wählen oder auf

komplexe regel-basierte Strategien zurückgreifen. Bei sechs Probanden konnte sogar beobachtet werden, dass zwischen verschiedenen regel-basierten Strategien gewechselt wurde. Da die Selbsteinschätzungen der Probanden aus der Nachbefragung letztlich keine Antwort auf die Frage brachten, worin sich Probanden, die regel-basiert vorgehen, von solchen unterscheiden, die nicht regel-basiert vorgehen, soll in Experiment 2 der Einfluss des numerischen Denkens, insbesondere der numerischen Verarbeitungskapazität, auf die abhängigen Variablen untersucht werden.

In Verbindung mit der Analyse der Strategiewahl konnte ein Trend ausgemacht werden, dass der Faktor *Informationsart* in einem gewissen Maß, wenn auch keinesfalls durchgängig und bei allen Probanden Einfluss darauf hat, welche regel-basierte Strategie von den Probanden generiert oder abgerufen wird. Auch wenn der Einsatz der  $\Delta P$ -Regel in allen Bedingungen häufiger zu beobachten war als eine Bestimmung der kausalen Stärke  $p$  nach der PPC-Theorie, so war doch die relative Anzahl von Probanden, welche die kausale Stärke entsprechend den Vorhersagen der PPC-Theorie ermittelten, unter Häufigkeitsbedingungen (vor allem unter der Bedingung *SH*) größer als unter Wahrscheinlichkeitsbedingungen. Eine mögliche Erklärung dafür ist, dass sich - unter der Voraussetzung, dass die Probanden von der einfacheren Gleichung (6) und nicht von Gleichung (4) Gebrauch machen - die kausale Stärke  $p$  unter Häufigkeitsbedingungen mindestens genauso einfach berechnen lässt wie unter Wahrscheinlichkeitsbedingungen (vorausgesetzt die Anzahl der Fälle bei gegebener und nicht gegebener Ursache ist wie in Experiment 1 identisch). Umgekehrt lässt sich das Kontingenzmaß  $\Delta P$  unter Wahrscheinlichkeitsbedingungen sehr viel leichter berechnen als unter Häufigkeitsbedingungen, weil den Probanden unter Wahrscheinlichkeitsbedingungen bereits Werte für Erfolgs- und Basisrate vorliegen, während unter den beiden Häufigkeitsbedingungen Erfolgs- und Basisrate erst noch aus den in Prozentwerte umzuwandelnden Zellhäufigkeiten errechnet werden müssen. Wird davon ausgegangen, dass die Probanden den Aufwand für den Einsatz einer Strategie im Sinne der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) antizipieren können und entsprechend auf rechnerisch einfachere Strategien bei der Bestimmung der Wirksamkeitsbeurteilung zurückgreifen, so kann vermutet werden, dass unter Häufigkeitsbedingungen weniger Probanden von der  $\Delta P$ -Regel Gebrauch machten als unter Wahrscheinlichkeitsbedingungen und sich auch einige der Probanden dafür entschieden haben, die kausale Stärke entsprechend der verhältnismäßig einfachen Gleichung (6) zu berechnen.



## 9 Experiment 2: Einfluss von Darbietungsart und visueller Strukturierung des darstellenden Mediums

Experiment 1 zeigte, dass sich tabellarische Vierfeldertafeln und Stapeldiagramme nicht signifikant hinsichtlich ihres Einflusses auf die abhängigen Variablen *Anzahl regel-konformer Lösungen*, *Bearbeitungszeit* und *Verarbeitungseffizienz* unterscheiden. Beide Medien haben gemeinsam, dass die zu verarbeitenden, summarisch dargebotenen Informationen durch ein hohes Maß an visueller Strukturierung übersichtlich angeordnet sind. Dadurch wurden im Vergleich zur Protokollliste Suchprozesse, die Gruppierung zusammen gehörender Informationen oder Schlussfolgerungsprozesse wie zum Beispiel das Anstellen von Vergleichen vereinfacht. Als Folge dieser Vereinfachung benötigten die Probanden in einer statisch-deskriptiven Experimentalumgebung mit summarischer Ergebniszusammenfassung weniger Zeit und waren eher in der Lage, komplexe kognitive Schemata zur Berechnung von Wirksamkeiten abzurufen oder zu generieren.

Was geschieht nun aber, wenn die Probanden vor die vermeintlich schwierigere Aufgabe gestellt werden, die Wirksamkeit einer Ursache an Hand von einzelfallweise dargebotenen Kausalinformationen zu bestimmen? In Experiment 1 fiel es den Probanden bei summarischer Informationsdarbietung relativ leicht, von regel-basierten Strategien Gebrauch zu machen (etwa 60 % setzten durchgängig, über 75 % aller Probanden zumindest teilweise regel-basierte Strategien ein). In Experiment 2 sollte die Fragestellung untersucht werden, welchen Einfluss eine Variation der *Darbietungsart* (summarisch versus einzelfallweise) auf die Beurteilung kausaler Zusammenhänge und die Einschätzung der Wirksamkeit einer Ursache haben. Wir gehen davon aus, dass sich bei einzelfallweiser Darbietung die Aufgabenkomplexität und damit auch die mentale Beanspruchung für die Probanden erhöhen. Das hat wiederum Auswirkungen auf die *Anzahl regel-konformer Lösungen* und auf die *Strategienwahl*. Wir nehmen an, dass die Probanden bei einzelfallweiser Darbietung weniger regel-konforme Lösungen produzieren und weniger regel-basierte Strategien einsetzen als bei einer summarischen Darbietung.

Bisherige Direktvergleiche von summarischer und einzelfallweiser Darbietung blieben des Öfteren auf einen Vergleich zwischen Protokolllisten und einzelfallweiser dargebotener Information beschränkt (Hagmayer, 2001, Ex. 3; Van Hamme & Wassermann, 1993), selten wurden anstelle von Protokolllisten Darbietungsarten eingesetzt, die sich durch eine im Vergleich zu einfachen Protokolllisten höhere visuelle Strukturierung der summarisch

dargebotenen Information auszeichnen (Ward & Jenkins, 1965; Kao & Wasserman, 1993). Neben dem Einfluss der *Darbietungsart* (einzelfallweise versus summarische Darbietung) sollte deshalb der Einfluss des *Ausmaßes an visueller Strukturierung* (hoch vs. gering) als zweiter Faktor untersucht werden. Produzieren Probanden häufiger regel-konforme Lösungen in weniger Zeit, wenn ihnen die Informationen über Ursache-Wirkungs-Zusammenhänge in einer visuell hoch strukturierten Form präsentiert werden? In Experiment 1 war die visuelle Strukturierung der eingesetzten Medien in den vier Experimentalbedingungen durchgängig hoch.

Experiment 1 hat auch gezeigt, dass trotz der vermuteten Deckeneffekte individuelle Unterschiede bei der Beurteilung von Ursache-Wirkungszusammenhängen bedeutsam sind. Auch wenn sich die Aufgabenstellung in Experiment 1 insgesamt als relativ einfach herausstellte, so traf dies doch nicht auf alle Probanden in gleicher Weise zu. Dies wurde nicht nur an der heterogenen Strategiewahl deutlich, sondern auch daran, dass die Zusammensetzung der Stichprobe den Einfluss unterschiedlicher Präsentationsformen und Inhaltsarten moderieren kann. Neben Unterschieden hinsichtlich der Motivation wurde die numerische Verarbeitungskapazität der Probanden als ein möglicher moderierender Faktor erwähnt. In Experiment 2 sollte deshalb der Frage nachgegangen werden, ob das numerische Denken einschließlich der numerischen Verarbeitungskapazität der Probanden als Kovariate neben den beiden Faktoren *visuelle Strukturierung des darstellenden Mediums* und *Darbietungsart* einen signifikanten Einfluss aufweist.

Nachfolgend werden Untersuchungsdesign, Variablen und Forschungshypothesen aus Experiment 2 vorgestellt. Im Anschluss an die Vorstellung der teilnehmenden Probanden und Methoden, sowie der Untersuchungsdurchführung werden dann ausgewählte Ergebnisse aus Experiment 2 näher beschrieben und diskutiert.



## 9.1 Untersuchungsdesign

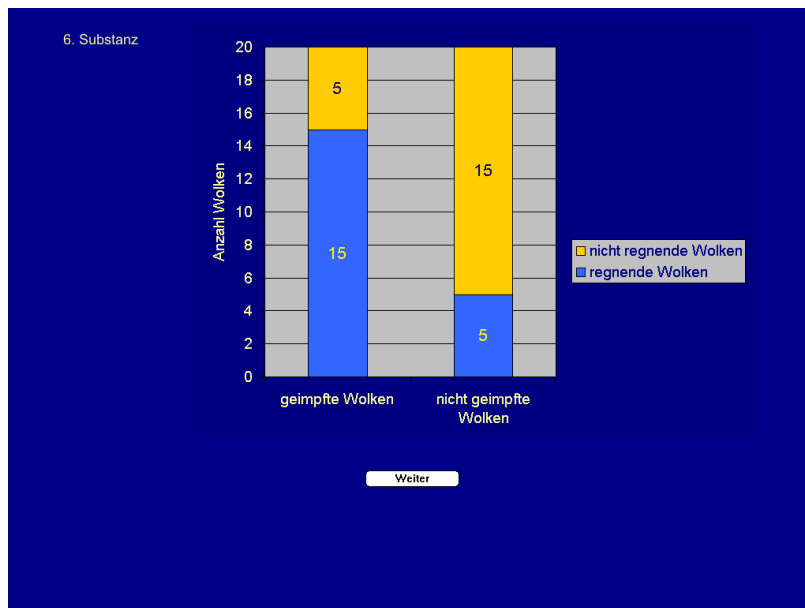
### Versuchsplan und unabhängige Variablen

Untersucht wurde der Einfluss der beiden zweistufigen Between-Faktoren *Darbietungsart* und *visuelle Strukturierung des darstellenden Mediums*, kurz *VS*. Hinsichtlich des Faktors *Darbietungsart* wurde der Einfluss von summarisch (kurz *sum*) dargebotenen Kausalinformationen mit einzelfallweise (kurz *tbt* für englisch *trial by trial*) dargebotenen Kausalinformationen verglichen, hinsichtlich des Faktors *visuelle Strukturierung des darstellenden Mediums* hoch visuell strukturierte Grafiken (*VS hoch*) mit visuell gering strukturierten Fliesstexten (*VS gering*).

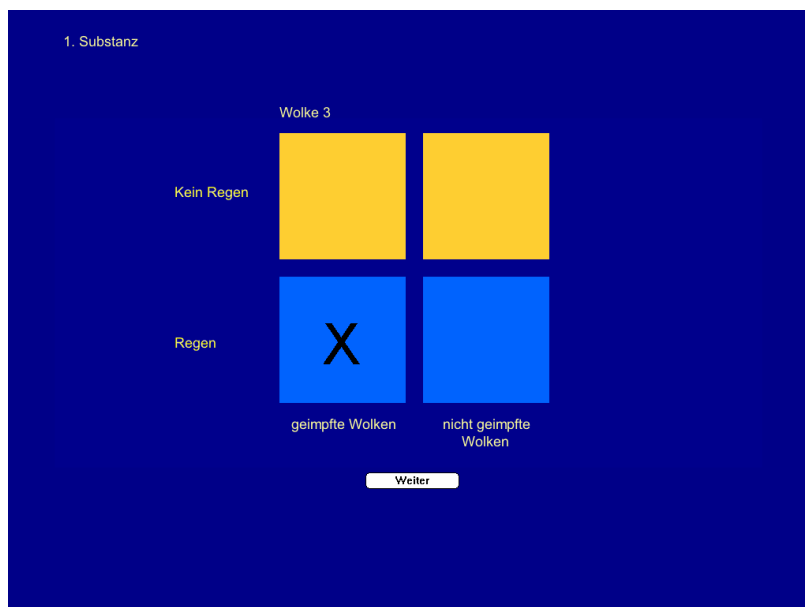
Für die summarische Ergebnisdarbietung bei hoher visueller Strukturierung haben wir ein Stapeldiagramm mit Häufigkeiten gewählt (kurz *SD sum*). Für Häufigkeiten haben wir uns zum einen entschieden, weil wir auf der Basis der Resultate aus Experiment 1 Deckeneffekte vermeiden wollten. Zum anderen sollte eine Vergleichbarkeit von Einzelfall- und summarischer Darbietung gewährleistet bleiben, indem für alle Probanden gleichermaßen Prozentberechnungen zwingend gemacht wurden, falls sie beabsichtigen sollten, die Wirksamkeiten mit Hilfe der ab-Regel oder einer regel-basierten Strategie zu bestimmen. Der Einfluss des Faktors *Informationsart* sollte so möglichst konstant gehalten werden. Für ein Stapeldiagramm haben wir uns entschieden, da dieses Medium in Experiment 1 effizienter als eine Tabelle mit Häufigkeiten verarbeitet wurde und für den Unterschied zwischen Stapeldiagrammen und Tabellen hinsichtlich der *Anzahl regel-konformer Lösungen* zumindest für die studentische Teilstichprobe ein Trend erzielt wurde. Die Ergebnisse aus Experiment 1 lassen den Schluss zu, dass Stapeldiagramme im Vergleich zu Tabellen ein zwar nicht immer signifikant höheres, aber doch tendenziell höheres Ausmaß an visueller Strukturierung aufweisen.

Bei Einzelfalldarbietung mit einem visuell hoch strukturierten Medium ist der Einsatz eines Stapeldiagramms nicht möglich, es wurde deshalb eine beschriftete und visuell hoch strukturierte farbige Vierfeldertafel entworfen, auf welcher ein einzelfallweise dargebotenes Kreuz durch seine Position auf der Vierfeldertafel anzeigt, welches Ereignis gerade beobachtet werden konnte (links unten: geimpfte Wolke, die anfängt zu regnen; links oben: geimpfte Wolke, die nicht anfängt zu regnen; rechts unten: nicht geimpfte Wolke, die anfängt zu regnen; rechts oben: nicht geimpfte Wolke, die nicht anfängt zu regnen).

Die Abbildungen 21 und 22 veranschaulichen beispielhaft ein Stapeldiagramm und eine Vierfeldertafel mit Positionskreuz und Einzelfalldarbietung (kurz *VT tbt*) wie sie in Experiment 2 eingesetzt wurden:



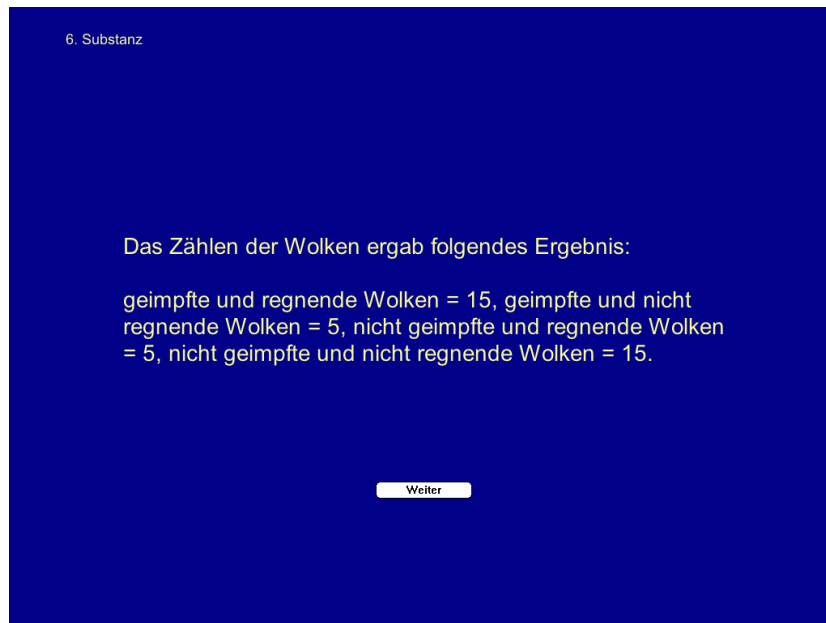
**Abbildung 20:** Experiment 2: Hohe visuelle Strukturierung des darstellenden Mediums und summarische Darbietung: Stapeldiagramm mit Häufigkeiten (kurz *SD sum*).



**Abbildung 21:** Experiment 2: Hohe visuelle Strukturierung des darstellenden Mediums und Einzelfalldarbietung: Vierfeldertafel mit Positionskreuz (kurz *VT tbt*).

Auf eine bildhafte Einzelfalldarbietung von Wolken (regnend oder nicht regnend; mit oder ohne Impfung) wurde verzichtet, da bei Bildern trotz ihrer Anschaulichkeit keine Ähnlichkeit mehr zu Stapeldiagrammen mit Häufigkeiten bestanden hätte und die Vergleichbarkeit zu Experiment 1 eingeschränkt gewesen wäre.

Die Abbildungen 22 und 23 zeigen beispielhaft Fliesstexte (kurz *FT*) bei summarischer und einzelfallweiser Darbietung der Kausalinformationen.



**Abbildung 22:** Experiment 2: Geringe visuelle Strukturierung und summarische Darbietung: Fliesstext mit Häufigkeitsangaben (kurz *FT sum*).



**Abbildung 23:** Experiment 2: Geringe visuelle Strukturierung und Einzelfalldarbietung: Fliesstext mit Aussagen zu Einzelereignissen (kurz *FT tbt*).

Aus der Kreuzung der beiden zweistufigen Faktoren *Darbietungsart* und *visuelle Strukturierung* ergab sich ein 2 x 2 - Design, das in Tabelle 8 abgebildet ist.

**Tabelle 8:** Experiment 2: 2 x 2 – Design.

	visuelle Strukturierung des darstellenden Mediums	
	hoch (Grafik)	gering (Fliesstext)
<b>Darbietungsart</b>		
summarische Darbietung	Stapeldiagramm ( <i>SD sum</i> )	Fliesstext ( <i>FT sum</i> )
Einzelfalldarbietung	Vierfeldertafel mit Positionskreuz ( <i>VT tbt</i> )	Fliesstext ( <i>FT tbt</i> )

Zur Reduktion der Fehlervarianz und zur Erhöhung von Teststärke und experimenteller Kontrolle über die Einflüsse der beiden Hauptfaktoren wurde als Kovariate das *numerische Denken* einschließlich der *numerischen Verarbeitungskapazität* eines jeden Probanden mit neun dafür ausgewählten Items aus dem Berliner Intelligenzstruktur-Test (BIS-Test, Jäger, Süß & Beauducel, 1997; Süß, 1996) erfasst.

Da es sich bei der Bestimmung der Wirksamkeiten auf der Basis von Häufigkeiten um eine Aufgabe handelt, die komplexe numerische Problemlöseprozesse erfordert, vermuten wir, dass die Fähigkeit zum numerischen Denken (insbesondere die numerische Verarbeitungskapazität) der Probanden einen Großteil der Variation in Experiment 1 verursacht hat und als Kovariate erheblichen Einfluss auf die abhängigen Variablen in Experiment 2 nehmen wird. Durch die Hinzunahme der Kovariate *numerisches Denken* verändert sich die statistische Auswertung, es kommt anstelle einer zweifaktoriellen Varianzanalyse eine zweifaktorielle Kovarianzanalyse zum Einsatz.

### Abhängige Variablen

Es wurden im Wesentlichen dieselben abhängigen Variablen erhoben wie in Experiment 1. Da Experiment 2 (wie auch Experiment 3) den Probanden rechnergestützt präsentiert wurde, konnte die Bearbeitungszeit genauer als in Experiment 1 erhoben werden. Neben der gesamten Bearbeitungszeit (*total time*) konnte durch die Anfertigung von Protokolldateien separat erfasst werden, wie lange die Probanden insgesamt jeweils vor dem dargebotenen Medium (*time on task*) und vor der anschließend präsentierten Wirksamkeitsskala (*rating time*) verweilten. Während mit der *rating time* (Bewertungszeit) ausschließlich die Dauer der kognitiven Prozesse in Verbindung mit der Urteilsbildung und der Integration gedächtnisbasierter Informationen erfasst werden, schließt die *time on task* (Zeit für die eigentliche Aufgabenbearbeitung) die Dauer von peripheren Prozessen im Zusammenhang

mit Informationsaufnahme und Wahrnehmung mit ein. Während der *time on task* können durchaus zentrale kognitive Prozesse in Zusammenhang mit der Urteilsbildung und der Integration von Vorwissen ablaufen, doch vermischen sie sich mit Prozessen und Handlungsrouinen in Verbindung mit der Datenaggregation und -zuordnung. Letztere sind mit dem Beginn der *rating time*, während der keine neuen Informationen über beobachtete Ursache-Wirkungs-Zusammenhänge dargeboten werden, abgeschlossen.

## **9.2 Forschungshypothesen**

### **9.2.1 Einfluss des Faktors Darbietungsart**

#### *Einfluss der Darbietungsart auf die Anzahl regel-konformer Lösungen*

Es wird erwartet, dass Personen, die die kausalen Informationen in summarischer Form dargeboten bekommen, signifikant häufiger regel-konforme Lösungen bei der Ermittlung der Wirksamkeit einer Ursache hervorbringen als Personen, denen die kausalen Informationen einzelfallweise dargeboten werden. Bei einzelfallweiser Informationsdarbietung fehlt den Probanden eine Gedächtnisstütze, in der die für die Berechnung der Wirksamkeiten relevanten Informationen mehr oder weniger übersichtlich angeordnet sind. Stattdessen stehen die Probanden vor der Aufgabe, die aufeinander folgenden Einzelinformationen mental nach Zellzugehörigkeit zu ordnen und aufzusummieren. Die verschiedenen Zellhäufigkeiten müssen dann bis zum Ende der Einzeldurchgänge aktiv im Gedächtnis behalten werden und entsprechend dem Vorgehen einer regel-konformen Strategie miteinander verrechnet werden. Bei einer summarischen Informationsdarbietung wird der kognitive Apparat hingegen entlastet, die Probanden müssen keine kognitive Kapazität für das Behalten, Aufsummieren und Gruppieren kausaler Informationen aufbringen und können sich ganz auf den Abruf oder das Hervorbringen einer Strategie zur Bestimmung der Wirksamkeiten konzentrieren.

Das Hervorbringen oder der Abruf einer regel-konformen Lösung erfordert ein hohes Maß an kognitiver Kapazität. Bei Einzelfalldarbietung dürfte die Verarbeitungskapazität vieler Probanden durch das zusätzliche Memorieren von relevanten Informationen an ihre Grenzen stoßen und damit zu einer kognitiven Überlastung (*cognitive overload*) führen, die das Generieren einer regel-konformen Lösung während der Aufgabenbearbeitung fast unmöglich macht. Nach der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) führt eine kognitive Überlastung oder eine hohe mentale Beanspruchung zu einem erhöhten Aufwand bei der Generierung oder beim Abruf einer regel-konformen Lösung. Diese hohen Kosten werden von den Probanden antizipiert, so dass sie sich dafür entscheiden, auf einfachere Strategien,

Heuristiken oder unklare Schätzungen zurückzugreifen. Entsprechend steigt so die mittlere Anzahl von nicht regel-konformen Lösungen und die Anzahl regel-konformer Lösungen sinkt. Alternativ kann angenommen werden, dass sich bei Einzelfalldarbietung die Anzahl von fehlerhaften oder verzerrten Kausalurteilen, die von den Vorhersagen der  $\Delta P$ -Regel oder Power PC-Theorie abweichen (Fiedler, 2000; Kao & Wasserman, 1993; Shaklee & Mims, 1982), erhöht. Bei einer Einzelfalldarbietung der für die Ermittlung der Wirksamkeit einer Ursache relevanten Informationen können regel-konforme Lösungen vermutlich nur dann erreicht werden, wenn sich die Probanden schon von vorneherein über den Einsatz einer regel-basierten Strategie im Klaren sind und genau wissen, wie sie die aufeinander folgenden Einzelinformationen zu ordnen und zueinander in Beziehung zu setzen haben. In Kurzschreibweise zusammengefasst ergibt sich folgende Hypothese:

H5a: # regel-konformer Lösungen<sub>sum</sub> > # regel-konformer Lösungen<sub>tbt</sub>

#### *Einfluss der Darbietungsart auf die Bearbeitungszeit*

Es wird erwartet, dass die mittlere Gesamtbearbeitungszeit bei Personen, die mit summarisch dargebotenen Inhalten arbeiten, geringer ausfällt als bei Personen, denen die Kausalinformationen einzelfallweise dargeboten werden, weil die Menge an zu verarbeitender Information unter der Bedingung summarische Darbietung geringer ist. Bei einzelfallweiser Darbietung geht Zeit für das im Vergleich zur summarischen Darbietung zusätzliche Ordnen, Aufsummieren und Memorieren von Einzelinformationen verloren. Im Zusammenhang mit der Kosten-Nutzen Theorie von Kleinmuntz & Schkade (1993) repräsentiert die Bearbeitungszeit den zeitlichen Aufwand der von Probanden zur Bildung eines Kausalurteils aufgebracht werden muss. Weiter nehmen wir an, dass die Unterschiede vor allem auf Unterschiede in der *time on task* zurückzuführen sind und weniger auf Unterschiede in der *rating time*.

H5b: Gesamtbearbeitungszeit<sub>sum</sub> < Gesamtbearbeitungszeit<sub>tbt</sub>

#### *Einfluss der Darbietungsart auf die Verarbeitungseffizienz*

Die Verarbeitungseffizienz pro Person wird aus dem Quotienten *Anzahl regel-konformer Lösungen / Gesamtbearbeitungszeit* ermittelt. Aus den Hypothesen H5a und H5b kann eine im Vergleich zur Einzelfallbedingung höhere mittlere Verarbeitungseffizienz bei summarischer Darbietung abgeleitet werden.

H5c: Verarbeitungseffizienz<sub>sum</sub> > Verarbeitungseffizienz<sub>tbt</sub>

*Einfluss der Darbietungsart auf die Strategiewahl*

Es wird erwartet, dass die Zahl der Probanden, die entweder von der  $\Delta P$ -Regel oder der Power PC Theorie Gebrauch machen, bei summarischer Informationsdarbietung signifikant höher liegt als bei Einzelfalldarbietung, weil bei einer summarischen Informationsdarbietung der Aufwand oder die mentale Beanspruchung für die Berechnung von  $\Delta P$  oder kausaler Stärke  $p$  geringer als bei einer einzelfallweisen Darbietung ausfällt.

H5d: # Pbn mit  $\Delta P$  oder PPC<sub>sum</sub> > # Pbn mit  $\Delta P$  oder PPC<sub>tbt</sub>

**9.2.2 Einfluss des Faktors visuelle Strukturierung des darstellenden Mediums***Einfluss der visuellen Strukturierung auf die Anzahl regel-konformer Lösungen*

Es wird erwartet, dass Personen, die mit einer visuell hoch strukturierten Grafik arbeiten, insgesamt häufiger regel-konforme Lösungen bei der Ermittlung der Wirksamkeit einer Ursache generieren als Personen, denen die kausalen Informationen in einem Fliesstext präsentiert werden. Allerdings vermuten wir weiter, dass eine hohe visuelle Strukturierung nur dann zu einer höheren Anzahl regel-konformer Lösungen führt, wenn die Informationen in summarischer Form dargeboten werden.

Bei summarischer Darbietung liefern beide Präsentationsformen die Informationen über Ursache-Wirkungs-Zusammenhänge in aggregierter und geordneter Form. Im Gegensatz zum Stapeldiagramm stellt ein einfacher Fliesstext aber keine externe Strukturierungshilfe bereit, wie die Informationen zur Bestimmung der Wirksamkeiten zueinander in Beziehung zu setzen sind. Demgegenüber erhalten die Probanden, denen Stapeldiagramme zur Aufgabenbearbeitung vorgelegt werden, durch die grafische Anordnung der Information visuelle Hinweise darauf, wie sie Erfolgs- und Basisrate zueinander in Beziehung zu setzen sollen. Durch die Gegenüberstellung von einer Säule mit relevanten Informationen zur Berechnung der Erfolgsrate und einer Säule mit relevanten Informationen zur Berechnung der Basisrate werden Vergleichsprozesse induziert, die salient machen, wie Erfolgs- und Basisrate im Sinne der  $\Delta P$ -Regel oder entsprechend der PPC-Theorie zueinander in Beziehung zu setzen sind. Indem der Vergleich zwischen Erfolgs- und Basisrate salient gemacht wird, werden regel-konforme Prozesse des Schlussfolgerns vereinfacht. Zu vermuten ist, dass ein regel-konformes Zueinander-in-Beziehung-setzen von Erfolgs- und Basisrate durch die Anordnung der Informationen im Stapeldiagramm visuell veranschaulicht, beziehungsweise nahe gelegt und partiell auf die Wahrnehmungsebene verlagert wird. Durch die Verlagerung auf die Wahrnehmungsebene können die kognitive Informationsverarbeitung entlastet und die mentale Beanspruchung reduziert (Fiedler, 2000; Wickens & Carswell, 1995), sowie

perzeptuelle Inferenzen (Larkin & Simon, 1987) ermöglicht werden. Eine mentale Aufwandsreduktion führt nach der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) dazu, dass die Probanden verstärkt von komplexeren regel-basierten Strategien Gebrauch machen und dadurch auch zu einer höheren Anzahl regel-konformer Lösungen gelangen. Ebenso kann davon ausgegangen werden, dass sich durch eine verringerte mentale Beanspruchung die Anzahl verzerrter oder fehlerhafter Kausalurteile reduziert (Shaklee & Mims, 1982; Fiedler, 2000).

Bei Einzelfalldarbietung ist hingegen anzunehmen, dass die visuelle Strukturierung des darstellenden Mediums keine Auswirkungen auf die Anzahl regel-konformer Lösungen hat. Zum einen kann vermutet werden, dass bei Einzelfalldarbietung die Aufgabenkomplexität im Vergleich zur summarischen Darbietung erhöht ist und dadurch zu hohe Anforderungen an die kognitive Kapazität der Probanden gestellt werden. Zum anderen werden im Gegensatz zur summarischen Darbietung durch die grafische Anordnung der einzelnen Informationen keine Prozesse unmittelbar nahe gelegt, die die Probanden dazu anregen oder ihnen Strukturierungshilfen an die Hand geben, wie sie Erfolgs- und Basisrate regel-konform zueinander in Beziehung setzen sollen. Durch die hohen Anforderungen an die kognitive Kapazität sind die Probanden bei einzelfallweiser Darbietung nicht in der Lage, von einer visuellen Strukturierung der Kausalinformationen zu profitieren, da sie erst einmal damit beschäftigt sind, die einzelnen Informationen zu gruppieren und aufzuaddieren. Etwas anders formuliert: Die als Folge einer einzelfallweisen Darbietung hohe mentale Beanspruchung überlagert oder verhindert eine Entlastung durch eine visuell hoch strukturierte Informationsanordnung

Zusammengefasst erwarten wir einen signifikanten Haupteffekt des Faktors *visuelle Strukturierung des darstellenden Mediums*, der auf eine Induktion, ein Salientmachen des Vergleichs von Erfolgs- und Basisrate zurückzuführen ist. Allerdings kann die Induktion eines Vergleichs zwischen Erfolgs- und Basisrate nur dann effektiv sein, wenn wie bei summarischer Ergebnisdarbietung die Beanspruchung der kognitiven Kapazität reduziert ist.

H6a: # regel-konformer Lösungen<sub>VS hoch</sub> > # regel-konformer Lösungen<sub>VS gering</sub>

#### *Einfluss der visuellen Strukturierung auf die Bearbeitungszeit*

Es wird erwartet, dass eine hohe visuelle Strukturierung des darstellenden Mediums bei beiden Darbietungsarten zu kürzeren Bearbeitungszeiten führt. Bei Einzelfalldarbietung können die Kausalinformationen von den Probanden durch die wechselnden Positionen des Kreuzes schneller gesammelt und einem der vier Ereignistypen (Wolke geimpft – Regen; Wolke nicht geimpft – Regen; Wolke geimpft – kein Regen; Wolke nicht geimpft – kein



Regen) zugeordnet werden. Bei summarischer Darbietung wird ebenfalls eine kürzere Bearbeitungszeit für die Stapeldiagramme vorhergesagt, da im Vergleich zum Fliesstext die Anforderungen an die kognitive Kapazität reduziert sind. Weiter ist zu erwarten, dass die vorhergesagten Unterschiede wesentlich auf zeitliche Unterschiede bei der eigentlichen Aufgabenbearbeitung (*time on task*) zurückzuführen sind und weniger auf zeitliche Unterschiede bei Abgabe eines Urteils über die Wirksamkeit eines Urteils (*rating time*).

H6b: Bearbeitungszeit<sub>VS hoch</sub> < Bearbeitungszeit<sub>VS gering</sub>

#### *Einfluss der visuellen Strukturierung auf die Verarbeitungseffizienz*

Aus den Hypothesen H6a und H6b kann abgeleitet werden, dass Personen, die mit visuell hoch strukturierten Grafiken arbeiten, eine im Vergleich zu Personen, die mit Fliesstexten arbeiten, höhere mittlere Verarbeitungseffizienz aufweisen werden. Weiter ist zu erwarten, dass der Unterschied bei summarischer Darbietung deutlicher ausfallen wird als bei Einzelfalldarbietung, da nur bei summarischer Informationsdarbietung Unterschiede im Hinblick auf die abhängige Variable *Anzahl regel-konformer Lösungen* vorhergesagt wurden.

H6c: Verarbeitungseffizienz<sub>VS hoch</sub> > Verarbeitungseffizienz<sub>VS gering</sub>

#### *Einfluss der visuellen Strukturierung auf die Strategiewahl*

Es wird erwartet, dass die Anzahl der Probanden, die entweder von der  $\Delta P$ -Regel oder der Power PC Theorie Gebrauch machen, bei hoher visueller Strukturierung des darstellenden Mediums höher liegt als bei einer geringen visuellen Strukturierung des darstellenden Mediums. Ebenso wie bei der abhängigen Variable *Anzahl regel-konformer Lösungen* wird auch hier die Einschränkung gemacht, dass die Anzahl regel-basiert vorgehender Probanden nur bei summarischer Informationsdarbietung signifikant erhöht sein wird, da es bei einzelfallweiser Darbietung insgesamt (unabhängig vom Einfluss des Faktors *visuelle Strukturierung des darstellenden Mediums*) nur sehr wenige Probanden geben dürfte, die von einer regel-basierten Strategie Gebrauch machen.

H6d: # Pbn mit  $\Delta P$  oder PPC<sub>VS hoch</sub> > # Pbn mit  $\Delta P$  oder PPC<sub>VS gering</sub>

### **9.2.4 Einfluss der Kovariate numerisches Denken**

Wir erwarten einen signifikanten Einfluss der Kovariate *numerisches Denken* auf alle abhängigen Variablen außer der *Bearbeitungszeit* ( $\text{Bearbeitungszeit}_{\text{ND hoch}} = \text{Bearbeitungszeit}_{\text{ND gering}}$ ). Die Bearbeitungszeiten werden vermutlich von einer Reihe weiterer Faktoren neben

dem numerischen Denken wie zum Beispiel der spezifischen Aufgabenmotivation und dem Interesse daran, sowie dem individuellen Arbeitsstil (z.B. Sorgfalt, Gewissenhaftigkeit) maßgeblich mit beeinflusst.

H7a: # regel-konformer Lösungen<sub>ND hoch</sub> > # regel-konformer Lösungen<sub>ND gering</sub>

H7b: Verarbeitungseffizienz<sub>ND hoch</sub> > Verarbeitungseffizienz<sub>ND gering</sub>

### 9.2.5 Einfluss des Faktors Studienausrichtung

Experiment 1 führte zu dem Ergebnis, dass Probanden aus Fächern mit mathematisch-naturwissenschaftlicher Ausrichtung zu mehr regel-konformen Lösungen kamen als Probanden aus geisteswissenschaftlichen Fächern, doch wurde der Unterschied entgegen den Erwartungen nicht signifikant. Immerhin konnte aber ein Trend festgestellt werden, dass der Anteil an Probanden, die durchgängig regel-basierte Strategien einsetzten bei Probanden aus mathematisch-naturwissenschaftlich ausgerichteten Fächern höher war als bei Probanden aus geisteswissenschaftlich ausgerichteten Fächern. Ein wesentlicher Grund für das Verfehlen des Signifikanzniveaus von  $\alpha = 5\%$  könnten die berichteten Deckeneffekte aus Experiment 1 sein. Für Experiment 2 erwarten wir keine Deckeneffekte mehr, da zumindest in drei von vier Bedingungen (beide Bedingungen mit Einzelfalldarbietung und summarische Bedingung mit Fliesstext) höhere Anforderungen an die kognitive Kapazität der Probanden als in Experiment 1 gestellt werden. Außerdem sollten an Experiment 2 wesentlich weniger Leibniz-Kollegiaten teilnehmen. In Experiment 1 waren die Leibniz-Kollegiaten maßgeblich am Entstehen der Deckeneffekte beteiligt gewesen.

Weil davon ausgegangen werden kann, dass Deckeneffekte in Experiment 2 ausbleiben, vermuten wir, dass der Faktor *Studienausrichtung* einen signifikanten Einfluss auf die abhängige Variable *Anzahl regel-konformer Lösungen* haben wird. Allerdings dürfte der Faktor *Studienausrichtung* nur dann signifikant werden, wenn die Kovariate *numerisches Denken* unberücksichtigt bleibt, da angenommen wird, dass Probanden aus mathematisch-naturwissenschaftlichen Studienfächern durchschnittlich höhere Intelligenzwerte beim numerischen Denken aufweisen als Probanden aus geisteswissenschaftlichen Fächern. Wenn die Kovariate *numerisches Denken* mitberücksichtigt wird, dann ist zu erwarten, dass der Faktor *Studienausrichtung* nicht signifikant wird, da ein erheblicher Teil der Treatment-Varianz durch die Kovariate erklärt wird. Bei Nichtberücksichtigung der Kovariate *numerisches Denken* ist zu erwarten:

H8a: # regel-konformer Lösungen<sub>NW</sub> > # regel-konformer Lösungen<sub>GW</sub>

Hinsichtlich der *Bearbeitungszeit* erwarten wir im Gegensatz zu Experiment 1 keine signifikanten Auswirkungen des Faktors *Studienausrichtung*, da bei dieser Variablen keine Deckeneffekte auftreten und in Experiment 1 keine signifikanten Unterschiede in Abhängigkeit von dieser Variablen zu beobachten waren. Zudem wurde vorhergesagt, dass das *numerische Denken* keinen signifikanten Einfluss auf die Bearbeitungszeiten hat, weil die Bearbeitungszeit in Verbindung mit individuellen Voraussetzungen vermutlich eher von motivationalen Faktoren und individuellen Arbeitsstilen als von der Befähigung zum *numerischen Denken* abhängt. Da wir von der Annahme ausgehen, dass sich Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung hinsichtlich ihrer Fähigkeiten im *numerischen Denken* von Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung unterscheiden, dürfte auch der Faktor *Studienausrichtung* nicht signifikant werden ( $\text{Bearbeitungszeit}_{\text{NW}} = \text{Bearbeitungszeit}_{\text{GW}}$ ).

In Verbindung mit der *Verarbeitungseffizienz* kann aus Hypothese H8a ein im Vergleich zur abhängigen Variable *Anzahl regel-konformer Lösungen* gemäßigt signifikanter Einfluss des Faktors *Studienausrichtung* vorhergesagt werden.

H8b:  $\text{Verarbeitungseffizienz}_{\text{NW}} > \text{Verarbeitungseffizienz}_{\text{GW}}$

Für den Einfluss von *Studienausrichtung* auf die Strategiewahl werden ähnliche Ergebnisse vorhergesagt wie in Experiment 1: Wir erwarten, dass die Zahl der Probanden, die von regelbasierten Strategien Gebrauch macht, in der Probandengruppe mit mathematisch-naturwissenschaftlicher Studienausrichtung höher ist als in der Probandengruppe, die nicht mathematisch-naturwissenschaftlich ausgerichtet studiert.

H8c:  $\# \text{Pbn mit } \Delta P \text{ oder PPC}_{\text{NW}} > \# \text{Pbn mit } \Delta P \text{ oder PPC}_{\text{GW}}$

### 9.3 Methodik und Vorgehensweise

#### Probanden

An Experiment 2 nahmen 80 Versuchspersonen (56 Frauen und 24 Männer) teil. Das Alter variierte zwischen 19 und 57 Jahren ( $M = 24.08$ ;  $s = 5.8$ ). Voraussetzung für die Teilnahme war erneut eine gute Beherrschung der deutschen Sprache. Alle Teilnehmer waren Studierende der Universität Tübingen oder angehende Studenten am Leibnizkolleg in Tübingen (70 Studierende, 10 Leibniz-Kollegiaten). Wie schon in Experiment 1 handelte es sich bei der einen Hälfte um Teilnehmer, die einen Studiengang absolvierten oder absolvieren wollten, in denen mathematische Fächer (wie z.B. Statistik) oder naturwissenschaftliche Methoden gelehrt wurden. Die andere Hälfte der Probanden setzte sich aus Studierenden oder angehenden Studierenden geisteswissenschaftlicher Disziplinen zusammen, in denen naturwissenschaftliche oder mathematische Methoden kaum oder gar nicht vermittelt werden. Tabelle A2 im Anhang A2.1 informiert über die Zusammensetzung der Probanden hinsichtlich ihrer Studienausrichtung und sonstiger demografischer Merkmale. Die Versuchsteilnehmer wurden extern durch Aushänge oder Ansprache angeworben und erhielten für die Teilnahme € 16,-. Die Probanden wurden einzeln untersucht und gleichmäßig nach dem Zufallsprinzip auf die vier Experimentalbedingungen verteilt.

#### Material

Wie in Experiment 1 hatten alle Probanden die Aufgabe zu beurteilen, ob und in welchem Ausmaß ein Ereignis A (die Ursache) ein Ereignis B (die Wirkung) hervorruft. Es lagen keine konkurrierenden Ursachen vor. Die Instruktion (s. Anhang A2.2) war im Wesentlichen dieselbe wie in Experiment 1, erneut kam das Wolkenimpfparadigma zum Einsatz. Unterschiede im Vergleich zu Experiment 1 ergaben sich aus der veränderten Variablenauswahl und der Ersetzung der Nachbefragung durch eine Testung der Fähigkeit zum rechnerischen Denken. Zudem wurden in Experiment 2 die Informationen über Ursache-Wirkungs-Zusammenhänge nicht mehr in Papierform dargeboten, sondern computergesteuert. Die computergestützte Versuchsumgebung wurde mit der Software Marcomedia-Director erstellt. Darüber hinaus konnten mittels Log-Dateien Eingabeinhalte und der zeitliche Ablauf des Hauptversuchs festgehalten werden.

#### Untersuchungsdurchführung

Wie schon in Experiment 1 beantworteten die Probanden zu Beginn Fragen zu Geschlecht, Alter und Studienfach (bzw. Studienwunsch), danach wurde ihnen die Instruktion ausgeteilt. Der Versuchsleiter war während des gesamten Versuchs anwesend.

Über alle Versuchsbedingungen hinweg wurde den Probanden in der Instruktion mitgeteilt, dass von einem Chemieunternehmen sechs Testflüge durchgeführt und dabei jeweils sechs verschiedene Substanzen getestet worden seien. Zusätzlich wurden sie darüber informiert, dass zu Beginn eine Probesubstanz eingesetzt worden sei. Je nach Bedingung unterschieden sich die Instruktionen wie folgt:

Wurden den Probanden die Informationen über Ursache-Wirkungs-Zusammenhänge summarisch in einem *Stapeldiagramm* dargeboten, so wurden sie darüber aufgeklärt, dass die Testergebnisse für jede Substanz auf dem Computer-Bildschirm in einem Säulendiagramm zusammengefasst seien. Zur leichteren Orientierung später auf dem Bildschirm wurde den Probanden das Diagramm selbst und sein Inhalt vorab auf einem separaten Instruktionsblatt erläutert. Durch diese Zusatzeinweisung in die Nutzung des Diagramms sollten die Teilnehmer mit dieser Präsentationsform vertraut werden.

Wurden den Probanden hingegen die Informationen über Ursache-Wirkungs-Zusammenhänge summarisch in einem *Fliesstext* vorgelegt, so wurde ihnen in der Instruktion mitgeteilt, dass die Testergebnisse für jede Substanz auf dem Computer-Bildschirm in einem kurzen Text zusammengefasst seien.

Bei summarisch dargebotenen Informationen wurde den Probanden weiter mitgeteilt, dass die zusammen gefassten Ergebnisse vier verschiedene Arten von Informationen enthalten würden:

1. *Anzahl* geimpfter Wolken, die im Anschluss an eine Impfung zu regnen begonnen haben.
2. *Anzahl* geimpfter Wolken, die im Anschluss an eine Impfung nicht zu regnen begonnen haben
3. *Anzahl* nicht geimpfter Wolken, die spontan zu regnen begonnen haben
4. *Anzahl* nicht geimpfter Wolken, die nicht zu regnen begonnen haben.

Wurden den Probanden jedoch die Informationen über Ursache-Wirkungs-Zusammenhänge einzelfallweise in einer *Vierfeldertafel mit Positionskreuz* dargeboten, so wurde den Probanden mitgeteilt, dass ihnen die Testergebnisse für jede der insgesamt vierzig beobachteten Wolken nacheinander auf dem Computer-Bildschirm in einer Vierfeldertafel dargeboten werden würden. Weiter wurden sie darüber aufgeklärt, dass durch die Platzierung eines Kreuzes auf der Vierfeldertafel für jeden der vierzig Fälle angezeigt werden würde, welches Ereignis (Wolke geimpft – Regen; Wolke geimpft – kein Regen; Wolke nicht geimpft – Regen; Wolke nicht geimpft – kein Regen) eingetroffen sei. Wie schon beim Stapeldiagramm wurde den Probanden die Vierfeldertafel und sein Inhalt vorab zur leichteren Bildschirmorientierung auf einem separaten Instruktionsblatt erläutert. Im Falle einer Einzelfalldarbietung in einem *Fliesstext* wurden die Probanden darüber informiert, dass die

Testergebnisse für jede der vierzig beobachteten Wolken nacheinander auf dem Computer-Bildschirm in einer Textzeile dargeboten werden würden. Auch bei Einzelfalldarbietung erhielten die Probanden die Mitteilung, dass die zusammen gefassten Ergebnisse vier verschiedene Arten von Informationen enthalten würden:

1. *Die Wolke* wurde geimpft und fing zu regnen an.
2. *Die Wolke* wurde geimpft und fing nicht zu regnen an.
3. *Die Wolke* wurde nicht geimpft und fing zu regnen an.
4. *Die Wolke* wurde nicht geimpft und fing nicht zu regnen an.

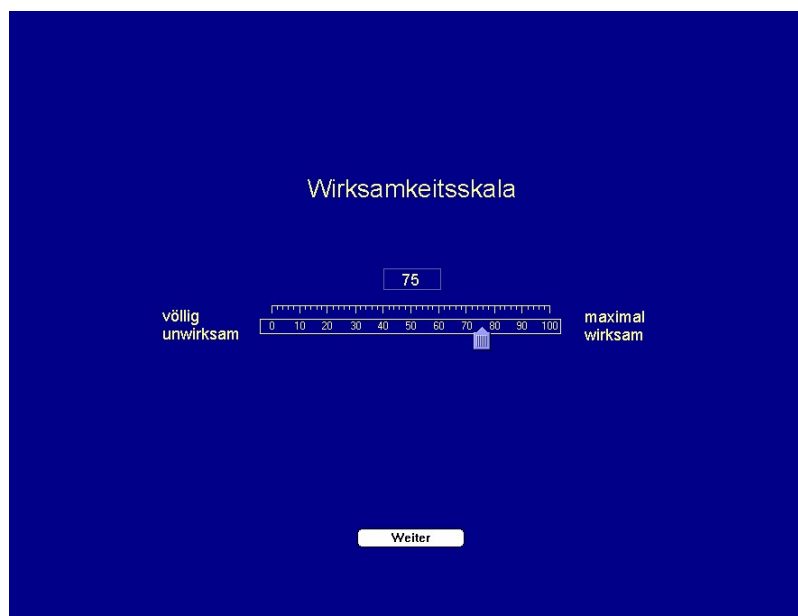
Nach dem Lesen der Instruktion sollten die Probanden eventuelle Fragen klären und die Instruktion kurz zusammenzufassen. Dann wurden allen Probanden dieselben Daten mit Angaben zu sechs verschiedenen Substanzen und vorab einer Probesubstanz präsentiert. Wie in Experiment 1 war die Reihenfolge der einzelnen Testaufgaben zufällig. Während der eigentlichen Testphase wurden den Probanden summarisch oder einzelfallweise für jede Substanz vierzig Beobachtungsergebnisse von je zwanzig geimpften und nicht geimpften, ursprünglich nicht regnenden Wolken präsentiert. Die Erfolgsrate für die Probesubstanz betrug .60, die Basisrate .20. Für die sechs Testsubstanzen wurden folgende Erfolgs- und Basisraten festgelegt: .25 vs. .25; .40 vs. .20; .50 vs. .25; .75 vs. .25; .80 vs. .20, sowie 1.00 vs. .25.

Wie bereits erwähnt erfolgte die Datenpräsentation im Gegensatz zu Experiment 1 nicht auf Papier, sondern rechnergesteuert auf einem Monitor. Dadurch konnte eine genaue zeitliche Taktung eingehalten werden. Bei Einzelfalldarbietung hatten die Probanden maximal 10 Sekunden Zeit, sich den Inhalt für jeden Einzelfall einzuprägen. Falls sie dazu weniger als 10 Sekunden benötigten, konnten sie frühestens nach einer Sekunde eine Weiertaste anklicken, um den Ablauf des Experiments zu beschleunigen, ansonsten wurde nach 10 Sekunden und einer Pause von 500 Millisekunden automatisch der nächste Einzelfall dargeboten. Durch die im Vergleich zu anderen Untersuchungen (Anderson & Sheu, 1995, Ex 1; Arkes & Harkness, 1983, Ex 1, 2, 3, 6, 7; Kao & Wasserman, 1993, Ex 2) längere Darbietungszeit sollte den Probanden genug Zeit gegeben werden, sich die Inhalte auf dem Bildschirm einzuprägen, außerdem sollte eine Vergleichbarkeit zur summarischen Bedingung hergestellt werden, bei welcher die Probanden für jede einzelne Substanz die summarische Ergebniszusammenfassung maximal sieben Minuten betrachten konnten (sieben Minuten entsprechen 420 Sekunden = 40 x 10.5 Sekunden). Wie die Probanden, denen die Informationen einzelfallweise dargeboten wurden, so konnten auch die Probanden, denen die Informationen in summarischer Form dargeboten wurden, nach Ablauf von einer Sekunde

eine Weitertaste anklicken, um auf der nächsten Bildschirmseite die Wirksamkeit der jeweiligen Substanz einzuschätzen.

Durch die computerunterstützte Aufgabenpräsentation und die Einführung einer konstanten zeitlichen Obergrenze bei der Aufgabenbearbeitung ("*time on task*") über alle Versuchsbedingungen hinweg unterscheidet sich Experiment 2 auch von Experiment 1, bei dem die Probanden die Aufgaben auf Papier ohne Zeitbegrenzung bearbeiten konnten. Zudem war es den Probanden in Experiment 2 nicht erlaubt, sich Notizen zu machen. Durch die Einführung einer zeitlichen Obergrenze und die fehlende Möglichkeit, sich Notizen zu machen, sollte zum einen die Vergleichbarkeit zwischen allen Versuchsbedingungen in Experiment 2 gewährleistet sein, zum anderen sollten die in Experiment 1 aufgetretenen Deckeneffekte für die beiden summarischen Bedingungen vermieden werden.

Wie in Experiment 1 sollten die Probanden im Anschluss an die Datenpräsentation für jede Substanz die Wirksamkeit der jeweiligen Substanz bestimmen. Zur Einschätzung der Wirksamkeit einer Substanz konnten die Probanden mit Hilfe eines Schiebereglers auf dem Bildschirm einen Wert auf einer wie in Experiment 1 von 0 bis 100 reichenden Skala eingeben. Abbildung 24 zeigt die in den Experimenten 2 und 3 über alle Versuchsbedingungen hinweg eingesetzte Skala zur Beurteilung der Wirksamkeit einer Substanz. Die Probanden konnten erst dann zur nächsten Substanz übergehen, wenn sie sich auf einen Wert festgelegt hatten, es gab keine zeitliche Obergrenze für das Ausfüllen der Skala ("*rating time*").



**Abbildung 24:** Experiment 2: Ratingskala zur Wirksamkeitsbeurteilung.

Wenn die Probanden mit der Bearbeitung der Aufgaben für die sechs Testsubstanzen fertig waren, erfolgte wie schon in Experiment 1 eine kurze schriftliche und mündliche Nachbefragung zur Vorgehensweise bei der Aufgabenbearbeitung (s. Anhang A2.3). Nach Durchführung des eigentlichen Experiments (Dauer ca. 45 Minuten) hatten sich die Probanden noch einem Test zum numerischen Denken mit neun Items aus dem Berliner Intelligenzstruktur-Test (BIS-Test; Jäger, Süß & Beauducel, 1997) zu unterziehen. Pro Proband wurden für alle Untersuchungen ca. 90 Minuten benötigt.

Der BIS-Test wurde verwendet, weil er eine gezielte Messung verschiedener Teilkomponenten der Intelligenz erlaubt. Das Intelligenzstrukturmodell, welches dem BIS-Test zugrunde liegt, beruht unter anderem auf einem Mehrmodalitätsprinzip, wonach Intelligenztestleistungen und Fähigkeitskonstrukte nach Inhalten (figural-bildhaft, verbal, numeral) und Operationen (Einfallsreichtum, Merkfähigkeit, Verarbeitungskapazität, Bearbeitungsgeschwindigkeit) klassifiziert werden können. Als inhaltliche Fähigkeit sollte für unsere Untersuchungszwecke nur numerisches Denken (N) in Verbindung mit den drei operationalen Fähigkeiten Verarbeitungskapazität (K – 5 Items), Merkfähigkeit (M – 1 Item) und Bearbeitungsgeschwindigkeit (B – 3 Items) erhoben werden. Unter "Verarbeitungskapazität" verstehen Jäger (1984, S.30f) und Süß (1996, S. 36) die Fähigkeit zur "Verarbeitung komplexer Informationen, die nicht auf Anhieb zu lösen sind, sondern Heranziehen, Verfügbarhalten, vielfältiges Beziehungsstiften, formallogisch exaktes Decken und sachgerechtes Beurteilen von Informationen erfordern" (Süß, 1996, S. 36). "Merkfähigkeit" beinhaltet "aktives "Einprägen und kurzfristiges Wiedererkennen oder Reproduzieren von verschiedenartigem Material"; unter "Bearbeitungsgeschwindigkeit" verstehen Jäger und Süß "Arbeitstempo, Auffassungsleichtigkeit und Konzentrationskraft beim Lösen einfach strukturierter Aufgaben von geringem Schwierigkeitsniveau" (Süß, 1996, S.36). Ursprünglich wurde der BIS-Test für Jugendliche und junge Erwachsene mit Ober- und Mittelschulbildung entwickelt und kam in Verbindung mit der Prädiktion von Strategiewahl, Lernerfolg und Problemlösekompetenz schon in mehreren Forschungsarbeiten zur Erfassung der kognitiven Verarbeitungskapazität für numerische Inhalte zum Einsatz (Bröder & Eichler, 2001; Friege & Lind, 2003; Süß, 1996, 1999). Neben einer Kurzbeschreibung (s.a. Süß, 1996, S. 230 ff.) befinden sich die neun aus dem BIS-Test ausgewählten Items im Anhang A4.



## 9.4 Explorative Datenanalysen

Die statistischen Analysen wurden erneut mit dem Programmpaket SPSS durchgeführt. Die Voraussetzung der Normalverteilung wurde mit Hilfe des Lilliefors-Tests und des Shapiro-Wilks-Tests, die Voraussetzung der Varianzhomogenität wurde mit dem Levene-Test überprüft. Hinsichtlich der Voraussetzung der Normalverteilung wurden vereinzelt Verstöße festgestellt, die im Fall der abhängigen Variable *Anzahl regel-konformer Lösungen* auf Bodeneffekte bei Einzelfalldarbietung sowie summarischer Darbietung mit Text zurückzuführen ist. Auch bei allen erhobenen Zeitmaßen ergaben sich Häufungen im unteren Bereich und vereinzelt Ausreißern im oberen Bereich mit zum Teil beträchtlichen Unterschieden in Abhängigkeit von den vier Untersuchungsbedingungen. Deckeneffekte blieben diesmal aus, auch wenn in der Bedingung *SD sum* knapp die Hälfte der Probanden zu regel-konformen Lösungen bei der Aufgabenbearbeitung gelangte. Die Ergebnisse des Levene-Tests weisen auf heterogene Fehlervarianzen aller abhängigen Variablen hin. Die fehlende Homogenität der Fehlervarianzen ist vermutlich auf unterschiedliche Einflüsse der Versuchsbedingungen auf die abhängigen Variablen zurückzuführen. Eine Mitberücksichtigung des Faktors *Studienausrichtung* wirkte sich homogenisierend auf die Fehlervarianzen aus. Trotz der Verstöße haben wir wegen der Robustheit der Varianzanalyse gegenüber Verletzungen ihrer Voraussetzungen Varianzanalysen durchgeführt, da ausreichende Zellgrößen ( $n = 20$ ) und gleichgroße Stichproben vorlagen.

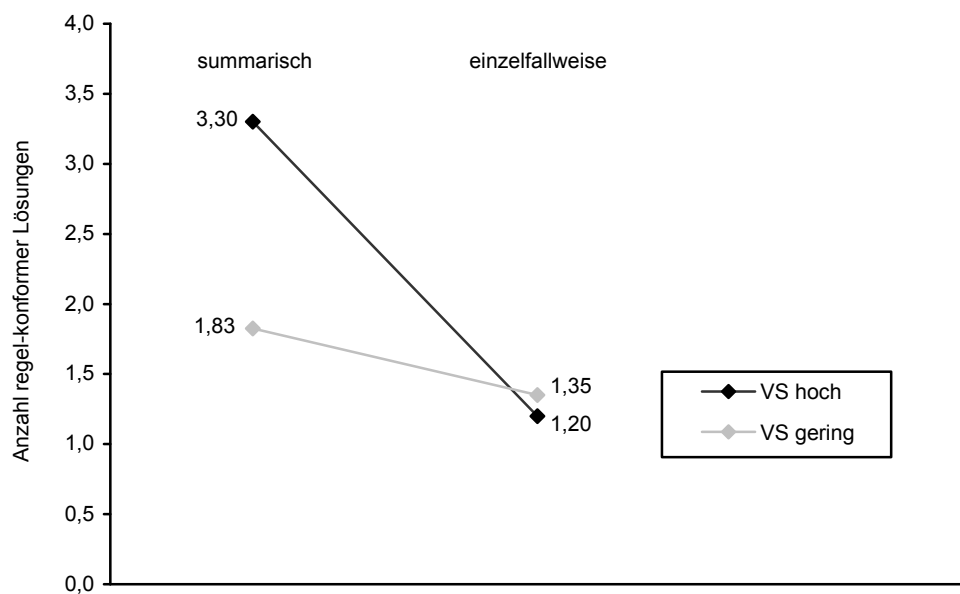
Für die Kovariate *numerisches Denken* ergaben sich in keiner Versuchsbedingung Verletzungen der Voraussetzungen der Varianzanalyse, was darauf hinweist, dass die gemessenen Werte zum numerischen Denken sich in allen vier Versuchsbedingungen normal verteilten und homogene Fehlervarianzen aufwiesen. Weitere Analysen ergaben, dass sich die vier Mittelwerte im *numerischen Denken* in Abhängigkeit von der Versuchsbedingung nicht signifikant voneinander unterschieden;  $F_{3,76} = .532$ ,  $MSE = 41.445$ ;  $p = .662$ .

## 9.5 Einfluss der Faktoren Darbietungsart und visuelle Strukturierung

### 9.5.1 Ergebnisse

#### Anzahl regel-konformer Lösungen

Abbildung 25 veranschaulicht die Mittelwerte für die *Anzahl regel-konformer Lösungen* in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*. In der angehängten Tabelle können zudem Mediane und Modalwerte eingesehen werden.



	summarisch	einzelfallweise	Gesamt
<b>visuelle Strukturierung hoch</b>			
Arithmetisches Mittel	3.3 (2.6)	1.2 (1.47)	<b>2.25</b> (2.34)
Median	4	1	1
Modalwert	6	1	0
<b>visuelle Strukturierung gering</b>			
Arithmetisches Mittel	1.83 (2.01)	1.35 (1.84)	<b>1.59</b> (1.92)
Median	1	0.5	1
Modalwert	1	0	0
<b>Gesamt</b>			
Arithmetisches Mittel	<b>2.56</b> (2.41)	<b>1.275</b> (1.65)	<b>1.92</b> (2.15)
Median	2	1	1
Modalwert	0	0	0

**Abbildung 25:** Experiment 2: Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*.

Für die abhängige Variable *Anzahl regel-konformer Lösungen* wurde eine Kovarianzanalyse mit den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung* sowie der Kovariate *numerisches Denken* durchgeführt. Die Ergebnisse lassen sich aus Tabelle 9 entnehmen.

**Tabelle 9:** Experiment 2: Ergebnisse aus den 2 (Darbietungsart) x 2 (visuelle Strukturierung) Kovarianzanalysen für die abhängige Variable *Anzahl regel-konformer Lösungen*. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen.

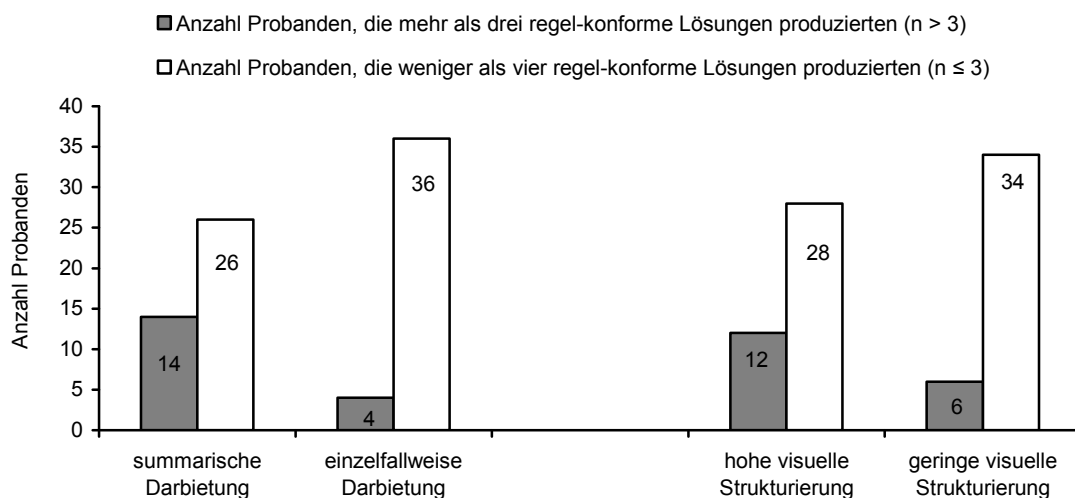
Faktor	F-Wert	MSE	P-Wert
Darbietungsart <sub>(1,75)</sub>	9.032	3.532	<b>.004</b>
visuelle Strukturierung <sub>(1,75)</sub>	3.382		<u>.070</u>
Darbietungsart x visuelle Strukturierung <sub>(1,75)</sub>	2.187		.147
numerisches Denken <sub>(1,75)</sub>	12.937		<b>&lt; .001</b>

Aus Tabelle 9 ist ersichtlich, dass der Faktor *Darbietungsart* hochsignifikant wurde, das heißt, dass die Probanden bei summarischer Darbietung der Kausalinformationen häufiger regel-konforme Lösungen einsetzten als bei Einzelfalldarbietung. Ebenso wurde auch die Kovariate *numerisches Denken* hochsignifikant.

Der Einfluss des Faktors *visuelle Strukturierung* lässt sich hingegen wie erwartet nicht eindeutig interpretieren. Ein Blick auf Abbildung 25 legt nahe, dass der Trend des Faktors *visuelle Strukturierung* ausschließlich auf die im Vergleich zu den anderen Experimentalbedingungen höhere mittlere Anzahl regel-konformer Lösungen unter der Bedingung *SD sum* zurückzuführen ist. Um feststellen zu können, ob der Faktor *visuelle Strukturierung* - wie in Hypothese H6a vorhergesagt - bei summarischer Informationsdarbietung signifikant wird und bei Einzelfalldarbietung nicht, wurden ohne Berücksichtigung der Kovariate *numerisches Denken* a priori Einzelkontraste berechnet, in denen die Versuchsbedingung Stapeldiagramm (*SD sum*) mit der Bedingung Fliesstext summarisch (*FT sum*), sowie die Bedingung Vierfeldertafel (*VT tbt*) mit der Bedingung Fliesstext einzelfallweise (*FT tbt*) verglichen wurde. Für die abhängige Variable Anzahl regel-konformer Lösungen wurde der Kontrast *SD sum* – *FT sum* erwartungsgemäß signifikant ( $p = .024$  bei angenommener Varianzgleichheit bzw.  $p = .052$  bei angenommener Varianzungleichheit), während der Kontrast *VT tbt* – *FT tbt* klar nicht signifikant wurde ( $p > .77$ ). Durchgeführte a posteriori Kontraste (Student-Newman-Keuls-Test und Tukey honestly significant difference Test) zeigten, dass die beiden Einzelfallbedingungen zusammen mit *FT sum* eine homogene Untergruppe bildeten (SNK:  $p = .593$ ; Tukey HSD:  $p = .763$ ), während *SD sum* für sich allein stand (SNK:  $p = 1.00$ ) oder gerade noch gemeinsam mit *FT sum* eine homogene Untergruppe formte (Tukey HSD:  $p = .105$ ). Die Einzelkontraste sprechen dafür,

dass der Faktor *visuelle Strukturierung* in Experiment 2 nur in Kombination mit einer summarischen Darbietung der Kausalinformationen zu einer signifikant höheren Anzahl regel-konformer Lösungen führt, nicht aber bei Einzeldarbietung.

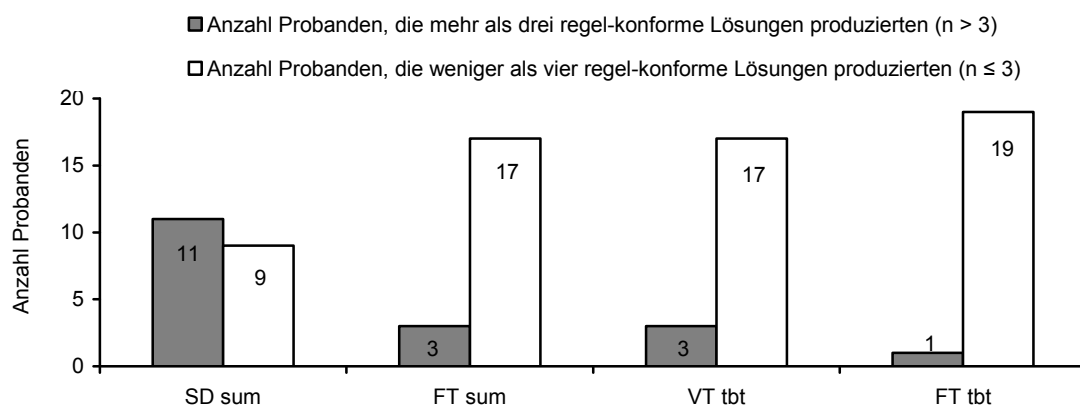
Rückschlüsse auf den Einfluss der Faktoren *Darbietungsart* und *visuelle Strukturierung* auf die *Anzahl regel-konformer Lösungen* können auch durch Häufigkeitsanalysen gewonnen werden. Diese haben wir zusätzlich durchgeführt, weil in Verbindung mit der abhängigen Variable *Anzahl regel-konformer Lösungen* die Voraussetzungen der Varianzanalyse verletzt wurden und vor allem auch weil Median und Modalwert unter der Bedingung *SD sum* deutlich höher ausfielen als unter den anderen drei Bedingungen. Zu diesem Zweck wurden alle Probanden dahingehend dichotomisiert, ob sie mehr als drei regel-konforme Lösungen produzierten oder nicht. Untersucht werden sollte, wie viele Probanden in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung* mehr als drei ( $n > 3$ ) oder weniger als vier ( $n \leq 3$ ) regel-konforme Lösungen produzierten. Abbildung 26 veranschaulicht, wie viele Probanden jeweils in Abhängigkeit von den Faktoren *Darbietungsart* und *visuelle Strukturierung* mehr als drei oder weniger als vier regel-konforme Lösungen hervorbrachten.



**Abbildung 26:** Experiment 2: Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*.

Zur Überprüfung der Annahme, ob sich die Anzahl der Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen unabhängig von den beiden Variablen *Darbietungsart* und *visuelle Strukturierung* verteilten, wurden je ein  $\chi^2$ -Test und ein exakter Test nach Fischer durchgeführt. Für den Zusammenhang zwischen *Darbietungsart* und der Anzahl der Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen

fürten sowohl der  $\chi^2$ -Test als auch der exakte Test nach Fischer erwartungsgemäß zu dem hoch-signifikanten Ergebnis, dass bei summarischer Darbietung die Anzahl von Probanden mit mehr als drei regel-konformen Lösungen höher war als bei Einzelfalldarbietung ( $\chi^2 = 7.168$ ,  $p_{\text{asymptotische Signifikanz}} = .007$ ;  $p_{\text{Fisher, einseitig}} = .007$ ). Für den Zusammenhang zwischen visueller Strukturierung und der Anzahl der Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen ergab sich hingegen lediglich eine Trendnähe (bzw. ein Trend), dass die Anzahl von Probanden mit mehr als drei regel-konformen Lösungen bei hoher visueller Strukturierung des darstellenden Mediums etwas höher lag als bei geringer visueller Strukturierung des darstellenden Mediums ( $\chi^2 = 2.581$ ,  $p_{\text{asymptotische Signifikanz}} = .108$ ;  $p_{\text{Fisher, einseitig}} = .090$ ). Eine weitere Häufigkeitsanalyse, mit welcher überprüft wurde, ob sich die Anzahl der Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen unabhängig von der Versuchsbedingung verteilte, belegte eindrucksvoll, dass unter der Bedingung *SD sum* die Anzahl der Probanden, die mehr als drei regel-konforme Lösungen produzieren, hochsignifikant höher war als unter den drei anderen Bedingungen *FT sum*, *VT tbt* und *FT tbt* ( $\chi^2 = 16.918$ ,  $p_{\text{asymptotische Signifikanz}} = .001$ ). Abbildung 27 veranschaulicht, wie sich die Anzahl der Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen in Abhängigkeit von der Versuchsbedingung verteilte:



**Abbildung 27:** Experiment 2: Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen in Abhängigkeit von der Versuchsbedingung.

Analog zur Berechnung paarweiser Einzelkontraste bei varianzanalytischen Untersuchungen zeigt Tabelle 10 die Ergebnisse von Häufigkeitsanalysen, in denen untersucht wurde, ob die Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen von je zwei Versuchsbedingungen unabhängig ist.

**Tabelle 10:** Experiment 2: Ergebnisse aus Häufigkeitsvergleichen, in denen untersucht wurde, ob die Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen von je zwei Versuchsbedingungen unabhängig ist oder nicht. Signifikante P-Werte ( $\alpha < .05$ ) sind **fett** markiert.

Vergleich	$\chi^2$ -Wert	df	Asymptotische Signifikanz (2-seitig)	exakte Signifikanz (exakter Test nach Fisher)
SD sum – FT sum	7.033	1	<b>.008</b>	<b>.009</b>
SD sum – VT tbt	11.905	1	<b>&lt; .001</b>	<b>&lt; .001</b>
SD sum – FT tbt	7.033	1	<b>.008</b>	<b>.009</b>
FT sum – VT tbt	1.111	1	.292	.302
FT sum – FT tbt	0.000	1	1.000	1.000
VT tbt – FT tbt	1.111	1	.292	.302

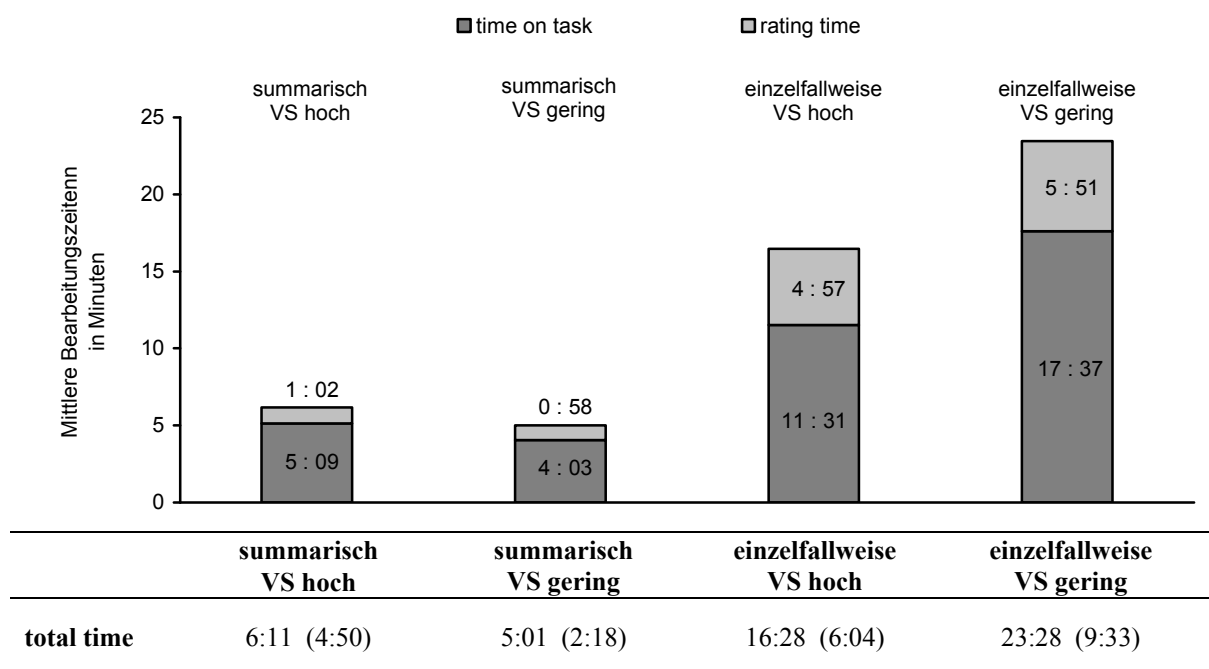
Tabelle 10 zeigt deutlich die Sonderstellung der Bedingung *SD sum* im Vergleich zu den drei anderen Bedingungen an. Während alle Häufigkeitsanalysen in Abhängigkeit von paarweisen Kombinationen mit der Bedingung *SD sum* hochsignifikant wurden, verteilte sich die Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen unabhängig von den drei paarweisen Kombinationen ohne die Versuchsbedingung *SD sum*. Die Interpretierbarkeit der Vergleiche ohne die Bedingung *SD sum* ist allerdings eingeschränkt, da wegen der Bodeneffekte in den drei anderen Bedingungen der Anteil an Zellen mit erwarteten Häufigkeiten < fünf 50 % betrug.

Insgesamt lassen sich die Ergebnisse für die abhängige Variable *Anzahl regel-konformer Lösungen* wie folgt zusammenfassen: Die Hypothese H5a kann beibehalten werden. Bei summarischer Informationsdarbietung werden im Vergleich zur Einzelfalldarbietung häufiger regel-konforme Lösungen produziert. In Verbindung mit dem Faktor *visuelle Strukturierung* bestätigen die Ergebnisse auch die Hypothese H6a. Nur bei summarischer Darbietung und nicht bei Einzelfalldarbietung wirkt sich eine hohe im Vergleich zu einer geringen visuellen Strukturierung des darstellenden Mediums dahingehend aus, dass die Probanden im Durchschnitt eine höhere Anzahl regel-konformer Lösungen produzieren. Durchgeführte Häufigkeitsanalysen, in denen die Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen in Abhängigkeit von der *Darbietungsart* und der *visuellen Strukturierung*, sowie den vier Versuchsbedingungen *SD sum*, *FT sum*, *VT tbt* und *FT tbt* untersucht wurde, unterstreichen sowohl den starken Einfluss des Faktors *Darbietungsart*, als auch den schwächeren, ausschließlich auf die Wirkung der Bedingung *SD sum* beschränkten Einfluss des Faktors *visuelle Strukturierung* auf die *Anzahl regel-konformer Lösungen*. Schließlich konnte im Zusammenhang mit der Kovariate *numerisches Denken* die Hypothese H7a klar bestätigt werden. Inhaltlich besagt dies, dass Probanden mit höheren

Fertigkeiten im numerischen Denken eine höhere Anzahl regel-konformer Lösungen produzieren.

### Bearbeitungszeit

Die mittleren Bearbeitungszeiten in Abhängigkeit von den beiden unabhängigen Variablen wurden mit der Hilfe von Stapeldiagrammen in Abbildung 28 veranschaulicht. Jede Säule steht für die Gesamtbearbeitungszeit, die sich aus den beiden Komponenten "time on task" und "rating time" zusammensetzt.



**Abbildung 28:** Experiment 2: Mittlere Bearbeitungszeit in Minuten : Sekunden (in Klammern Standardabweichungen) in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*.

Für die Gesamtbearbeitungszeit ("*total time*") und ihre beiden Teilkomponenten *time on task* und *rating time* wurden Kovarianzanalysen mit den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*, sowie der Kovariate *numerisches Denken* berechnet. Die Ergebnisse dieser drei Analysen finden sich in Tabelle 11:

**Tabelle 11:** Experiment 2: Ergebnisse aus den 2 x 2 Kovarianzanalysen, in denen der Einfluss der Faktoren *Darbietungsart* und *visuelle Strukturierung*, sowie der Kovariate *numerisches Denken* auf die drei Zeitmaße *total time*, *time on task* und *rating time* untersucht wurde. Signifikante P-Werte sind **fett** gedruckt.

	Faktor	F-Wert	MSE	P-Wert
<b>total time</b>	Darbietungsart <sub>(1,75)</sub>	104.094	39.59	< <b>.001</b>
	visuelle Strukturierung <sub>(1,75)</sub>	4.42		<b>.039</b>
	Darbietungsart x visuelle Strukturierung <sub>(1,75)</sub>	8.58		<b>.005</b>
	numerisches Denken <sub>(1,75)</sub>	.184		.669
<b>time on task</b>	Darbietungsart <sub>(1,75)</sub>	124.801	15.919	< <b>.001</b>
	visuelle Strukturierung <sub>(1,75)</sub>	8.053		<b>.006</b>
	Darbietungsart x visuelle Strukturierung <sub>(1,75)</sub>	16.475		< <b>.001</b>
	numerisches Denken <sub>(1,75)</sub>	.029		.866
<b>rating time</b>	Darbietungsart <sub>(1,75)</sub>	26.58	14.477	< <b>.001</b>
	visuelle Strukturierung <sub>(1,75)</sub>	.249		.619
	Darbietungsart x visuelle Strukturierung <sub>(1,75)</sub>	.345		.559
	numerisches Denken <sub>(1,75)</sub>	.265		.609

Im Hinblick auf die *Gesamtbearbeitungszeit* zeigen die Ergebnisse zwei signifikante Haupteffekte und eine signifikante Interaktion. Die Haupteffekte besagen inhaltlich, dass die Probanden sowohl bei einer summarischen Informationsdarbietung als auch bei einer hohen visuellen Strukturierung des darstellenden Mediums signifikant weniger Zeit benötigen. Die signifikante Interaktion kommt dadurch zustande, dass die Probanden unter der Bedingung *VT tbt* durchschnittlich sieben Minuten schneller arbeiteten als unter der Bedingung *FT tbt*, während sie unter der Bedingung *SD sum* über eine Minute langsamer waren als unter der Bedingung *FT sum*. Erwartungsgemäß wirken sich die beiden Hauptfaktoren und ihre Interaktion vor allem auf die Zeit für die eigentliche Aufgabenbearbeitung (*time on task*) aus. Für die abhängige Variable *rating time* wurde nur der Faktor *Darbietungsart* signifikant, das heißt die Probanden brauchen bei Einzelfalldarbietung mehr Zeit zur Beurteilung der Wirksamkeiten als bei summarischer Darbietung. Der Faktor *visuelle Strukturierung des darstellenden Mediums* hat ebenso wie die Interaktion keinen bedeutenden Einfluss auf die abhängige Variable *rating time*. Der nicht signifikante Einfluss der Kovariate *numerisches Denken* auf die Bearbeitungszeit und ihre beiden Teilkomponenten entsprach ebenfalls den Erwartungen, da die Bearbeitungszeit in Verbindung mit individuellen Voraussetzungen vermutlich eher von motivationalen Faktoren und individuellen Arbeitsstilen als von der Befähigung zum numerischen Denken abhängt.



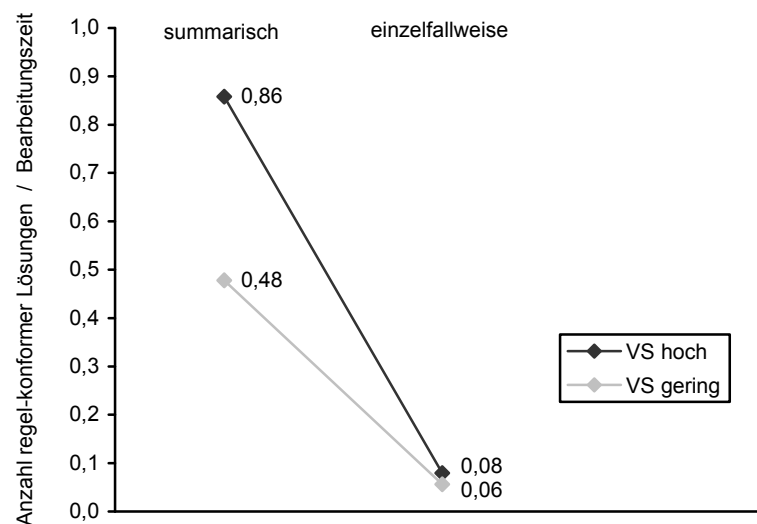
Die beobachteten Ergebnisse für die Gesamtbearbeitungszeit entsprechen weitgehend den gemachten Vorhersagen, die Hypothesen H5b und H6b können beibehalten werden. Ebenfalls den Erwartungen entsprach es, dass die Kovariate *numerisches Denken* keinen signifikanten Einfluss auf die Bearbeitungszeit hatte. Bei Einzelfalldarbietung und geringer visueller Strukturierung des darstellenden Mediums benötigen die Probanden für die Aufgabenbearbeitung mehr Zeit als bei summarischer Darbietung und hoher visueller Strukturierung, wobei Probanden mit hohen Fertigkeiten im numerischen Denken weder schneller noch langsamer sind als Probanden mit geringen Fertigkeiten im numerischen Denken. Überraschend war lediglich die signifikante Interaktion, da es nicht unseren Erwartungen entsprach, dass die Probanden bei summarischer Informationsdarbietung Fliesstexte schneller bearbeiten würden als Stapeldiagramme. A priori und posteriori Kontraste zeigen zwar, dass der Unterschied zwischen den beiden summarischen Bedingungen nicht signifikant ist (Tukey HSD:  $p = .936$ ), aber die Tatsache, dass summarische Kausalinformationen in Textform schneller bearbeitet werden als solche in einem Stapeldiagramm überrascht doch, zumal Stapeldiagramme in Experiment 1 schneller bearbeitet wurden als Tabellen und die Protokollliste. Ein möglicher Grund hierfür könnte darin liegen, dass Probanden mit Texten in der Regel hoch vertraut sind und ihnen schneller Informationen entnehmen können als den im allgemeinen weniger vertrauten Diagrammen. Durch das hohe Ausmaß an visueller Strukturierung im Stapeldiagramm werden zwar Inferenz- und Vergleichsprozesse vereinfacht, was sich in einer im Vergleich zur summarischen Textbedingung höheren mittleren Anzahl regelkonformer Lösungen zeigt, nicht aber Prozesse der Informationsaufnahme. Es scheint im Gegenteil so zu sein, dass die Aufnahme summarischer Kausalinformationen durch die zwar sinnvolle, aber im Vergleich zu Fliesstexten doch relativ komplexe Anordnung der Informationen geringfügig erschwert wird.

Im Gegensatz zu den summarischen Bedingungen unterscheiden sich die beiden Einzelfallbedingungen im Hinblick auf die Bearbeitungszeit hoch signifikant voneinander. Dies belegen sowohl a priori ( $p < .01$ ) als auch a posteriori Einzelkontraste (Tukey HSD:  $p = .004$ ). Während sich die beiden Einzelfallbedingungen im Hinblick auf die *Anzahl regelkonformer Lösungen* nicht unterscheiden, so bearbeiten die Probanden eine Vierfeldertafel hoch signifikant schneller als einen Fliesstext. Bei Einzelfalldarbietung hat die visuelle Strukturierung des darstellenden Mediums keine Auswirkungen auf das Ziehen von Schlussfolgerungen und das richtige Zueinander-in-Beziehung-setzen von Erfolgs- und Basisrate, aber sie beeinflusst die Geschwindigkeit, mit der die Probanden vorgehen.

Offensichtlich wirkt sich eine hohe visuelle Strukturierung des darstellenden Mediums beschleunigend auf Prozesse der Informationsaufnahme aus. Dieser begünstigende Einfluss gilt aber nur dann, wenn die mentale Beanspruchung wie bei Einzelfalldarbietung hoch ist. Bei summarischer Darbietung und geringer mentaler Beanspruchung müssen die Probanden die einzelnen Informationen nicht mehr bündeln und ordnen.

### Verarbeitungseffizienz

Abbildung 29 veranschaulicht die Ergebnisse für das Effizienzmaß *Anzahl regel-konformer Lösungen / Bearbeitungszeit* in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*.



	summarisch	einzelfallweise	Gesamt
visuelle Strukturierung hoch	.86 (.86)	.08 (.11)	.47 (.72)
visuelle Strukturierung gering	.48 (.63)	.06 (.08)	.27 (.49)
Gesamt	.67 (.77)	.07 (.09)	.37 (.62)

**Abbildung 29:** Experiment 2: Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*.

Erneut wurde eine Kovarianzanalyse mit den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*, sowie der Kovariate *numerisches Denken* berechnet. Die Ergebnisse der beiden Analysen sind in Tabelle 12 abgebildet:

**Tabelle 12:** Experiment 2: Ergebnisse aus den 2 (Darbietungsart) x 2 (visuelle Strukturierung) Kovarianzanalysen für das Effizienzmaß *Anzahl regel-konformer Lösungen / Bearbeitungszeit*. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen.

Faktor	F-Wert	MSE	P-Wert
Darbietungsart <sub>(1,75)</sub>	27.091	.261	< <b>.001</b>
visuelle Strukturierung <sub>(1,75)</sub>	3.938		<u>.051</u>
Darbietungsart x visuelle Strukturierung <sub>(1,75)</sub>	1.382		.243
numerisches Denken <sub>(1,75)</sub>	9.362		<b>.003</b>

Wie Tabelle 12 zeigt, wurde der Faktor *Darbietungsart* ebenso wie die Kovariate *numerisches Denken* hochsignifikant. Für den Faktor *visuelle Strukturierung des darstellenden Mediums* ergab sich ein Trend. Die Interaktion wurde nicht signifikant, auch wenn deutlich wurde, dass der Einfluss des Faktors *visuelle Strukturierung des darstellenden Mediums* bei summarischer Informationsdarbietung ausgeprägter ist als bei Einzelfalldarbietung. Durchgeführte a priori und a posteriori Einzelkontraste ohne die Kovariate *numerisches Denken* bestätigen, dass die Unterschiede zwischen den beiden summarischen Bedingungen entweder signifikant werden (a priori:  $p < .050$ ) oder das Signifikanzniveau von  $\alpha = 5\%$  relativ knapp verfehlen (Tukey HSD:  $p < .17$ ); diejenigen zwischen den beiden Einzelfallbedingungen werden dagegen klar nicht signifikant (a priori:  $p > .87$ ; Tukey HSD:  $p > .99$ ).

Die für die Einflüsse des Faktors *Darbietungsart* und der Kovariate *numerisches Denken* gemachten Vorhersagen entsprechen den Erwartungen, die Hypothesen H5c und H7b können also beibehalten werden. Der Einfluss des Faktors *visuelle Strukturierung des darstellenden Mediums* kann hingegen nicht eindeutig interpretiert werden, weil je nach Darbietungsart differenziert werden muss. Insgesamt sprechen die Ergebnisse dafür, dass die Hypothese H6c bei summarischer Darbietung beizubehalten ist, nicht aber bei Einzelfalldarbietung. Bei summarischer Darbietung arbeiten die Probanden mit visuell hoch strukturierten Medien wie einem Stapeldiagramm tendenziell effizienter als mit einem gering visuell strukturierten Medium wie einem Fliesstext. Bei Einzelfalldarbietung scheint der Faktor *visuelle Strukturierung des darstellenden Mediums* dagegen keinerlei Einfluss auf die Verarbeitungseffizienz zu nehmen. Angesichts der signifikanten Unterschiede zwischen den mittleren Zeiten für die Bearbeitung von einzelfallweise dargebotener Information in Vierfeldertafeln und Fliesstexten mag dieser Befund überraschen, er findet jedoch eine einfache Erklärung in der Tatsache, dass knapp die Hälfte aller Probanden, die mit einzelfallweise dargebotener Information gearbeitet haben, überhaupt keine regel-konformen

Lösungen hervorgebracht hat. Dieser Bodeneffekt in Verbindung mit der *Anzahl regel-konformer Lösungen* hat verhindert, dass Probanden, die mit Vierfeldertafeln und Einzelfalldarbietung gearbeitet haben, trotz kürzerer Bearbeitungszeiten höhere Effizienzmaße erzielten.

### 9.5.2 Diskussion

#### Einfluss des Faktors Darbietungsart

Alle aufgestellten Hypothesen können beibehalten werden. Der Faktor *Darbietungsart* wirkt sich auf alle abhängigen Variablen hochsignifikant aus.

*Anzahl regel-konformer Lösungen:* Die höhere mittlere Anzahl regel-konformer Lösungen bei summarischer Darbietung von Informationen über Ursache-Wirkungs-Zusammenhängen lässt sich darauf zurückführen, dass die Probanden im Gegensatz zur Einzelfalldarbietung bei der Aufgabenbearbeitung mental entlastet sind. Bei summarischer Informationsdarbietung werden das Arbeitsgedächtnis und andere kognitive Funktionen weniger beansprucht, da die Probanden keine Einzelfälle memorieren und ordnen müssen. Stattdessen können sich die Probanden ganz darauf konzentrieren, die bereits vorsortierten und zusammengefassten Informationen zueinander in Beziehung zu setzen und miteinander zu verrechnen. Die im Vergleich zur Einzelfalldarbietung geringere mentale Beanspruchung bei summarischer Darbietung führt auf kognitiver Ebene dazu, dass mehr Probanden dazu in die Lage versetzt werden, eine hohe Anzahl regel-konformer Lösungen zu produzieren. Bei einzelfallweiser Darbietung greifen dagegen mehr Probanden auf einfachere Strategien zurück (Kleinmuntz & Schkade, 1993; Shaklee & Mims, 1982) oder sie bringen eine geringere Anzahl von verzerrten oder fehlerhaften Kausalurteilen hervor (Fiedler, 2000; Shaklee & Mims, 1982).

*Bearbeitungszeit:* Die bei summarischer Darbietung hochsignifikant geringere Bearbeitungszeit kann im Sinne der Kosten-Nutzen Theorie (Kleinmuntz & Schkade, 1993) so interpretiert werden, dass der mentale Aufwand (die Kosten), den die Probanden bei der Aufgabenbearbeitung betreiben müssen, bei summarischer Darbietung geringer ist als bei einzelfallweiser Darbietung. Die Aufwandsreduzierung bei summarischer Darbietung zeigt sich sowohl bei der *time on task* als auch bei der *rating time*, und nicht nur bei der *time on task*, wie es in Verbindung mit der Hypothese H5b zusätzlich formuliert wurde. Es kann davon ausgegangen werden, dass sich der mediale Faktor *Darbietungsart* sowohl auf die peripheren Prozesse und Handlungsrouitinen in Verbindung mit der Datenaggregation und –zuordnung, als auch auf die zentralen Prozesse bei der Urteilsbildung auswirkt. Die bei summarischer Darbietung sehr kurzen Bearbeitungszeiten bei der eigentlichen Urteilsbildung

(*rating time*) kommen vermutlich dadurch zu Stande, dass sich die Probanden schon während der *time on task* mit der Urteilsbildung auseinandersetzen. Bei einzelfallweiser Darbietung beschäftigten sich die Probanden in der *time on task* dagegen mit dem Sammeln und Ordnen der neu aufgenommenen Information, was den Probanden, denen die Informationen in summarischer Form dargeboten wurden, erspart blieb.

*Verarbeitungseffizienz*: Per definitionem wirken sich die bei summarischer Darbietung höhere mittlere Anzahl regel-konformer Lösungen und die geringere mittlere Bearbeitungszeit auch auf die Verarbeitungseffizienz aus. Summarisch dargebotene Informationen können hoch signifikant effizienter verarbeitet werden als einzelfallweise dargebotene Informationen.

Zusammenfassend kann festgestellt werden, dass sich der Faktor *Darbietungsart* auf die gesamte Informationsverarbeitung bei der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen auswirkt. Je nach Darbietungsart fällt die mentale Beanspruchung bei der Beurteilung der Stärke von Kausalzusammenhängen unterschiedlich aus. Entsprechend moderiert der Faktor *Darbietungsart*, in welchem Umfang Probanden regel-konforme Lösungen hervorbringen und wie schnell die Prozesse bei der Beurteilung der Stärke von Ursache-Wirkungsprozessen ablaufen.

#### *Einfluss des Faktors Visuelle Strukturierung des darstellenden Mediums*

Die Einflüsse des Faktors *visuelle Strukturierung des darstellenden Mediums* waren heterogener und insgesamt weniger stark ausgeprägt, als die Einflüsse des Faktors *Darbietungsart*. Lediglich bei der Bearbeitungszeit konnte ein signifikanter Einfluss nachgewiesen werden. Für die abhängige Variable Anzahl regel-konformer Lösungen ergab sich ebenso wie für die Verarbeitungseffizienz ein Trend.

*Anzahl regel-konformer Lösungen*: Der wesentliche Grund für den mäßigen Einfluss des Faktors *visuelle Strukturierung* auf die Anzahl regel-konformer Lösungen liegt erwartungsgemäß darin, dass eine hohe visuelle Strukturierung den Probanden nur dann das Ziehen von regel-konformen oder komplexen Schlussfolgerungen über die zu beurteilenden Ursache-Wirkungs-Zusammenhänge erleichtert, wenn die mentale Beanspruchung durch eine summarische Darbietung der dafür relevanten Informationen reduziert ist. Ist die mentale Beanspruchung durch eine einzelfallweise Darbietung dagegen erhöht, so macht es keinen Unterschied, ob die visuelle Strukturierung des darstellenden Mediums gering oder hoch ist. Dies belegen auch die durchgeführten Einzelkontraste, die zeigen dass die Probanden unter der Bedingung *SD sum* eine signifikant höhere Anzahl regel-konformer Lösungen als unter der Bedingung *FT sum* erzielten, während der Unterschied zwischen den Bedingungen *VT tbt*

und *FT tbt* wegen der hohen mentalen Beanspruchung in diesen beiden Einzelfallbedingungen klar nicht signifikant wurde.

Der signifikante Unterschied zwischen den beiden Bedingungen *SD sum* und *FT sum* zeigt, dass - im Unterschied zu Experiment 1, wo sich tabellarische Vierfeldertafel und Stapeldiagramm nicht signifikant voneinander unterscheiden - auch die Präsentationsform einen wesentlichen Einfluss auf die Anzahl regel-konformer Lösungen haben kann, wenn sich die beiden Präsentationsformen nur hinlänglich hinsichtlich ihrer Informationsanordnung und ihres Angebots an Strukturierungshilfen unterscheiden. Durch die Informationsanordnung im Stapeldiagramm kann für die Probanden ein regel-konformes Zueinander-in-Beziehungsetzen von Erfolgs- und Basisrate salienter gemacht, unmittelbar nahe gelegt und partiell auf die Wahrnehmungsebene verlagert werden (Wickens & Carswell, 1995). Fliesstexte hingegen ermöglichen weder perzeptuelle Inferenzen (Larkin & Simon, 1987), noch kann die Informationsverarbeitung auf die Wahrnehmungsebene verlagert werden (Wickens & Carswell, 1995). In Verbindung mit der Kosten-Nutzen Theorie (Kleinmuntz & Schkade, 1993) besteht die Möglichkeit, dass die Probanden unter der Fliesstext-Bedingung eine höhere Aufgabenschwierigkeit antizipieren und deshalb auf einfachere Strategien zurückgreifen und entsprechend weniger regel-konforme Lösungen hervor bringen.

Bei Einzelfalldarbietung hat der Faktor *visuelle Strukturierung* wie erwähnt keine Auswirkungen auf die Anzahl regel-konformer Lösungen, weil Arbeitsgedächtnis und Aufmerksamkeit durch das Memorieren und Gruppieren der Einzelinformationen absorbiert sind. Dennoch hat eine Minderheit von Probanden auch bei einzelfallweiser Darbietung regel-konforme Lösungen produziert. Wir vermuten, dass diese Probanden entweder schon vor der Aufgabenbearbeitung gewusst haben, wie sie die Einzelinformationen aufaddieren, gruppieren und zueinander in Beziehung setzen müssen, oder dass sie aufgrund besonderer individueller Voraussetzungen dazu in der Lage waren, trotz ungünstiger medialer Rahmenbedingungen regel-konforme Lösungen zu generieren.

*Bearbeitungszeit:* Auf die Bearbeitungszeit wirkt sich der Faktor *visuelle Strukturierung des darstellenden Mediums* anders aus als auf die Anzahl regel-konformer Lösungen. Hier haben wir festgestellt, dass eine hohe visuelle Strukturierung nur bei Einzelfalldarbietung - und nicht bei summarischer Darbietung von kausalen Informationen - zu signifikant schnelleren Bearbeitungszeiten führt. Dies erklärt auch das von uns unerwartete Signifikantwerden der Interaktion *Darbietungsart x visuelle Strukturierung*.

Wir vermuten, dass bei Einzelfalldarbietung die übersichtliche Anordnung der Informationen in einer Vierfeldertafel mit Positionskreuz die Wahrnehmung und Aufnahme einzelner

Informationen im Vergleich zu einem Fliesstext mit Einzelfalldarbietung deutlich vereinfacht und beschleunigt. Denkbar ist auch, dass das immer wieder neue Erscheinen des Positionskreuzes in der Grafik eine immer wieder neue Orientierungsreaktion auslöst und die Aufmerksamkeit im Gegensatz zum Fliesstext mit Einzelfalldarbietung zusätzlich fixiert und strukturiert wird. Trotz beschleunigter Informationsaufnahme werden die Probanden, die mit einer Vierfeldertafel und Positionskreuz arbeiten, aber aufgrund der insgesamt hohen mentalen Beanspruchung und der im Unterschied zur Bedingung *SD sum* seriellen Abfolge von Einzelinformationen nicht darin unterstützt, wie sie die einzelnen Informationen bei der Beurteilung der Wirksamkeiten gruppieren und zueinander regel-konform in Beziehung setzen sollen.

Durch die Aufschlüsselung der Gesamtbearbeitungszeit in die Teile *time on task* und *rating time* konnte gezeigt werden, dass der Faktor *visuelle Strukturierung* im Gegensatz zum Faktor *Darbietungsart* keinen Einfluss auf die *rating time* hat, sondern nur auf die *time on task*. Dies kann als eine Bestätigung der Annahme verstanden werden, dass der Faktor *visuelle Strukturierung* die Dauer von Prozessen bei der Datenaggregation und –zuordnung, nicht aber die Dauer von Prozessen bei der eigentlichen Urteilsbildung beeinflusst.

Wenn die Annahme zutreffend ist, dass der Einfluss des Faktors *visuelle Strukturierung* auf die Bearbeitungszeit vor allem periphere Prozesse und Handlungsroutinen betrifft, dann erklärt dies auch das für uns überraschende Ergebnis, dass bei summarischer Darbietung keine signifikanten Zeitunterschiede in Abhängigkeit von der *visuellen Strukturierung* zu beobachten waren. Unter den Bedingungen *SD sum* und *FT sum* werden den Probanden im Gegensatz zu den beiden Einzelfallbedingungen *VT tbt* und *FT tbt* das zeitaufwendige Aggregieren und Zuordnen von Daten abgenommen. Das Ergebnis, dass Fliesstexte mit summarischer Informationsdarbietung sogar geringfügig schneller verarbeitet werden als Stapeldiagramme haben wir auf eine möglicherweise größere Vertrautheit der Probanden im Umgang mit Texten zurückgeführt. Da sich aber die Bedingungen *SD sum* und *FT sum* signifikant hinsichtlich der Anzahl regel-konformer Lösungen unterscheiden, ist es wahrscheinlich, dass die Probanden in den beiden Bedingungen unterschiedlich komplexe Strategien ausgewählt haben. Möglicherweise haben eine Reihe von Probanden unter der Bedingung *FT sum* darauf verzichtet, komplexe regel-konforme Berechnungen vorzunehmen und stattdessen auf einfachere und weniger zeitaufwendige Schätzungen zurückgegriffen.

*Verarbeitungseffizienz:* Die Ergebnisse für die Verarbeitungseffizienz lassen sich erneut aus den Ergebnissen für die Anzahl regel-konformer Lösungen und Bearbeitungszeit ableiten. Dabei zeigte es sich, dass visuell hoch strukturierte und übersichtlich gestaltete

Präsentationsformen nur dann effizienter verarbeitet werden, wenn sie summarisch dargeboten werden und die mentale Beanspruchung gering ist. In den beiden Einzelfallbedingungen haben die Probanden jeweils so wenig regel-konforme Lösungen hervor gebracht, dass die Unterschiede im Hinblick auf die Bearbeitungszeiten bei der Ermittlung der Verarbeitungseffizienz nicht ins Gewicht gefallen sind.

#### Einfluss der Kovariate numerisches Denken

Die Kovariate *numerisches Denken* hatte erwartungsgemäß einen hochsignifikanten Einfluss auf die Anzahl regel-konformer Lösungen und die Verarbeitungseffizienz. Individuelle Unterschiede in der Befähigung zum numerischen Denken sind also hoch bedeutsam, wenn es darum geht, in welchem Ausmaß die Probanden regel-konforme Lösungen generieren und wie effizient sie die dargebotenen Informationen verarbeiten können. Ebenso entsprach es den Erwartungen, dass die Bearbeitungszeit nicht von individuellen Unterschieden im numerischen Denken beeinflusst wurde. Hier spielen vermutlich andere Variablen wie zum Beispiel individuelle Arbeitsstile und die Motivation eine wichtigere Rolle.

Der bislang noch nicht nachgewiesene signifikante Einfluss der Kovariate *numerisches Denken* auf die Anzahl regel-konformer Lösungen bestätigt die Annahme von Shaklee & Hall (1983), dass Unterschiede im mathematischen Denken die Häufigkeit des Einsatzes und die Genauigkeit regel-basierter Strategien beeinflussen. Darüber hinaus zeigt der hochsignifikante Einfluss der Kovariate *numerisches Denken* auch, dass die Kritik von Perales & Shanks (2003) an der geringen Forschungsaktivität zum Einfluss individueller Unterschiede auf die Beurteilung der Stärke kausaler Zusammenhänge absolut berechtigt ist. Unser Experiment 2 erbrachte den Nachweis, dass hohe Werte im numerischen Denken mit einer höheren Anzahl regel-konformer Lösungen einhergehen.



## 9.6 Häufigkeitsanalysen zum Einfluss der Versuchsbedingungen auf die Strategiewahl

### 9.6.1 Ergebnisse

Wie in Experiment 1 wurde für jeden Probanden anhand der abgegebenen Wirksamkeitseinschätzungen, sowie der schriftlichen und mündlichen Nachbefragung ermittelt, welche Strategie(n) von ihnen eingesetzt wurden. Die Kategorien für Strategietypen, denen die einzelnen Probanden zugeordnet wurden, sind dieselben wie in Experiment 1.

1. Strategie unklar
2. Wechsler unklar – ab-Regel
3. ab-Regel oder eine andere Heuristik
4. Wechsler zu einer regel-basierten Strategie
5. durchgängig regel-basierte Strategie

Tabelle 13 zeigt, wie sich die 80 Probanden insgesamt und in Abhängigkeit von den vier Experimentalbedingungen auf die sechs Kategorien verteilen:

**Tabelle 13:** Experiment 2: Anzahl der Probanden in Abhängigkeit von Strategietyp und Versuchsbedingung (in Klammern Prozentangaben). Zelhäufigkeiten > 5 sind **fett** markiert.

	Strategietyp					Gesamt
	unklar	Wechsler ab-Regel	ab-Regel	Wechsler regel-basiert	regel-basiert	
<b>Versuchsbedingung</b>						
SD sum	3	2	1	<b>7</b>	<b>7</b>	20
FT sum	3	<b>9</b>	1	4	3	20
VT tbt	5	5	4	5	1	20
FT tbt	4	5	2	<b>8</b>	1	20
<b>Gesamt</b>	15 (18.75 %)	21 (26.25 %)	8 (10 %)	24 (30 %)	12 (15 %)	80 (100 %)

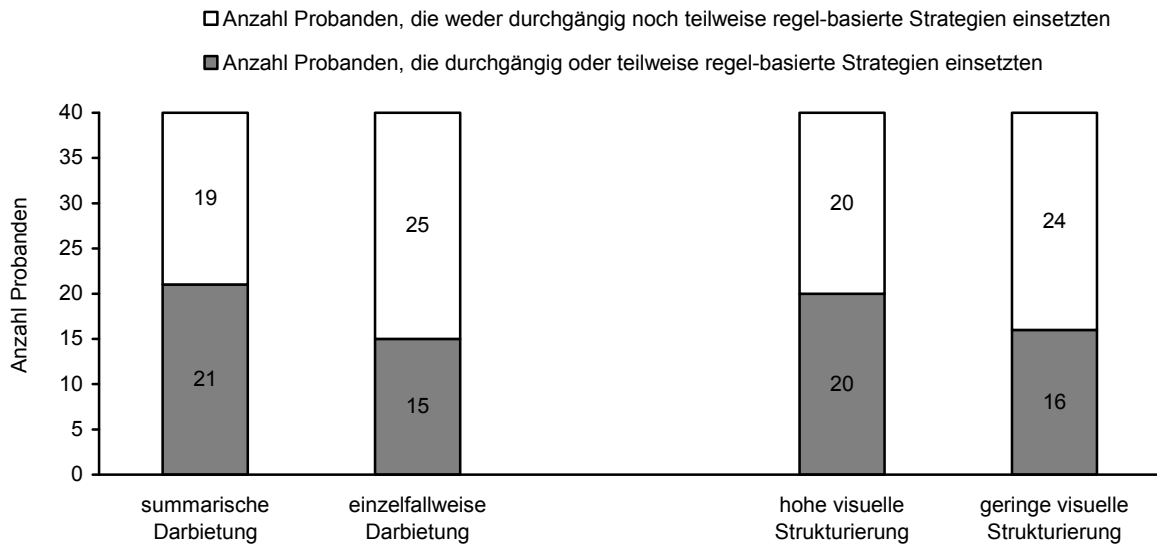
Im Vergleich zu Experiment 1 ist der Anteil derjenigen Probanden, die ausschließlich regel-basierte Strategien eingesetzt haben, mit 15 % sehr gering, werden noch diejenigen Probanden hinzugezählt, die während der Versuchsdurchführung zu einer regel-basierten Strategie übergewechselt haben, so haben insgesamt 45 % aller Probanden ausschließlich oder teilweise regel-basierte Strategien eingesetzt. Die in Experiment 1 beobachteten Deckeneffekte treten in Experiment 2 wie erwartet nicht mehr auf, was wesentlich mit der Bedingungsmanipulation zusammenhängen dürfte. Auffallend und den Erwartungen entsprechend an Tabelle 13 ist, dass der Anteil derjenigen, die ausschließlich regel-basierte Strategien eingesetzt haben, bei summarischer Darbietung höher ist, als bei Einzelfalldarbietung (10 : 2). Etwas überraschend

ist vielleicht, dass der Anteil derjenigen, die durchgängig unklare Strategien eingesetzt haben, bei Einzelfalldarbietung nur unwesentlich höher ist als bei summarischer Darbietung (9 : 6). Offensichtlich versuchen auch bei Einzelfalldarbietung über 75 % der Probanden von mehr oder weniger regel-basierten Strategien Gebrauch zu machen.

Im Hinblick auf die Strategiewahl nimmt die Bedingung *SD sum* eine Sonderstellung ein, da unter dieser Bedingung der Einsatz komplexerer Strategien wie der  $\Delta P$ -Regel oder der Berechnung der kausalen Stärke nach der Power PC-Theorie überwiegt. 70 % aller Probanden, die mit Stapeldiagrammen arbeiten, setzten solche Strategien durchgängig oder teilweise ein. Bei Einzelfalldarbietung oder unter der Bedingung *FT sum* bevorzugten die Probanden hingegen eher den Wechsel zu oder den durchgängigen Einsatz von einfachen Heuristiken wie der ab-Regel. Während unter der Bedingung *SD sum* 35 % der Probanden durchgängig regel-basierte Strategien einsetzen, so beträgt dieser Anteil über die drei anderen Bedingungen gemittelt nur 8,3 %.

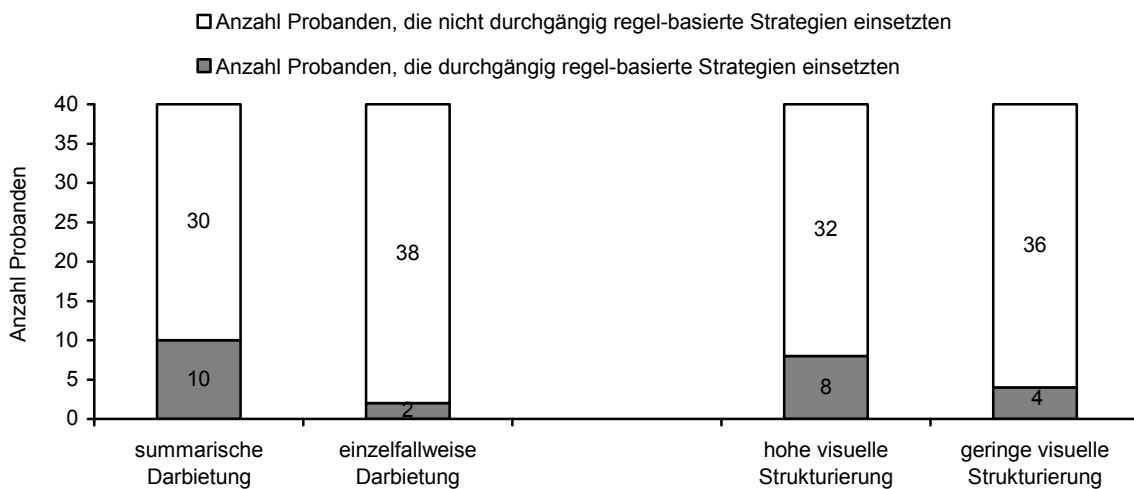
Weitere Häufigkeitsanalysen ergaben, dass 70 % aller Probanden ihre Strategie bei der Bearbeitung der verschiedenen Aufgaben verändert haben. Die verschiedenen Wechsler-Typen und ihre Häufigkeit sind in Tabelle A3 im Anhang A2.4 eingetragen und geben anders als Tabelle 13 einen Einblick in die tatsächliche Vielfalt der eingesetzten Strategien.

Zur Überprüfung der beiden Hypothesen H5d und H6d wurden nun alle Probanden binär hinsichtlich des Strategietyps verteilt. Die beiden Hypothesen besagen, dass die Anzahl der Probanden, die regel-basierte Strategien einsetzen, bei summarischer Darbietung (H5d) und hoher visueller Strukturierung des darstellenden Mediums (H6d) höher ist als bei Einfalldarbietung und geringer visueller Strukturierung. Die erste Gruppe setzte sich aus den 36 Probanden zusammen, die teilweise oder durchgängig regel-basierte Strategien eingesetzt haben. Die zweite Gruppe bestand aus den restlichen 44 Probanden, die keinen Gebrauch von regel-basierten Strategien gemacht haben. Im Anschluss daran wurde mit der Hilfe eines  $\chi^2$ -Tests und des exakten Tests nach Fisher untersucht, wie sich diese beiden Strategietypen jeweils in Abhängigkeit von der *Darbietungsart* und der *visuellen Strukturierung* verteilten. Abbildung 30 zeigt, wie sich die beiden Strategietypen auf die beiden zweistufigen Faktoren *Darbietungsart* und *visuelle Strukturierung* verteilten:



**Abbildung 30:** Experiment 2: Verteilung von Probanden, die regel-basierte oder keine regel-basierten Strategien eingesetzt haben, in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*.

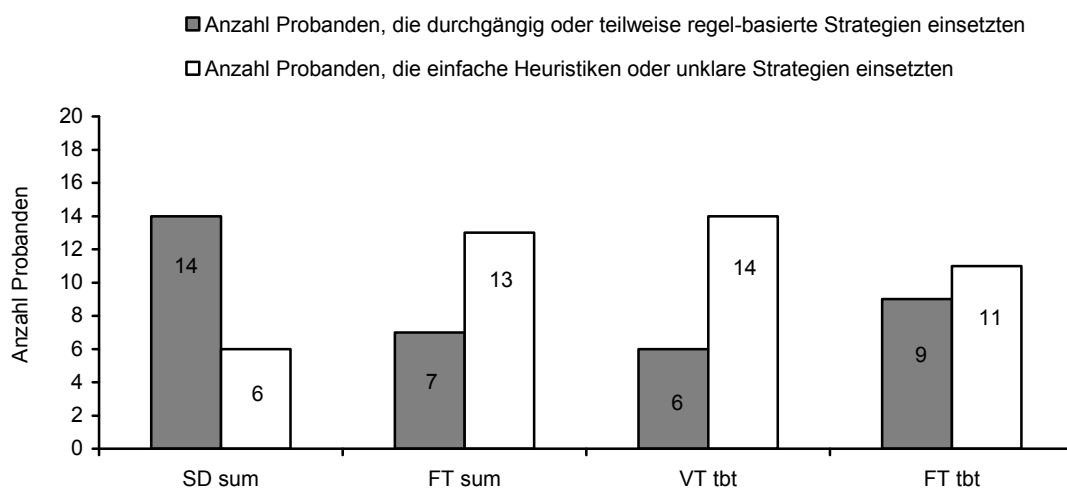
Auch wenn die Häufigkeitsunterschiede jeweils in die vorhergesagte Richtung gingen, so verteilten sich sowohl *Darbietungsart* und *Strategietyp* ( $\chi^2 = 1.818$ ,  $p = .178$ ; einseitiger P-Wert nach Fisher's exaktem Test = .131), als auch *visuelle Strukturierung des darstellenden Mediums* und *Strategietyp* unabhängig voneinander ( $\chi^2 = .808$ ,  $p_{\text{asymptotische Signifikanz}} = .369$ ;  $P_{\text{Fisher, einseitig}} = .250$ ), die Zusammenhänge wurden also nicht signifikant. Andere Ergebnisse stellen sich ein, wenn die 12 Probanden, die ausschließlich durchgängig von regel-basierten Strategien Gebrauch machten, den 68 Probanden gegenübergestellt werden, die nicht durchgängig regel-basierte Strategien einsetzten. Sie werden in Abbildung 31 gezeigt.



**Abbildung 31:** Experiment 2: Verteilung von Probanden, die entweder durchgängig regel-basierte oder nicht durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*.

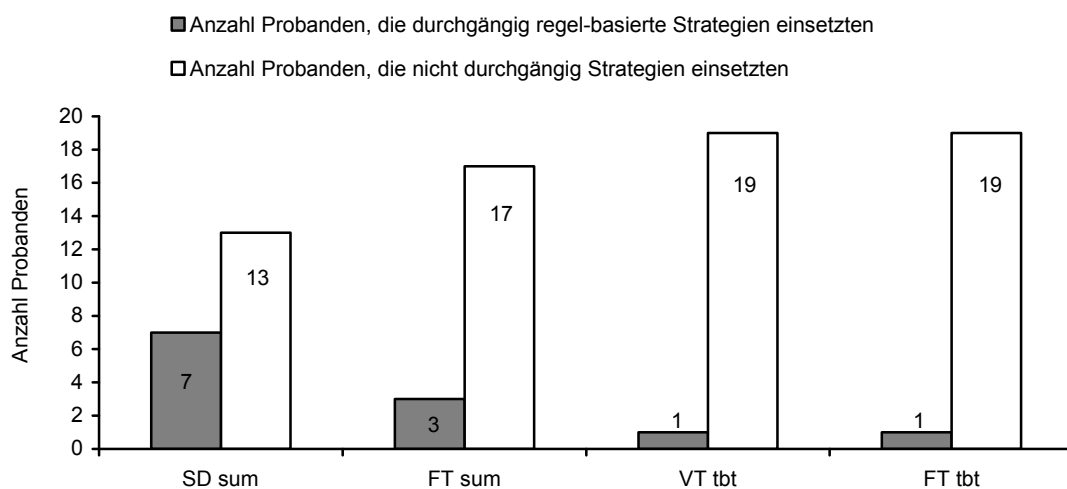
Ein  $\chi^2$ -Test und ein exakter Test nach Fischer führten zu dem signifikanten Ergebnis, dass bei summarischer Darbietung die Anzahl von Probanden, die durchgängig von einer regelbasierten Strategie Gebrauch machten, höher war als bei Einzelfalldarbietung ( $\chi^2 = 6.275$ ,  $p_{\text{asymptotische Signifikanz}} = .012$ ;  $p_{\text{Fisher, einseitig}} = .013$ ). Dagegen hatte die Variable *visuelle Strukturierung des darstellenden Mediums* weiterhin keinen signifikanten Einfluss auf die Häufigkeitsverteilung der beiden Probandengruppen ( $\chi^2 = 1.569$ ,  $p_{\text{asymptotische Signifikanz}} = .210$ ;  $p_{\text{Fisher, einseitig}} = .174$ ). Während die Hypothese H5d zumindest für Probanden, die durchgängig von regelbasierten Strategien Gebrauch machten, beibehalten werden kann, so konnte die Hypothese H6d weder für Probanden, die durchgängig regelbasierte Strategien einsetzten, noch für solche, die dies zumindest teilweise taten, bestätigt werden.

Die Auswertungsergebnisse in Tabelle 9 ergab ebenso wie die Abbildungen 25 und 27 einen Hinweis darauf, dass die Probanden unter der Bedingung *SD sum* eine höhere Anzahl regelkonformer Lösungen hervorbrachten als die Probanden in den anderen drei Bedingungen. Dass die Bedingung *SD sum* auch im Hinblick auf die Anzahl der Probanden, die regelbasierte Strategien einsetzten, einen Sonderstatus im Vergleich zu den anderen drei Bedingungen einnimmt, veranschaulicht sehr gut Abbildung 32. In ihr werden die Probanden nach Strategietyp entweder in eine Gruppe von Probanden, die teilweise oder durchgängig regelbasierte Strategien angewendet haben, oder in eine Gruppe, die nur von einfachen Heuristiken wie der ab-Regel oder unklaren Strategien Gebrauch gemacht hat, unterteilt.



**Abbildung 32:** Experiment 2: Verteilung von Probanden, die durchgängig oder teilweise regelbasierte Strategien eingesetzt haben und von Probanden, die einfache Heuristiken oder unklare Strategien eingesetzt haben in Abhängigkeit von der Versuchsbedingung.

Ein  $\chi^2$ - Test erbrachte einen Trend, dass die Verteilung von regel-basiert vorgehenden Strategieanwendern und nicht regel-basiert vorgehenden Strategieanwendern in Abhängigkeit von der Versuchsbedingung nicht zufällig ist ( $\chi^2 = 7,677$ ;  $p = .053$ ; Phi-Wert = .310). Noch deutlicher zeigt sich die Sonderstellung der Bedingung *SD sum* wenn in Abhängigkeit von der Versuchsbedingung die Anzahl der Probanden, die durchgängig regel-basierte Strategien einsetzen, der Anzahl von Probanden, die nicht durchgängig regel-basierte Strategien einsetzen, gegenübergestellt wird. Die sich hieraus ergebende Häufigkeitsverteilung veranschaulicht Abbildung 33.



**Abbildung 33:** Experiment 2: Verteilung von Probanden, die entweder durchgängig regel-basierte oder nicht durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von der Versuchsbedingung.

Der auch zu dieser Verteilung durchgeführte  $\chi^2$ - Test führte zu dem signifikanten Ergebnis, dass die Verteilung von durchgängig regel-basiert vorgehenden Strategieanwendern und nicht durchgängig regel-basiert vorgehenden Strategieanwendern in Abhängigkeit von der Versuchsbedingung nicht zufällig ist ( $\chi^2 = 9,412$ ;  $p = .024$ ; Phi-Wert = .343). Allerdings ist dieses signifikante Ergebnis mit Vorsicht zu interpretieren, da der Anteil der erwarteten Zellhäufigkeiten, die kleiner als 5 sind, 50 % erreicht.

Wie schon in Experiment 1, wurden auch die Probanden aus Experiment 2, die durchgängig regel-basierte Strategien eingesetzt haben, weiter ausdifferenziert und daraufhin untersucht, ob sie von der  $\Delta P$ -Regel, der Power PC-Regel oder einer anderen regel-basierten Strategie Gebrauch gemacht haben. Tabelle 14 zeigt, wie viele Probanden in welcher Versuchsbedingung welche Strategie angewendet haben. Im Einzelnen wurden für Tabelle 14 die folgenden Kategorien, denen die Probanden zugeordnet wurden, gebildet:

1. Delta P-Regel, kurz " $\Delta$ P-Regel"
2. Power PC-Theorie, kurz "Power PC"
3. Wechsler Delta P-Regel - Power PC Theorie, kurz " $\Delta$ P & PPC"

**Tabelle 14:** Experiment 2: Anzahl der Probanden, die durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von der Art der regel-basierten Strategie und der Versuchsbedingung (in Klammern Prozentangaben).

	Art der regel-basierten Strategie			Gesamt
	$\Delta$ P-Regel	Power PC	$\Delta$ P & PPC	
<b>Versuchsbedingung</b>				
SD sum	3	3	1	7 (58.3 %)
FT sum	0	2	1	3 (25.0 %)
VT tbt	0	1	0	1 ( 8.3 %)
FT tbt	0	1	0	1 ( 8.3 %)
<b>Gesamt</b>	3 (25 %)	7 (58,3 %)	2 (16,7 %)	12 (100 %)

Erwartungsgemäß setzen unter der Bedingung *SD sum* mehr Probanden durchgängig eine regel-basierte Strategie ein. Etwas überraschend machen über alle vier Bedingungen hinweg mehr Probanden von der Power PC-Theorie als von der  $\Delta$ P-Regel Gebrauch (7 : 3). In der zahlenmäßig weit größeren Gruppe der Wechsler (siehe Tabelle A3 im Anhang A2.4) hatten 15 Probanden von unklaren Strategien oder einfachen Heuristiken zur  $\Delta$ P-Regel (ohne die Kategorie "25"), 8 zur Power PC Theorie und 2 zu beiden regel-basierten Strategien gewechselt.

Zusammengefasst erbrachten die Häufigkeitsanalysen zu Experiment 2 folgende Ergebnisse: Ungefähr 45 % der Probanden setzten durchgehend oder teilweise regel-basierte Strategien ein, der Anteil derjenigen, die durchgängig regel-basierte Strategien eingesetzt haben, beläuft sich auf 15 %. Bei summarischer Darbietung ist die Anzahl derjenigen, die ausschließlich regel-basierte Strategien eingesetzt haben, signifikant höher als bei Einzelfalldarbietung. Bei hoher visueller Strukturierung des darstellenden Mediums ist die Anzahl der Probanden, die regel-basierte Strategien einsetzen zwar ebenfalls höher als bei geringer visueller Strukturierung, doch wurde dieser Unterschied nicht signifikant. Während unter der Bedingung *SD sum* die Mehrheit der Probanden durchgängig oder teilweise auf regel-basierte Strategien einsetzt (70 %), machen die Probanden bei Einzelfalldarbietung und auch unter der Bedingung *FT sum* überwiegend von unklaren Strategien oder von einfachen Heuristiken wie der ab-Regel Gebrauch (*FT sum*: 65 %; *VT tbt* 70 %; *FT tbt* 55 %).

### 9.6.2 Diskussion

Insgesamt bestätigen die Häufigkeitsanalysen die Ergebnisse der Kovarianzanalyse mit der abhängigen Variable *Anzahl regel-konformer Lösungen*. Unter der Bedingung *SD sum* machen mehr Probanden von regel-basierten Strategien Gebrauch als unter den anderen drei Bedingungen. Wir sehen darin eine Bestätigung der Annahme, dass bei summarischer Darbietungsform und der hohen visuellen Strukturierung des Mediums Stapeldiagramm die mentale Beanspruchung geringer ausfällt als bei den beiden einzelfallweisen und auch bei der summarischen Darbietungsform im Fliesstext. Darüber hinaus erhalten die Probanden unter der Bedingung *SD sum* durch die Informationsanordnung mehr Unterstützung dabei, die übersichtlich und strukturiert zusammengefassten Informationen regel-konform zueinander in Beziehung zu setzen und zu verrechnen.

Das Ergebnis, dass unter der Bedingung *SD sum* die meisten Probanden hoch komplexe regel-basierte Strategien einsetzten, während in den drei anderen Bedingungen die meisten Probanden von unklaren Strategien oder einfachen Heuristiken Gebrauch machten, kann nach unserer Auffassung als eine Bestätigung der Kosten-Nutzen-Theorie von Kleinmuntz & Schkade (1993) angesehen werden. Danach würde die Mehrzahl der Probanden unter den Bedingungen *FT sum*, *VT tbt* und *FT tbt* den Einsatz einer komplexen regel-basierten Strategie als zu aufwendig und kostenintensiv ansehen und deshalb auf einfachere und ungenauere Strategien zurückgreifen.

Im Vergleich zu Experiment 1 wurden weniger regel-basierte Strategien eingesetzt und es ergaben sich keine Deckeneffekte. In Verbindung mit einem durchgängigen Einsatz regel-basierter Strategien kam es in den beiden Einzelfallbedingungen sogar eher zu Bodeneffekten. Das spricht dafür, dass die mentale Beanspruchung der Probanden in Experiment 1 insgesamt höher war als in Experiment 2. Die in Experiment 1 gefundene Tendenz, dass der Anteil der Probanden, die eine regel-basierte Strategie gewählt haben, mit zunehmender mentaler Beanspruchung abnimmt, zeigt sich auch in Experiment 2.

Eine genauere Analyse der Probanden (s.a. Tabelle 14 und Tabelle A3 im Anhang A2.4), die teilweise oder durchgängig regel-basierte Strategien gewählt haben, kam zu dem Ergebnis, dass wie in Experiment 1 die meisten dieser Probanden von der  $\Delta P$ -Regel Gebrauch gemacht haben. Über alle Versuchsbedingungen hinweg haben von allen Probanden, die regel-basiert vorgegangen sind, 50 % die  $\Delta P$ -Regel eingesetzt und etwa 40 % die Stärke des Kausalzusammenhanges nach der Power PC-Theorie ermittelt. In Experiment 1 waren unter den beiden Häufigkeitsbedingungen 33,3 % aller regel-basiert vorgehenden Probanden

entsprechend der Power PC-Theorie und 54,2 % nach der  $\Delta P$ -Regel vorgegangen. In weiteren Teilanalysen stellte es sich heraus, dass eine zumindest teilweise den Vorhersagen der  $\Delta P$ -Regel entsprechende Vorgehensweise eher von Probanden eingesetzt wurde, die unter Einzelfallbedingungen (Verhältnis Pbn  $\Delta P$  zu Pbn PPC 10 : 5) arbeiten mussten oder die die Strategie gewechselt haben (Verhältnis Wechsler  $\Delta P$  zu Wechsler PPC 15 : 8). Von einer den Vorhersagen der Power PC-Theorie entsprechenden Strategie machten hingegen eher solche Probanden Gebrauch, denen die Informationen über Ursache-Wirkungs-Zusammenhänge summarisch dargeboten wurden (Verhältnis Pbn  $\Delta P$  zu Pbn PPC 8 : 10) oder die die Strategie nicht gewechselt haben (Verhältnis Pbn durchgängig  $\Delta P$  zu Pbn durchgängig PPC 3 : 7). Auch wenn sich solche Teilergebnisse nur schwer interpretieren lassen, so legen sie doch den Schluss nahe, dass medienbezogene Faktoren nicht nur einen Einfluss darauf haben können, in welchem Ausmaß regel-basiert vorgegangen wird, sondern darüber hinaus auch noch, von welcher der verschiedenen regel-basierten Strategien ( $\Delta P$  oder PPC) eine Mehrheit der regel-basiert vorgehenden Probanden Gebrauch macht.

In Experiment 2 haben 44 der insgesamt 80 Probanden (55 %) weder durchgängig noch teilweise Gebrauch von regel-basierten Strategien Gebrauch gemacht. Dieser Anteil ist höher als in Experiment 1, wo nur 21 von 80 Probanden (ca. 26 %) durchgängig von keiner regel-basierten Strategie gebraucht gemacht haben. Auch dies spricht eine insgesamt höhere mentale Beanspruchung der Probanden in Experiment 2. Von den 44 nicht regel-basiert vorgehenden Probanden haben 29 Probanden (ca. 36 %) durchgängig oder teilweise von der einfacheren ab-Regel Gebrauch gemacht, die restlichen 15 Probanden (ca. 19 %) haben Vorgehensweisen gewählt, die sich auch nach der Nachbefragung als unklar herausgestellt haben und meist auf intuitiven Schätzungen beruhen. Fehlerhaft eingesetzte regel-basierte Strategien konnten in Experiment 2 nicht ausgemacht werden, doch kann ihr Vorkommen auch nicht ausgeschlossen werden. In Verbindung mit den Ergebnissen aus Experiment 1 spricht der geringe Anteil an fehlerhaft eingesetzten regel-basierten Strategien eher dafür, dass die Probanden aufgrund der subjektiv wahrgenommenen Aufgabenkomplexität auf einfachere und ungenauere Strategien zurückgegriffen haben (Kleinmuntz & Schkade, 1993; Shaklee & Mims, 1982) und weniger dafür, dass sie Fehler bei der Anwendung regel-basierter Strategien begehen.

Die Häufigkeitsanalysen aus Experiment 2 bestätigen auch das in Experiment 1 gefundene Ergebnis, dass die Probanden weder über alle Versuchsbedingungen hinweg noch innerhalb einer einzigen Versuchsbedingung einheitlich vorgegangen sind. Dies spricht erneut gegen die Annahme, dass es einen einheitlichen Verarbeitungsmechanismus bei der Beurteilung der



Stärke von Kausalzusammenhängen gibt, und dafür, dass individuellen Unterschieden dabei eine wichtige Rolle zukommt. Das Vorkommen eines assoziativen Lernmechanismus wurde von uns nicht eigens untersucht, kann aber für die beiden Einzelfallbedingungen nicht ausgeschlossen werden. Allerdings unterscheiden sich die beiden Einzelfallbedingungen und die Bedingung *FT sum* hinsichtlich der Strategiewahl kaum voneinander. Da assoziative Lernmechanismen einer Verarbeitung von summarisch dargebotenen Informationen nicht zugrunde gelegt werden können (Lober & Shanks, 2000; Shanks, 1991), könnte die ähnliche Strategiewahl in den beiden Einzelfallbedingungen und die Bedingung *FT sum* nicht auf einen gemeinsamen assoziativen Lernmechanismus zurückgeführt werden.

Mehr als die Hälfte (45 von 80) aller Probanden hat ihre Strategie im Verlauf von Experiment 2 gewechselt. Wenn man noch diejenigen Probanden hinzunimmt, die von der  $\Delta P$ -Regel zu einer den Vorhersagen der Power PC Theorie entsprechenden Vorgehensweise gewechselt sind, dann betrug der Wechsler-Anteil sogar knapp 60 % (47 von 80). Viele Probanden haben sogar mehrmals die Strategie gewechselt. Am wenigsten wurde unter der Bedingung *SD sum* gewechselt. Wenn unter der Bedingung *SD sum* gewechselt wurde, dann wurde vor allem zu einer oder mehreren komplexen regel-basierten Strategie/n gewechselt, während die Probanden in den anderen drei Bedingungen häufiger zu einfacheren Heuristiken wie der ab-Regel und später zu regel-basierten Strategien wechselten. Der im Vergleich zu Experiment 1 höhere Wechsler-Anteil in Experiment 2 zeigt ebenso wie das Ergebnis, dass unter der Bedingung *SD sum* am wenigsten gewechselt wurde, dass das Vorkommen von Strategiewechseln mit zunehmender mentaler Beanspruchung und bei einer Informationsanordnung, die kaum Strukturierungsangebote enthält, steigt. Offensichtlich steigt dann bei einem Anwachsen der mentalen Beanspruchung und bei einer Abnahme von Strukturierungsangeboten die Unsicherheit der Probanden, welche sich in einem höheren Wechsler-Anteil zeigt. Die Probanden versuchen dann, ihre Strategien entweder zu optimieren oder individuell an die subjektiv wahrgenommene Aufgabenkomplexität anzupassen (Kleinmuntz & Schkade 1993).

In den beiden Einzelfallbedingungen haben insgesamt nur 2 von 40 Probanden durchgängig von regel-basierten Strategien Gebrauch gemacht. Der Anteil derer, die während des Experiments in den beiden Einzelfallbedingungen zu einer regel-basierten Strategie gewechselt haben, betrug immerhin 32,5 % (13 von 40) und war damit verhältnismäßig hoch. Dieser relativ hohe Anteil an Probanden, die zu regel-basierten Strategien übergewechselt sind, hängt vielleicht damit zusammen, dass die Probanden im Vergleich zu anderen Untersuchungen (Anderson & Sheu, 1995, Ex 1; Arkes & Harkness, 1983, Ex 1, 2, 3, 6, 7;

Kao & Wasserman, 1993, Ex 2) wegen der langen Darbietungszeit von 10 Sekunden viel Zeit hatten, die einzelfallweise dargebotenen Informationen über Ursache-Wirkungs-Zusammenhänge zu verarbeiten und sich dafür eine geeignete Strategie zu überlegen. Bei kürzeren Intervallen zwischen zwei Einzelfalldarbietungen wäre mit einem weiteren Anstieg der mentalen Beanspruchung zu rechnen, so dass sich der Anteil von Probanden, der zumindest teilweise von regel-basierten Strategien Gebrauch gemacht hätte, vermutlich noch weiter reduziert hätte.

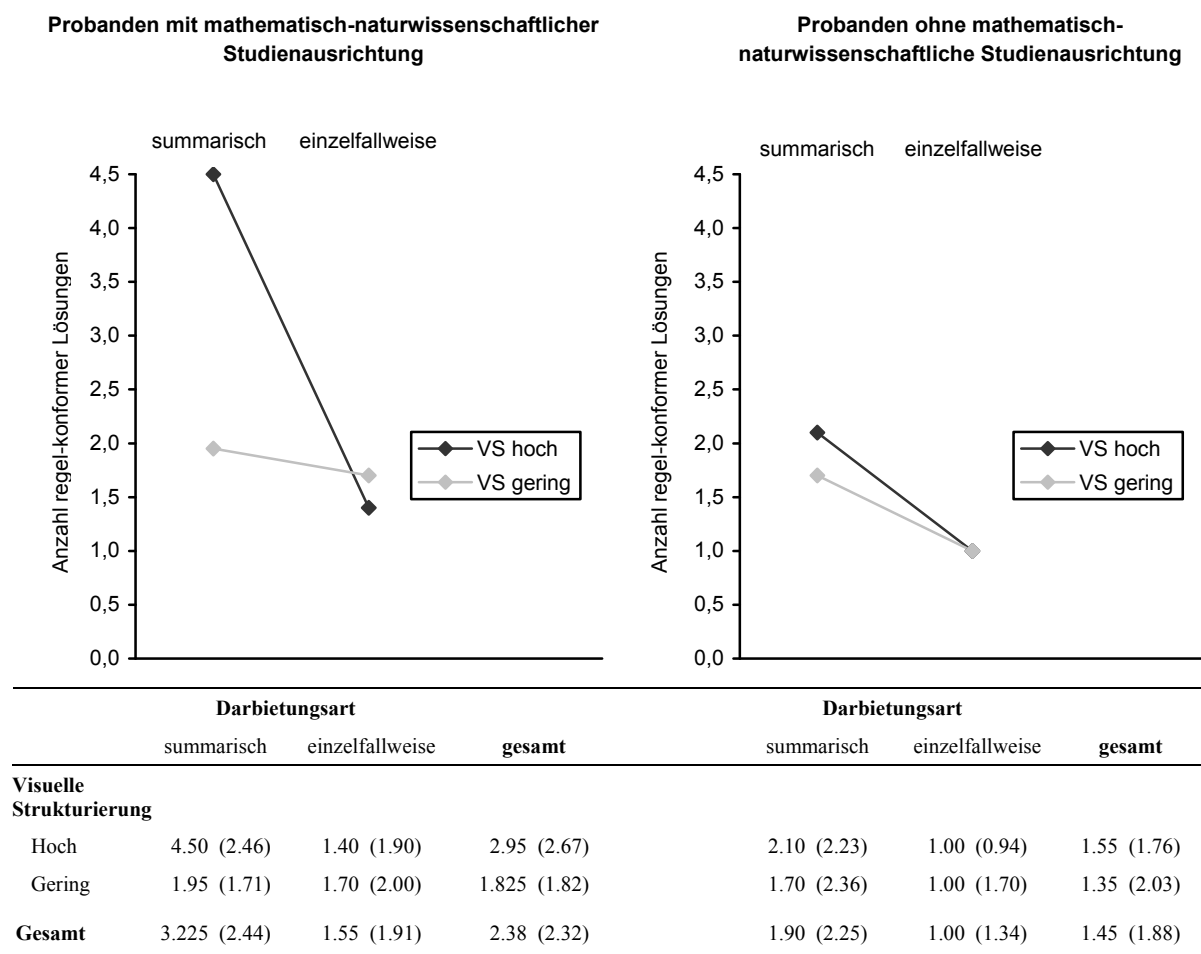
## 9.7 Einfluss von Studiausrichtung und Unterschieden im numerischem Denken

### 9.7.1 Einfluss der Studiausrichtung

#### 9.7.1.1 Ergebnisse

##### Anzahl regel-konformer Lösungen

Wie in Experiment 1 wurde vor Durchführung des Experiments erfasst, ob die Probanden in ihrem Studium eher mathematisch-naturwissenschaftlich oder eher geisteswissenschaftlich orientiert sind. Dementsprechend wurden sie in zwei gleichgroße Gruppen aufgeteilt und zufällig auf die vier Experimentalbedingungen verteilt. Abbildung 34 zeigt, dass die Gruppe mit mathematisch-naturwissenschaftlicher Studiausrichtung in allen Bedingungen höhere Mittelwerte im Hinblick auf die *Anzahl regel-konformer Lösungen* aufweist als die Gruppe ohne mathematisch-naturwissenschaftliche Studieninhalte.



**Abbildung 34:** Experiment 2: Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) getrennt nach Studiausrichtung in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*.

Unter Nichtberücksichtigung der Kovariate *numerisches Denken* wurde mit den Daten eine dreifaktorielle Varianzanalyse mit dem neuen Faktor *Studienausrichtung* und den beiden bereits untersuchten Faktoren *Darbietungsart* und *visuelle Strukturierung* gerechnet. Darüber hinaus wurden für jede Stufe des Faktors *Studienausrichtung* getrennte zweifaktorielle Varianzanalysen durchgeführt. In den Analysen mit Probanden aus mathematisch-naturwissenschaftlichen Fächern ist wegen des beträchtlichen Abstandes von *SD sum* zu den anderen drei Experimentalbedingungen mit stärkeren Effekten als in den Analysen mit Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung zu rechnen. Tabelle 15 gibt einen Überblick über die varianzanalytischen Ergebnisse:

**Tabelle 15:** Experiment 2: Ergebnisse für die abhängige Variable *Anzahl regel-konformer Lösungen* aus der dreifaktoriellen Varianzanalyse mit dem zusätzlichen Faktor *Studienausrichtung* über alle Probanden hinweg und den 2 (*Darbietungsart*) x 2 (*visuelle Strukturierung*) Varianzanalysen für die beiden nach *Studienausrichtung* getrennten Probandengruppen. Signifikante P-Werte sind **fett gedruckt**, Trends wurden unterstrichen.

Faktor	F-Wert	MSE	P-Wert
<b>Alle Probanden</b>			
Studienausrichtung (1,72)	4.549	3.864	<b>.036</b>
Darbietungsart (1,72)	8.579		<b>.005</b>
Visuelle Strukturierung (1,72)	2.272		.136
Studienausrichtung x Darbietungsart (1,72)	1.303		.264
Studienausrichtung x visuelle Strukturierung (1,72)	.777		.381
Darbietungsart x visuelle Strukturierung (1,72)	3.417		<u>.069</u>
Studienausrichtung x Darbietungsart x visuelle Strukturierung (1,72)	1.942		.168
<b>Pbn mit mathematisch-naturwissenschaftlicher Studienausrichtung</b>			
Darbietungsart (1,36)	7.179	3.583	<b>.011</b>
Visuelle Strukturierung (1,36)	2.788		.104
Darbietungsart x visuelle Strukturierung (1,36)	4.561		<b>.040</b>
<b>Pbn ohne mathematisch-naturwissenschaftliche Studienausrichtung</b>			
Darbietungsart (1,36)	2.260	4.145	.141
Visuelle Strukturierung (1,36)	.112		.740
Darbietungsart x visuelle Strukturierung (1,36)	.112		.740

Die an dieser Stelle wichtigen Ergebnisse aus Tabelle 15 sind die Signifikanz des Faktors *Studienausrichtung* und die Unterschiede zwischen den beiden Probandengruppen: Während in den Analysen mit Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung nicht einmal der Faktor *Darbietungsart* signifikant wurde, so wurden in den Analysen über Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung sowohl *Darbietungsart* als auch die Interaktion *Darbietungsart x visuelle Strukturierung* signifikant. Der P-Wert für *visuelle Strukturierung* kam in Trendnähe. Hauptgrund für die Unterschiede

zwischen den beiden Probandengruppen ist ihr unterschiedliches Abschneiden unter der Bedingung *SD sum*, wo Probanden mit mathematisch-naturwissenschaftlicher Studiaausrichtung wesentlich höhere Werte erzielen als Probanden ohne mathematisch-naturwissenschaftliche Studiaausrichtung.

#### *Bearbeitungszeit und Verarbeitungseffizienz*

Mit den abhängigen Variablen *Bearbeitungszeit* und *Verarbeitungseffizienz* haben wir hinsichtlich des Designs entsprechende Analysen durchgeführt. Die Mittelwerte für diese beiden Variablen, die in Tabelle 16 eingetragen sind, zeigen ebenfalls, dass mathematisch-naturwissenschaftlich ausgerichtete Probanden schneller und effizienter arbeiten.

**Tabelle 16:** Experiment 2: Bearbeitungszeit in Minuten : Sekunden und Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von *Studiaausrichtung* ("NW" = mathematisch-naturwissenschaftliche Ausrichtung; "GW" = geisteswissenschaftliche Ausrichtung) und den vier Experimentalbedingungen *SD sum*, *FT sum*, *VT tbt* und *FT tbt*.

	<b>SD sum</b>		<b>FT sum</b>		<b>VT tbt</b>		<b>FT tbt</b>		<b>Gesamt</b>	
	NW	GW	NW	GW	NW	GW	NW	GW	NW	GW
<b>Bearbeitungszeit</b>	3 : 42 (1 : 14)	8 : 40 (5 : 51)	4 : 25 (2 : 05)	5 : 37 (2 : 28)	16 : 23 (7 : 38)	16 : 33 (4 : 23)	22 : 46 (7 : 13)	24 : 10 (11 : 48)	11 : 49 (9 : 41)	13 : 44 (9 : 58)
<b>Verarbeitungseffizienz</b>	1.32 (.89)	.40 (.55)	.55 (.46)	.41 (.79)	.09 (.13)	.07 (.08)	.08 (.10)	.04 (.06)	.51 (.71)	.23 (.50)

Die varianzanalytischen Vergleiche über alle Probanden hinweg (auf deren ausführliche Darstellung hier aus Platzgründen verzichtet wird, siehe Tabelle A4 im Anhang A2.4 mit einer Auflistung signifikanter und nicht signifikanter Faktoren und Interaktionen) zeigten wie vorhergesagt, dass der Faktor *Studiaausrichtung* zwar nicht für die *Bearbeitungszeit*, aber für die *Verarbeitungseffizienz* signifikant wurde. Darüber hinaus wurde für die *Verarbeitungseffizienz* die Interaktion *Studiaausrichtung* x *Darbietungsart* signifikant, außerdem ergaben sich Trends für die Interaktion *Studiaausrichtung* x *visuelle Strukturierung* und die dreifache Interaktion.

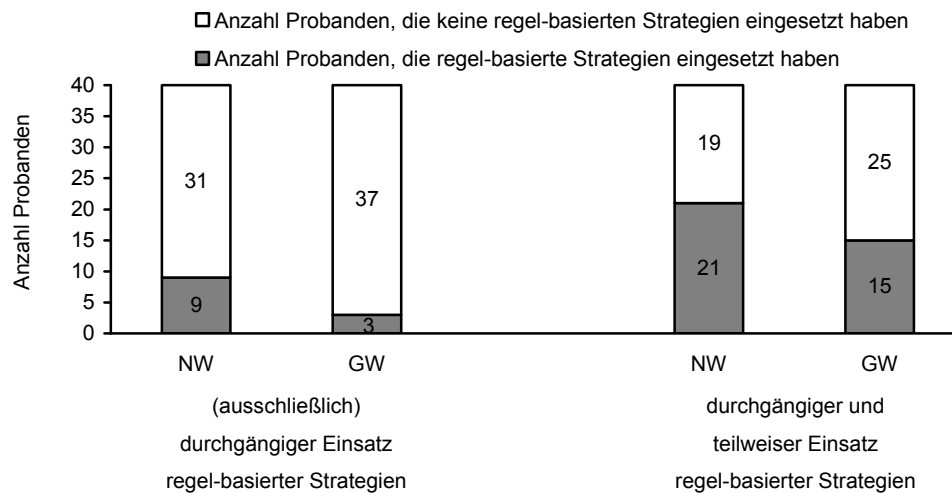
Für jede abhängige Variable wurden außerdem in Analogie zur *Anzahl regel-konformer Lösungen* für jede Stufe des Faktors *Studiaausrichtung* getrennte Varianzanalysen ohne Hinzunahme der Kovariate *numerisches Denken* durchgeführt. Eine Auflistung der signifikant gewordenen Faktoren und Interaktionen befindet sich ebenfalls in Tabelle A4 im Anhang A2.4. Wesentliches Ergebnis aus diesen Analysen sind auch hier wieder die Unterschiede zwischen den beiden Probandengruppen. In der Gruppe der mathematisch-naturwissenschaftlichen ausgerichteten Studenten wurden alle Haupteffekte signifikant ( $p < .05$ ). Die Interaktion *visuelle Strukturierung* x *Darbietungsart* hatte nur auf die

*Verarbeitungseffizienz*, nicht aber auf die *Bearbeitungszeit* ( $p = .106$ ) einen signifikanten Einfluss. Die signifikante Interaktion bei der *Verarbeitungseffizienz* rührt daher, weil die mathematisch-naturwissenschaftlich ausgerichteten Studenten bei summarischer Darbietung, nicht aber bei Einzelfalldarbietung effizienter (weil regel-konformer) mit Grafiken arbeiten. In der Gruppe der nicht mathematisch-naturwissenschaftlichen ausgerichteten Studenten waren die Effekte wie schon bei der abhängigen Variable *Anzahl regel-konformer Lösungen* deutlich schwächer oder gar nicht ausgeprägt, lediglich der Faktor *Darbietungsart* unterschritt bei allen abhängigen Variablen das Signifikanzniveau von  $\alpha = 5\%$ . Der Faktor *visuelle Strukturierung* verfehlte hingegen die 5 %-Marke bei allen abhängigen Variablen deutlich und die Interaktion *visuelle Strukturierung*  $\times$  *Darbietungsart* wurde nur für die *Bearbeitungszeit* signifikant ( $p = .022$ ), weil die nicht mathematisch-naturwissenschaftlich ausgerichteten Studenten bei summarischer Darbietung Texte und bei Einzelfalldarbietung die Vierfeldertafeln schneller bearbeiten. .

Die Ergebnisse aus den Varianzanalysen haben gezeigt, dass die Hypothesen H8a und H8b beibehalten werden können, der Faktor *Studienausrichtung* hat wie erwartet auf die beiden abhängigen Variablen *Anzahl regel-konformer Lösungen* und *Verarbeitungseffizienz* einen signifikanten Einfluss, nicht aber auf die *Bearbeitungszeit*. Mathematisch-naturwissenschaftlich ausgerichtete Studenten arbeiten effizienter und produzieren häufiger regel-konforme Lösungen als nicht mathematisch-naturwissenschaftlich ausgerichtete Studenten. Darüber hinaus zeigen die Ergebnisse, dass die Auswirkungen der Faktoren *Darbietungsart* und *visuelle Strukturierung des darstellenden Mediums* bei mathematisch-naturwissenschaftlich ausgerichteten Probanden wesentlich größer sind als bei nicht mathematisch-naturwissenschaftlich ausgerichteten Probanden. Die signifikant gewordenen Interaktionen lassen sich dahingehend interpretieren, dass mathematisch-naturwissenschaftlich ausgerichtete Studenten von summarisch dargebotenen Informationen im Stapeldiagramm mehr profitieren als nicht mathematisch-naturwissenschaftlich ausgerichtete Studenten.

### *Strategienwahl*

Häufigkeitsanalysen ergaben, dass die 40 mathematisch-naturwissenschaftlich ausgerichteten Studenten häufiger durchgängig oder zumindest teilweise regel-basierte Strategien eingesetzt haben als die 40 Probanden mit geisteswissenschaftlicher Studienausrichtung. Die Ergebnisse dieser Häufigkeitsanalysen können aus Abbildung 35 entnommen werden.



**Abbildung 35:** Experiment 2: Anzahl von Probanden, die durchgängig oder zumindest teilweise regel-basierte Strategien in Abhängigkeit von der *Studienausrichtung* ("NW" = mathematisch-naturwissenschaftlich"; "GW" = geisteswissenschaftlich bzw. nicht mathematisch-naturwissenschaftlich) eingesetzt haben.

Während für den Zusammenhang von *Studienausrichtung* und *Strategietyp* ein Trend ausgemacht werden konnte ( $\chi^2 = 3.529$ ,  $p = .060$ ; einseitiger P-Wert nach Fisher's exaktem Test = .057), dass mathematisch-naturwissenschaftlich ausgerichtete Probanden häufiger durchgängig regel-basierte Strategien einsetzten als nicht mathematisch-naturwissenschaftlich ausgerichtete Probanden, verschwand dieser Trend, wenn zusätzlich diejenigen Probanden miteinbezogen wurden, die während der Aufgabenbearbeitung zu einer regel-basierten Strategie gewechselt haben. Weder ein  $\chi^2$ -Test, noch Fisher's exakter Test ergaben einen signifikanten Zusammenhang zwischen den beiden Variablen *teilweiser Einsatz regel-basierter Strategien* und *Strategietyp* ( $\chi^2 = 1.818$ ,  $p = .178$ ; einseitiger P-Wert nach Fisher's exaktem Test = .131), obwohl eine Trendnähe unübersehbar ist.

#### 9.7.1.2 Diskussion

*Anzahl regel-konformer Lösungen:* Warum produzieren mathematisch-naturwissenschaftlich ausgerichtete Studenten häufiger regel-konforme Lösungen und arbeiten dadurch effizienter als nicht mathematisch-naturwissenschaftlich ausgerichtete Studenten? Neben möglichen Unterschieden hinsichtlich aufgabenspezifischem Vorwissen und Vertrautheit im Umgang mit Stapeldiagrammen kann angenommen werden, dass mathematisch-naturwissenschaftlich ausgerichtete Studenten höhere Werte im numerischen Denken erzielen als nicht mathematisch-naturwissenschaftlich ausgerichtete Studenten. Diese gerichtete Hypothese wurde mittels eines T-Tests für unabhängige Stichproben überprüft und bestätigt ( $MW_{NW} = 102.69$ ;  $MW_{GW} = 96,83$ ;  $T_{78} = 4.595$ ,  $p < .001$ ). Aus den Kovarianzanalysen mit den beiden

Faktoren *Darbietungsart* und *visuelle Strukturierung des darstellenden Mediums* wissen wir, dass der Einfluss der Kovariate *numerisches Denken* nur für die Anzahl regel-konformer Lösungen ( $p < .001$ ), nicht aber für die Bearbeitungszeit ( $p = .669$ ) hochsignifikant wurde. Wenn die Kovariate *numerisches Denken* einen signifikanten Einfluss auf die Anzahl regel-konformer Lösungen hat, dann sollte der Faktor *Studienausrichtung*, der zuvor in der Varianzanalyse (ohne Kovariate) mit den drei Faktoren *Darbietungsart*, *visuelle Strukturierung des darstellenden Mediums* und *Studienausrichtung* signifikant geworden war ( $p = .036$ ) in entsprechenden Kovarianzanalysen mit der Kovariate *numerisches Denken* nicht signifikant werden. Auch diese Annahme konnte bestätigt werden, der F-Wert für den Einfluss von *Studienausrichtung* auf die abhängige Variable *Anzahl regel-konformer Lösungen* war kleiner als 1, der entsprechende p-Wert betrug .465. Das bedeutet, dass der Einfluss des Faktors *Studienausrichtung* zu einem erheblichen Teil auf Unterschiede im numerischen Denken zurückzuführen ist.

Besonders ausgeprägt ist der Unterschied zwischen den Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung und solchen ohne mathematisch-naturwissenschaftliche Studienausrichtung unter der Bedingung *SD sum*. Dies bestätigt Befunde von Maichle (1992), aus denen sich ergab, dass Abiturienten, die sich stärker für Mathematik und Naturwissenschaften interessierten, Diagramme besser interpretieren konnten als Abiturienten, die mehr Interesse für andere Fächer aufbrachten.

*Bearbeitungszeit:* In Verbindung mit den Bearbeitungszeiten haben wir keine unterschiedlichen Einflüsse des Faktors *Studienausrichtung* vorhergesagt. Diese Annahme bestätigte sich auch. In Verbindung mit den Kovarianzanalysen haben wir bereits das Ergebnis diskutiert, dass die beiden medienbezogenen Faktoren *Darbietungsart* und *visuelle Strukturierung des darstellenden Mediums* einen großen Einfluss auf die Bearbeitungszeiten hatten. Durch die zusätzliche Unterteilung nach Studienausrichtung kann im Nachhinein auch das etwas überraschende Ergebnis plausibel gemacht werden, dass Probanden bei summarischer Informationsdarbietung Fliesstexte schneller bearbeiteten als Stapeldiagramme. Tabelle 16 zeigt nämlich, dass die insgesamt langsamere Verarbeitung von Stapeldiagrammen im Vergleich zu Fliesstexten mit summarischer Informationsdarbietung ausschließlich auf Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung zurückzuführen ist, weil diese im Gegensatz zu den Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung für die Informationsverarbeitung aus Stapeldiagrammen über drei Minuten länger benötigen als für die Informationsverarbeitung aus Fliesstexten mit summarischer Informationsdarbietung. Probanden mit mathematisch-naturwissenschaftlicher



Studienausrichtung verarbeiten hingegen Informationen aus Stapeldiagrammen knapp fünf Minuten schneller als Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung, vermutlich weil sie im Umgang mit Stapeldiagrammen vertrauter und weil sie aufgrund ihrer Vorteile im numerischen Denken zu einer schnelleren Berechnung der Stärke von Kausalzusammenhängen in der Lage sind.

*Verarbeitungseffizienz:* Die für die Verarbeitungseffizienz signifikante Interaktion *Studienausrichtung x Darbietungsart* lässt sich nach einem Blick auf Tabelle 16 dadurch erklären, dass Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung unter der Bedingung *FT sum* genauso effizient (weil trotz geringerer Anzahl regel-konformer Lösungen schneller) gearbeitet haben wie unter der Bedingung *SD sum*, während Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung unter der Bedingung *SD sum* deutlich effizienter (weil mit höherer Anzahl regel-konformer Lösungen) waren als unter der Bedingung *FT sum*. Die signifikanten Interaktionen und Trends mit dem Faktor *Studienausrichtung* lassen sich folgendermaßen plausibel erklären: Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung arbeiten mit Stapeldiagrammen (*SD sum*), die ihnen aus dem Studium sehr vertraut sein dürften, sehr viel effizienter und profitieren im Hinblick auf das Ziehen von Schlussfolgerungen, wie Erfolgs- und Basisrate regel-konform zueinander in Beziehung zu setzen sind, stark von ihnen. Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung haben vermutlich mit der Auswertung und Interpretation von aus dem Studium weniger vertrauten Stapeldiagrammen mehr Schwierigkeiten und arbeiten dementsprechend langsamer und weniger effizient. Offensichtlich können also Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung sowohl hinsichtlich der Anzahl regel-konformer Lösungen als auch hinsichtlich Bearbeitungszeit und Verarbeitungseffizienz von einer Informationsanordnung im Stapeldiagramm stark profitieren, während Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung dies nicht tun können.

*Strategienwahl:* Bei den Häufigkeitsanalysen ergab sich zwar ein Trend, dass Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung häufiger durchgängig regel-basierte Strategien einsetzen als Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung, doch fielen die beobachteten Häufigkeitsunterschiede geringer aus erwartet. Zudem ergab sich nicht einmal ein Trend für den Vergleich zwischen Probanden mit und ohne mathematisch-naturwissenschaftliche Studienausrichtung, wenn zusätzlich noch diejenigen Probanden mit berücksichtigt werden, die im Verlauf von Experiment 2 zum Gebrauch einer regel-basierten Strategie übergegangen sind. Hypothese H8c kann deshalb nur

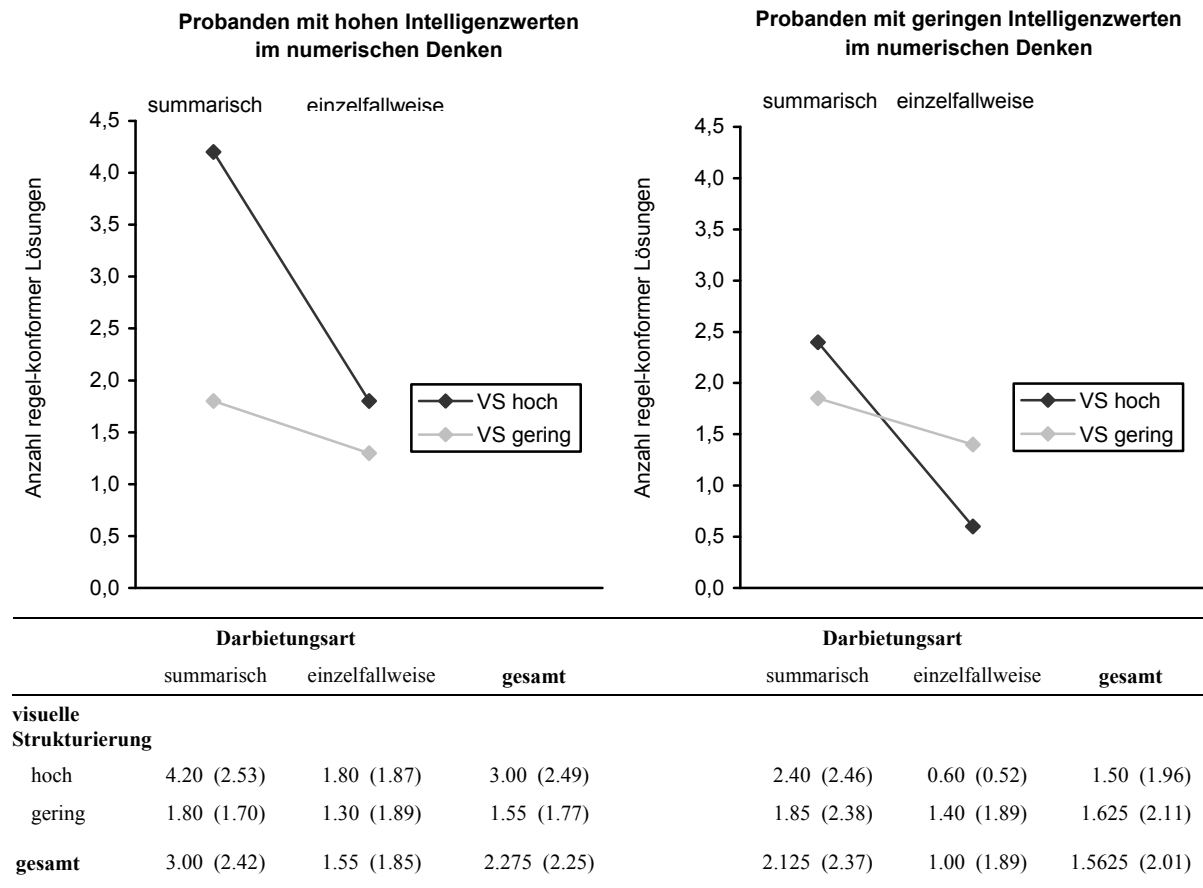
der Tendenz nach bestätigt werden. Wir vermuten, dass sich die Bodeneffekte, die sich aufgrund der hohen mentalen Beanspruchung in den drei Bedingungen *FT sum*, *VT tbt* und *FT sum* ergeben haben (nur 8,3 % der Probanden in diesen drei Bedingungen setzten durchgängig regel-basierte Strategien ein, weitere 28,3 % machten teilweise von regel-basierten Strategien Gebrauch), bei den Häufigkeitsanalysen stärker ausgewirkt haben als in den Varianzanalysen im Zusammenhang mit der Anzahl regel-konformer Lösungen. In Verbindung mit den Häufigkeitsanalysen zur Strategiewahl wird lediglich erfasst, ob ein Proband regel-basiert vorgegangen ist oder nicht. Im Gegensatz zur Bestimmung der Anzahl regel-konformer Lösungen wurde nicht erfasst, wie häufig der einzelne Proband regel-basiert vorgegangen ist oder nicht.

## 9.7.2 Einfluss von Unterschieden im numerischen Denken

### 9.7.2.1 Ergebnisse

#### *Anzahl regel-konformer Lösungen*

In der Kovarianzanalyse mit der abhängigen Variablen *Anzahl regel-konformer Lösungen* konnte ein signifikanter Einfluss der Kovariate *numerisches Denken* nachgewiesen werden. Um herauszufinden, ob die beiden Faktoren *Darbietungsart* und *visuelle Strukturierung des darstellenden Mediums* einen unterschiedlichen Einfluss auf Probanden mit hohen und geringen Intelligenzwerten im numerischen Denken haben, wurden die Probanden in jeder der vier Experimentalbedingungen *SD sum*, *FT sum*, *VT tbt* und *FT tbt* in zwei gleich große Probandengruppen mit entweder hohen oder geringen Intelligenzwerten im numerischen Denken aufgeteilt. Über alle Experimentalbedingungen hinweg konnten so zwei Teilstichproben mit je 40 Probanden hergestellt werden, die sich hinsichtlich ihrer erreichten Werte im *numerischen Denken* unterscheiden (*hoch* vs. *gering*). Getrennt für die beiden Teilstichproben wurden die Mittelwerte für die abhängige Variable *Anzahl regel-konformer Lösungen* berechnet. Sie sind in Abbildung 36 eingetragen.



**Abbildung 36:** Experiment 2: Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) getrennt nach Intelligenzunterschieden im numerischen Denken (*ND hoch* vs. *ND gering*) und in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*.

Bemerkenswert ist, dass sich die Mittelwerte für die beiden Teilstichproben nur dann im größeren Ausmaß unterscheiden, wenn die visuelle Strukturierung des darstellenden Mediums hoch ist (*SD sum* und *VT tbt*). Bei geringer visueller Strukturierung des darstellenden Mediums (*VT tbt* und *FT tbt*) unterscheiden sich die Mittelwerte der beiden Teilstichproben kaum, es spielt also für die Wirksamkeitsbeurteilung auf der Basis von visuell gering strukturierten Fliesstexten keine Rolle, ob die Probanden aus einer studentischen Stichprobe hohe oder geringe Intelligenzwerte im numerischen Denken aufweisen.

Die Mittelwerte aus Abbildung 36 legen nahe, dass nur der Faktor *visuelle Strukturierung* mit der Variable *numerisches Denken* (*ND hoch* versus *ND gering*) interagiert, nicht aber der Faktor *Darbietungsart*. Probanden mit einem hohen Wert im numerischen Denken können die visuell hoch strukturierte Informationsanordnung in Stapeldiagrammen und ansatzweise auch in Vierfeldertafeln nutzen, um häufiger regel-konforme Lösungen zu produzieren, während Probanden mit einem geringen Wert im numerischen Denken von einer hohen visuellen

Strukturierung des darstellenden Mediums vor allem in der Einzelfallbedingung eher benachteiligt werden.

Wie schon im Zusammenhang mit den beiden unterschiedlichen Probandengruppen in Abhängigkeit vom Faktor *Studienausrichtung*, so sollte auch für die beiden Gruppen von Probanden mit hohen und mit geringen Intelligenzwerten im *numerischen Denken* herausgefunden werden, ob sich in beiden Teilstichproben für die abhängige Variable *Anzahl regel-konformer Lösungen* unterschiedliche Effekte in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung des darstellenden Mediums* ergeben. Erneut wurden deshalb für jede der beiden Teilstichproben zweifaktorielle Varianzanalysen durchgeführt. Die Ergebnisse sind in Tabelle 17 abgebildet.

**Tabelle 17:** Experiment 2: Ergebnisse aus den 2 (*Darbietungsart*) x 2 (*visuelle Strukturierung*) Varianzanalysen für die beiden Teilstichproben (*ND hoch* und *ND gering*) mit der abhängigen Variable *Anzahl regel-konformer Lösungen*. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen.

Faktor	F-Wert	MSE	P-Wert
<b>Pbn mit hohen Intelligenzwerten im numerischen Denken</b>			
Darbietungsart <sub>(1,36)</sub>	5.316	4.092	<b>.027</b>
visuelle Strukturierung <sub>(1,36)</sub>	4.962		<b>.032</b>
Darbietungsart x visuelle Strukturierung <sub>(1,36)</sub>	2.091		.157
<b>Pbn mit geringen Intelligenzwerten im numerischen Denken</b>			
Darbietungsart <sub>(1,36)</sub>	3.392	3.892	<u>.074</u>
visuelle Strukturierung <sub>(1,36)</sub>	.058		.811
Darbietungsart x visuelle Strukturierung <sub>(1,36)</sub>	1.086		.304

Tabelle 17 zeigt, dass sich für jede Teilstichprobe unterschiedliche Haupteffekte ergeben. In der Gruppe von Probanden mit hohen Intelligenzwerten im numerischen Denken werden sowohl der Faktor *Darbietungsart* als auch der Faktor *visuelle Strukturierung* signifikant, die Interaktion kommt in Trendnähe. In der Gruppe der Probanden mit geringen Intelligenzwerten im numerischen Denken wird für den Faktor *Darbietungsart* nur ein Trend erzielt, der Faktor *visuelle Strukturierung* wird wie die Interaktion nicht einmal annähernd signifikant.

#### *Bearbeitungszeit und Verarbeitungseffizienz*

Da die *Kovariate* numerisches Denken erwartungsgemäß keinen signifikanten Einfluss auf die Bearbeitungszeiten hatte, wurden keine Varianzanalysen mit einem zweistufigen Faktor numerisches Denken durchgeführt, in denen der Einfluss dieses Faktors auf die abhängigen Variablen *Bearbeitungszeit* und *Verarbeitungseffizienz* untersucht wurde.

### 9.7.2.2 Diskussion

Signifikante Effekte treten in Experiment 2 vor allem dann auf, wenn nur Probanden mit mathematisch-naturwissenschaftlicher Studiaausrichtung oder solche mit hohen Intelligenzwerten im numerischen Denken untersucht werden. In diesen Teilstichproben wirkt sich vor allem bei summarischer Darbietung eine hohe visuelle Strukturierung, wie sie durch die gruppierte und übersichtliche Anordnung von Kausalinformationen im Stapeldiagramm gewährleistet ist, begünstigend im Sinne eines regel-konformen In-Beziehung-Setzens von Erfolgs- und Basisrate aus. Bei Probanden ohne mathematisch-naturwissenschaftliche Studiaausrichtung oder mit geringen Werten im numerischen Denken waren hingegen in allen vier Experimentalbedingungen ausgeprägte Bodeneffekte zu beobachten.

Kaum überraschend belegen die Ergebnisse aus Experiment 2 wie auch weitere inferenzstatistische Häufigkeitsanalysen den engen Zusammenhang zwischen *Studiaausrichtung* und Intelligenzwerten im numerischen Denken. So befanden sich nach einem Mediansplit mit allen Intelligenztestwerten in der Gruppe der mathematisch-naturwissenschaftlich ausgerichteten Probanden 30 Probanden mit hohen und 10 Probanden mit geringen Intelligenzwerten im numerischen Denken, in der Gruppe der nicht mathematisch-naturwissenschaftlich ausgerichteten Probanden verhielt es sich genau umgekehrt ( $\chi^2 = 20.000$ ,  $p < .001$ ; einseitiger P-Wert nach Fisher's exaktem Test  $< .001$ ; Phi-Koeffizient = .50).

Der Einsatz regel-basierter Strategien macht den Einsatz komplexer Rechenoperationen erforderlich. Deshalb lässt sich mit dem Ergebnis, dass mathematisch-naturwissenschaftlich ausgerichtete Studenten höhere Intelligenzwerte im numerischen Denken aufweisen, gut erklären, warum mathematisch-naturwissenschaftlich ausgerichtete Studenten komplexere und häufiger regel-konforme Lösungen generieren. Darüber hinaus ist auch zu vermuten, dass diese Probandengruppe insbesondere im Umgang mit Stapeldiagrammen vertrauter agiert als nicht mathematisch-naturwissenschaftlich ausgerichtete Studenten. Auf eine höhere Vertrautheit von mathematisch-naturwissenschaftlich ausgerichteten Studenten im Umgang mit Stapeldiagrammen und Prozentrechnen haben schon die Ergebnisse aus der Nachbefragung von Experiment 1 hingewiesen, in der mathematisch-naturwissenschaftlich ausgerichtete Studenten unter anderem angaben, dass sie im Vergleich zu nicht mathematisch-naturwissenschaftlich ausgerichteten Studenten häufiger Diagramme anfertigen, Prozentrechnen und sich sicherer im Prozentrechnen und in den Grundrechenarten fühlen.

Die beobachteten Teilstichprobeneffekte sind aus zwei Gründen bedeutsam: Zum einen belegen sie, welche unterschiedlichen Auswirkungen ein und dasselbe Medium bei der

Bearbeitung von Aufgaben zur Beurteilung kausaler Wirksamkeiten auf verschiedene Probandenstichproben haben kann. In Experiment 2 hat sich gezeigt, dass Probanden mit mathematisch-naturwissenschaftlich Studenausrichtung und solche mit hohen Intelligenzwerten im numerischen Denken durch die Präsentation von Stapeldiagrammen im hohen Maße dazu angeregt werden können, regel-konforme und komplexe Lösungen mit hoher Effizienz zu generieren, während Probanden aus nicht mathematisch-naturwissenschaftlichen Studienfächern und solche mit geringen Intelligenzwerten im numerischen Denken nur wenig von dem visuell hoch strukturierten und gruppierten Informationsangebot in Stapeldiagrammen profitieren können.

Zum anderen wird durch die beobachteten Teilstichprobeneffekte demonstriert, wie bedeutsam individuelle Unterschiede bei der kausalen Urteilsbildung werden können. Zuvor wurde die Bedeutsamkeit individueller Unterschiede durch die Signifikanz der Kovariate *numerisches Denken* belegt.

Neben Unterschieden im numerischen Denken und hinsichtlich der Vertrautheit im Umgang mit einzelnen Medien darf angenommen werden, dass auch ein unterschiedliches aufgabenspezifisches Vorwissen und motivationale Aspekte zur individuellen Variation bei der Wirksamkeitsbeurteilung beitragen. Die motivationalen Aspekte können zum einen innerpsychische Bereitschaften (wie z.B. Interesse an der Aufgabenstellung oder Tagesform) betreffen, zum anderen können sie sich auf ein sowohl qualitativ als auch quantitativ unterschiedliches Motivierungspotential von unterschiedlichen medialen Rahmenbedingungen bei der Aufgabenbearbeitung beziehen. So ist zum Beispiel in Experiment 2 nicht auszuschließen, dass der Unterschied zwischen den beiden Bedingungen *SD sum* und *FT sum* hinsichtlich der Anzahl regel-konformer Lösungen auch darin begründet ist, dass die Probanden durch Texte weniger als durch Diagramme dazu angeregt oder aufgefordert werden, Berechnungen zur Wirksamkeitsbestimmung durchzuführen oder komplexe Verhältnisse zwischen Erfolgs- und Basisrate rechnerisch zu ermitteln. Unterschiedliche mediale Rahmenbedingungen könnten also auch bei der kausalen Urteilsbildung zu unterschiedlichen Erwartungshaltungen oder Vorannahmen hinsichtlich der Aufgabenbearbeitung führen. Indem sie unterschiedliche Informationsverarbeitungsprozesse anregen, dürften mediale Rahmenbedingungen in einer engen Wechselbeziehung mit individuellen Voraussetzungen auf Probandenseite stehen. Beispielsweise könnte ein Geisteswissenschaftler mit unterdurchschnittlicher Intelligenz im numerischen Denken durch ein Stapeldiagramm davon abgehalten werden, komplexere Berechnungen zur Bestimmung der Wirksamkeit anzustellen, während ein Naturwissenschaftler mit hoher Intelligenz im

numerischen Denken gerade durch das Diagramm, aber nicht durch einfache Fliesstexte zur Wahl einer elaborierten Strategie mit komplexen Rechenanforderungen motiviert wird.

Individuelle Unterschiede wie zum Beispiel im numerischen Denken beeinflussen nicht nur die Anzahl regel-konformer Lösungen insgesamt, sie haben auch einen Einfluss darauf, von welcher regel-basierten Strategie die Probanden Gebrauch machen: So war in weiteren Häufigkeitsanalysen zu beobachten, dass Probanden mit geringen Werten im numerischen Denken vor allem unter der Bedingung *SD sum* eher dazu neigen, die  $\Delta P$ -Regel einzusetzen, wenn sie regel-basiert vorgehen (*ND gering*: mittlere Anzahl PPC-theoriekonformer Lösungen = .10; mittlere Anzahl  $\Delta P$ -konformer Lösungen = 2.3). Probanden mit hohen Werten im numerischen Denken machen dagegen häufiger von der Power PC-Theorie als von der weniger komplexen  $\Delta P$ -Regel Gebrauch (*ND hoch*: mittlere Anzahl PPC-theoriekonformer Lösungen = 2.6; mittlere Anzahl  $\Delta P$ -konformer Lösungen = 1.0). Ein über alle Probanden hinweg durchgeführter T-Test führte zu dem signifikanten Ergebnis, dass Probanden mit hohen Werten im numerischen Denken eine höhere Anzahl PPC-theoriekonformer Lösungen generieren als Probanden mit geringen Werten im numerischen Denken ( $MW_{ND\ hoch} = 1.213$ ;  $MW_{ND\ gering} = .375$ ;  $T_{78} = -2.123$ ;  $p < .039$ ). Hinsichtlich der Anzahl  $\Delta P$ -konformer Lösungen unterscheiden sich *ND hoch* und *ND gering* hingegen nicht ( $MW_{ND\ hoch} = .925$ ;  $MW_{ND\ gering} = 1.2$ ;  $T_{78} = .762$ ;  $p = .449$ ).

Zusammengefasst sprechen die Ergebnisse aus den Analysen der Teilstichproben dafür, dass die Wahl der Stichprobe wie in Experiment 1 das Auftreten oder Nicht-Auftreten von Haupteffekten erheblich modifizieren kann. Darüber hinaus bestätigen die Teilstichprobenanalysen die im Zusammenhang mit Abbildung 5 formulierte Annahme, dass medienbezogene (wie z.B. die *Darbietungsart* oder die *visuelle Strukturierung*) und personenbezogene Faktoren (z.B. *numerisches Denken*) miteinander interagieren und kognitive Prozesse bei der Beurteilung der Stärke von Kausalzusammenhängen unterschiedlich beeinflussen können.

## 9.8 Abschließende Diskussion von Experiment 2

Im Gegensatz zu Experiment 1 konnten die meisten Hypothesen aus Experiment 2 bestätigt werden:

*Einfluss des Faktors Darbietungsart:* Der Faktor *Darbietungsart* hatte einen hoch signifikanten Einfluss auf die Anzahl regel-konformer Lösungen, die Bearbeitungszeit, die Verarbeitungseffizienz und die Strategiewahl. Bei summarischer Darbietung haben die Probanden häufiger regel-basierte Strategien eingesetzt und mehr regel-konforme Lösungen gebildet, sowie schneller und effizienter gearbeitet. Im Hinblick auf die *Anzahl regel-konformer Lösungen* werden dadurch Studien bestätigt (Arkes & Harkness, 1983; Buehner et al., 2003; Kao & Wasserman, 1993; Lober & Shanks, 2000; Ward & Jenkins, 1965), in denen gezeigt werden konnte, dass Probanden bei summarisch dargebotenen Ursache-Wirkungs-Zusammenhängen eher regel-konforme Lösungen produzieren als bei einer einzelfallweisen Darbietung.

Shaklee & Mims (1982) haben die komplexere Strategiewahl und die höhere Urteilsgenauigkeit in einer Bedingung, in welcher die Probanden ihre einzelfallweise dargebotenen Ergebnisse summarisch in einer Häufigkeitstabelle zusammenfassen konnten, auf eine geringere Gedächtnisanforderung im Vergleich zu einer einzelfallweisen Darbietung ohne die Möglichkeit einer summarischen Ergebniszusammenfassung zurückgeführt. An Stelle des Begriffes *Gedächtnisanforderung* haben wir bislang den allgemeinen, aber vagen Begriff *mentale Beanspruchung* vorgezogen, weil wir uns dadurch die Möglichkeit offen lassen wollten, den häufigeren Einsatz komplexer regel-basierter Strategien bei summarischer Ergebnisdarbietung nicht nur auf reduzierte Gedächtnisanforderungen zurückführen zu müssen. In Experiment 3 sollen die unterschiedlichen Auswirkungen von summarischer und einzelfallweiser Darbietung genauer aufgeschlüsselt werden und untersucht werden, ob die höhere Anzahl regel-konformer Lösungen und der häufigere Einsatz regel-basierter Strategien bei summarischer Ergebnisdarbietung ausschließlich auf eine Entlastung des Arbeitsgedächtnisses durch externale Repräsentationen wie zum Beispiel eine Vierfeldertafel zurückzuführen ist.

Hinweise dafür, dass externale Repräsentationen nicht nur das Arbeitsgedächtnis entlasten, sondern darüber hinaus auch noch ein regel-konformes Zueinander-in-Beziehung-setzen von Erfolgs- und Basisrate salient machen können, stammen aus den Arbeiten von Zhang (2000, 1997; Zhang & Norman, 1994; s.a. Larkin & Simon, 1987) und aus Experiment 2. Nach Zhang (2000, 1997; Zhang & Norman, 1994; s.a. Larkin & Simon, 1987) fungieren externale



Repräsentationen nicht nur als Gedächtnisstütze, sondern sie können darüber hinaus die Beschaffenheit einer Aufgabe verändern und ihren Abstraktionsgrad einschränken. Zudem können sie den Suchaufwand für zusammengehörende Informationen reduzieren, die Informationsverarbeitung strukturieren und kognitives Verhalten steuern, sowie einfach zu verarbeitende perzeptuelle Inferenzen ermöglichen. Die Ergebnisse aus Experiment 2 belegen, dass bei summarischer Darbietung von Informationen in einem Stapeldiagramm die Probanden mehr regel-konforme Lösungen produzierten und häufiger regel-basierte Strategien einsetzten als bei einer summarischen Informationsdarbietung in einem Fliesstext. Der Unterschied zwischen den beiden Bedingungen *SD sum* und *FT sum* zeigt, dass es nicht nur eine Rolle spielt, ob die Informationen zur Beurteilung der Stärke von kausalen Zusammenhängen summarisch oder einzelfallweise dargeboten werden, sondern auch, wie diese Informationen in einem darstellenden Medium angeordnet sind und präsentiert werden. Im Gegensatz zu einem einfachen Fliesstext erhalten die Probanden in einem visuell hoch strukturierten Stapeldiagramm durch die räumliche Gegenüberstellung von verschiedenen Säulen für Erfolgs- und Basisrateninformationen, sowie durch die unterschiedliche Farbgebung für das Eintreten des Effekts verschiedene visuelle Hinweise darauf, welche Zellinformationen zusammen gehören und wie sie regel-konform zueinander in Beziehung zu setzen sind.

*Einfluss des Faktors visuelle Strukturierung des darstellenden Mediums:* Der Einfluss des Faktors *visuelle Strukturierung* war insgesamt schwächer als derjenige des Faktors *Darbietungsart*. Zudem hatte die visuelle Strukturierung des darstellenden Mediums in Abhängigkeit von der Darbietungsart einen unterschiedlichen Einfluss. So führte eine visuell hoch strukturierte Grafik bei summarischer Darbietung zu einer höheren Anzahl regel-konformer Lösungen, aber nicht zu schnelleren Bearbeitungszeiten. Umgekehrt wurde die visuell hoch strukturierte Vierfeldertafel bei einzelfallweiser Darbietung zwar schneller verarbeitet als ein einfacher Fliesstext, führte aber nicht zu einer höheren Anzahl regel-konformer Lösungen oder zu einem vermehrten Einsatz regel-basierter Strategien.

Die unterschiedlichen Auswirkungen des Faktors *visuelle Strukturierung* in Abhängigkeit von der *Darbietungsart* hängen damit zusammen, dass sich je nach Darbietungsart die Beschaffenheit der Aufgabe, die die Probanden bei der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen auszuführen haben, verändert. Bei einzelfallweiser Darbietung müssen die Einzelinformationen im Gedächtnis behalten werden und einer der vier Kategorien in Abbildung 1 zugeordnet werden. Mit jeder neuen Einzelinformation ändert sich eine der vier Zellhäufigkeiten und muss neu im Gedächtnis gespeichert werden. Bei

summarischer Darbietung entfallen diese Aufgaben von Datenzuordnung und Datenaggregation (s. Crocker, 1981), die Komplexität der Aufgabe reduziert sich und das Gedächtnis wird entlastet. Um zu einer regel-konformen Lösung zu gelangen, müssen stattdessen lediglich die bereits vorgegebenen Zellhäufigkeiten heran gezogen und regel-konform miteinander verrechnet werden.

Damit das kognitive System durch visuelle Hinweise dabei unterstützt werden kann, regel-konforme Vergleiche und Schlussfolgerungen vorzunehmen, ist es entscheidend, dass die dafür notwendigen Informationen erst einmal in summarischer Form als Häufigkeiten vorliegen. Nur wenn die Häufigkeiten external repräsentiert sind, können visuelle Hinweise Vergleichsprozesse induzieren, die salient machen, wie die vier Zellhäufigkeiten zueinander in Beziehung gesetzt werden, damit eine regel-konforme Lösung zu Stande kommt. Ebenso werden eine Verlagerung von Informationsverarbeitungsprozessen auf die Wahrnehmungsebene (Wickens & Carswell, 1995) und perzeptuelle Inferenzen (Larkin & Simon, 1987) erst dann möglich, wenn die zu verarbeitenden Informationen als Häufigkeiten vorliegen und durch ihre Anordnung mit zusätzlichen visuellen Hinweisen versehen sind.

Bei Einzelfalldarbietung müssen die zu verarbeitenden Häufigkeiten erst einmal gebildet und mental repräsentiert werden, d.h. sie liegen nicht als externale Repräsentationen vor. Liegen die zu verarbeitenden Häufigkeiten nicht in externalisierter Form vor, dann fehlen auch die visuellen Hinweise, durch welche Vergleiche und Schlussfolgerungen salient gemacht werden könnten, die zu einer Bildung regel-konformer Lösungen anregen. Bei summarischer Informationsdarbietung in einem Fliesstext wird den Probanden zwar das Ordnen und Anhäufen von Einzelinformationen vollständig abgenommen, aber es fehlen visuelle Hinweise darauf, wie die vorliegenden Häufigkeiten zueinander in Beziehung gesetzt werden könnten.

Während sich die beiden Einzelfallbedingungen hinsichtlich der Anzahl regel-konformer Lösungen nicht voneinander unterscheiden, wurde die visuell hoch strukturierte Vierfeldertafel mit Positionskreuz von den Probanden signifikant schneller verarbeitet als der Fliesstext mit Einzelfallinformationen. Wir vermuten, dass durch das Positionskreuz in der Vierfeldertafel und der daraus resultierenden Aufmerksamkeitsfokussierung die Organisation der Wahrnehmung und Handlungsabläufe wie das Zuordnen und Anhäufen von Einzelinformationen im Vergleich zum Fliesstext erleichtert und beschleunigt werden. Die in der Vierfeldertafel mit Positionskreuz enthaltenen Strukturierungshilfen wirken sich aber nicht auf die Anzahl regel-konformer Lösungen aus, da bei Einzelfalldarbietung keine

visuellen Hinweise enthalten sein können, wie die internal gebildeten Häufigkeiten regelkonform zueinander in Beziehung gesetzt werden können.

Der Blick auf die Ergebnisse von Experiment 2 zeigt, dass – anders als in Experiment 1 - der Einfluss medialer Faktoren unter Umständen auch sehr stark sein kann. Dies geschieht dann, wenn mediale Faktoren Aufgabenbeschaffenheit und Komplexität der Aufgabe entscheidend verändern. Dabei wirken sich mediale Faktoren nicht nur auf die Gedächtnisbeanspruchung aus, sondern auch auf Wahrnehmungsorganisation und Handlungsabläufe, sowie Prozesse des Schlussfolgerns. Besonders ausgeprägt können mediale Effekte dann sein, wenn sich die Einflüsse medialer Faktoren im Zusammenspiel mit der zu lösenden Aufgabe ergänzen. So ist auf der Grundlage der Ergebnisse aus Experiment 2 die kognitive Passung (Vessey, 1991, 1994) zwischen einem Stapeldiagramm mit summarischer Informationsdarbietung und der Aufgabe, die Stärke eines Ursache-Wirkungs-Zusammenhangs an Hand von Häufigkeiten zu ermitteln, besonders günstig.

*Einfluss der Fähigkeit zum numerischen Denken:* Die Annahme, dass unterschiedliche Fähigkeiten im numerischen Denken die Anzahl regelkonformer Lösungen und den Einsatz regelbasierter Strategien beeinflussen, konnte eindrucksvoll bestätigt werden. Dies bestätigt die Richtigkeit der im Zusammenhang mit dem Schaubild in Abbildung 5 aufgestellte Vermutung, dass die bei der Beurteilung der Stärke von kausalen Zusammenhängen ablaufenden kognitiven Prozesse und Mechanismen nicht nur von aufgaben- und medienbezogene Faktoren, sondern auch von individuellen Unterschieden und Voraussetzungen abhängen. Zusammengefasst lassen die Ergebnisse aus Experiment 2 den Schluss zu, dass regelbasierte Strategien vor allem von Probanden eingesetzt werden, die ein mathematisch-naturwissenschaftlich ausgerichtetes Studienfach absolvieren bzw. über eine hohe numerische Verarbeitungskapazität verfügen und denen die Informationen über Ursache-Wirkungs-Zusammenhänge summarisch in einem Stapeldiagramm dargeboten wurden. Andere Probanden haben weniger regelbasierte Strategien eingesetzt und weniger regelkonforme Lösungen gebildet, vor allem konnten sie kaum von einem visuell hoch strukturierten Stapeldiagramm mit summarischer Informationsdarbietung profitieren.

Experiment 2 lieferte wie schon Experiment 1 keine Belege für die Annahme, dass der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen ein einheitlicher Verarbeitungsmechanismus zugrunde liegen würde. Die Mehrheit der Probanden hat ihre Strategie über mehrere Urteilsprozesse hinweg gewechselt, außerdem wurden bei einzelfallweiser Darbietung und geringer visueller Strukturierung weniger regelbasierte

Strategien eingesetzt als bei summarischer Darbietung und hoher visueller Strukturierung. Von allen Probanden, die regel-basiert vorgehen, machten etwas mehr Probanden von der rechnerisch etwas weniger komplexen  $\Delta P$ -Regel Gebrauch. Probanden, die zu Lösungen gelangten, die einer Berechnung der kausalen Stärke entsprechend der Power PC Theorie entsprach, waren aber häufiger in der Gruppe der Probanden, die ihre Strategie nicht gewechselt haben, und unter Bedingungen mit summarischer Informationsdarbietung zu finden als Probanden, die eine den Vorhersagen der  $\Delta P$ -Regel entsprechende Strategie einsetzten.

## **10 Experiment 3: Einfluss von Notizen bei Einzelfalldarbietung**

Die Ergebnisse aus Experiment 2 belegen, dass die Probanden bei summarischer Darbietung von Informationen über Ursache-Wirkungs-Zusammenhängen häufiger regel-konforme Lösungen generierten als bei Einzelfalldarbietung. Eine Erklärung für dieses Ergebnis liegt darin, dass Informationsverarbeitung und insbesondere Arbeitsgedächtnis bei summarischer Darbietung geringer belastet werden als bei Einzelfalldarbietung. Im Gegensatz zur Einzelfalldarbietung wird den Probanden das Gruppieren und die parallele Durchführung mehrerer Zählvorgänge abgenommen, die verfügbare kognitive Kapazität kann so ganz auf das Zueinander-in-Beziehung-setzen der verschiedenen Häufigkeiten entsprechend den vier Zellen einer Kontingenztafel ausgerichtet werden. Den Probanden wird es dadurch leichter gemacht, regel-basierte Strategien entweder neu zu generieren oder die entsprechenden Schemata, die der Anwendung regel-basierter Strategien zugrunde liegen, aus dem Langzeitgedächtnis abzurufen und verfügbar zu machen.

Eine durch summarische Ergebniszusammenfassung reduzierte mentale Beanspruchung ist eine wesentliche Grundbedingung dafür, dass die Probanden bei der Wirksamkeitsbeurteilung zu regel-konformen Lösungen gelangen konnten. Bei Einzelfalldarbietung sind dazu nur wenige Probanden in der Lage – selbst dann nicht, wenn die Einzelinformationen übersichtlich in einer Vierfeldertafel angeordnet sind und somit den verschiedenen Ereignistypen leichter zugeordnet werden können. Eine hohe visuelle Strukturierung des darstellenden Mediums wurde in Experiment 2 als eine weitere mögliche Bedingung dafür genannt, dass Probanden häufiger regel-konforme Lösungen hervorbringen. Die Ergebnisse aus Experiment 2 zeigen aber, dass Probanden nur dann durch eine hohe visuelle Strukturierung des darstellenden Mediums häufiger zu regel-konformen und komplexen Schlussfolgerungen angeregt werden konnten, wenn ihnen die Informationen über Ursache-Wirkungs-Zusammenhänge in summarischer Form dargeboten wurden und wenn die Probanden über hohe Fähigkeiten im numerischen Denken verfügten. Offensichtlich ist eine hohe visuelle Strukturierung nur dann als Mittel zur Unterstützung bei Vergleichs- und Schlussfolgerungsprozessen sinnvoll, wenn Arbeitsgedächtnis und Informationsverarbeitung hinreichend entlastet sind.

In Experiment 3 sollte genauer die Fragestellung untersucht werden, warum Probanden bei summarischer Darbietung mehr regel-konforme Lösungen hervorbringen als bei

Einzelfalldarbietung. Wenn von der Annahme ausgegangen wird, dass das Arbeitsgedächtnis bei Einzelfalldarbietung zu stark beansprucht wird, dann müssten Probanden durch eine reduzierte Beanspruchung des Arbeitsgedächtnisses bei Einzelfalldarbietung in der Lage sein, mehr regel-konforme und komplexere Lösungen zu bilden. Um dieses Ziel zu erreichen, wurde im Vergleich zu Experiment 2 eine Variation der Aufgabe in zwei Versuchsbedingungen vorgenommen. Eine Möglichkeit, die mentale Beanspruchung bei Einzelfalldarbietung zu reduzieren, besteht darin, den Probanden über alle Einzelfalldarbietungen hinweg die Möglichkeit zu geben, sich Notizen zu machen. Durch diese Variation der Aufgabenstellung können die Probanden ihre internen Gedächtnisprozesse externalisieren und für weitere Denkopoperationen (wie z.B. zählen, gruppieren, vergleichen, berechnen, schlussfolgern etc.) verfügbar machen.

Shaklee & Mims (1982, Ex 1) haben ebenfalls eine aufgabenbezogene Variation vorgenommen und ihren Probanden in einer Bedingung die Möglichkeit gegeben, bei einzelfallweiser Darbietung Notizen anzufertigen und diese in eine vorgegebene leere Häufigkeitstabelle einzutragen. In den beiden anderen Bedingungen hatten die Probanden keine Möglichkeit, sich Notizen zu machen, sondern sollten aus dem Gedächtnis den kausalen Zusammenhang zwischen zwei Ereignissen beurteilen. Es zeigte sich, dass in der Bedingung mit Notizen komplexere und häufiger regel-basierte Strategien eingesetzt wurden als in den beiden Bedingungen ohne Notizen. Nach Shaklee & Mims (1982) sprechen diese Ergebnisse dafür, dass die Gedächtnisbeanspruchung einen klaren Einfluss auf die Genauigkeit bei der kausalen Urteilsbildung hat. Sie argumentieren dabei ähnlich wie Kleinmuntz & Schkade (1993), dass die einfachere Strategiewahl die stärkere Gedächtnisbeanspruchung in den beiden Bedingungen ohne Notizen ausgleicht.

Unsere Vermutung ist, dass externale Repräsentationen in Form von Notizen die Probanden durch eine reduzierte mentale Beanspruchung auch bei Einzelfalldarbietung darin unterstützen, von komplexen, regel-basierten Strategien Gebrauch zu machen. Offen ist aber dabei die Frage, in welchem Ausmaß sie das tun. Produzieren Probanden, deren mentale Beanspruchung bei Einzelfalldarbietung durch selbst angefertigte Notizen reduziert wird, ebenso komplexe und genau so viele regel-konforme Lösungen wie Probanden, denen ein Stapeldiagramm in summarischer Form dargeboten wird? Oder bringen Probanden, die ein summarisch dargebotenes Stapeldiagramm bearbeiten, häufiger regel-konforme Lösungen hervor, weil ihnen ein regel-konformes Zueinander-in-Beziehung-setzen von Informationen im Stapeldiagramm salienter gemacht wird als bei Einzelfalldarbietung und der Möglichkeit, Notizen anzufertigen.

Falls die Probanden bei Einzelfalldarbietung mit freiem Anfertigen von Notizen weniger regel-konforme Lösungen hervorbringen als bei summarischer Darbietung im Stapeldiagramm, dann hängt dies nicht mit einer reduzierten Gedächtnisbeanspruchung zusammen, sondern könnte darin begründet sein, dass die Informationen bei freien Notizen weniger salient als im Stapeldiagramm angeordnet sind. Eine Möglichkeit, die Salienz bei Einzelfalldarbietung mit Notizen zu erhöhen, besteht darin, die Probanden strukturiert anzuleiten, wie sie ihre Notizen übersichtlich in einer Vierfeldertafel anordnen können.

Strukturierte Notizen sollten im Gegensatz zum freien Anfertigen von Notizen die Salienz bei Einzelfalldarbietung erhöhen. Indem die Probanden angeleitet werden, wie sie die Notizen anzufertigen haben, wird die Aufgabenstellung für sie klarer, eingegrenzter und konkreter, sie müssen sich nicht mehr mit der Frage beschäftigen, welche Informationen sie beim Notizen erstellen wie aufbereiten müssen. Durch die Anleitung sollen die Notizen übersichtlicher gemacht werden. Für die Probanden wird es dadurch möglicherweise leichter, zu einer Schlussfolgerung darüber zu gelangen, wie die vier verschiedenen Häufigkeiten innerhalb einer Vierfeldertafel zueinander in Beziehung gesetzt und miteinander verglichen werden können, um die Wirksamkeit einer Substanz zu ermitteln.

In Experiment 3, welches parallel zu Experiment 2 durchgeführt wurde, sollte der Einfluss des Faktors *Notizen anfertigen* in seiner Funktion als Mittel zur Reduzierung der mentalen Beanspruchung auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen untersucht werden. Die zentrale Fragestellung für Experiment 3 war, ob eine Reduzierung der mentalen Beanspruchung durch freie oder strukturierte Notizen zu einer höheren Anzahl regel-konformer Lösungen führt als eine Einzelfalldarbietung ohne die Möglichkeit, Notizen anzufertigen. Weiter sollte die Frage untersucht werden, ob Probanden, die bei Einzelfalldarbietung durch das Anfertigen von Notizen darin unterstützt werden, einen summarischen Überblick über die Daten zu gewinnen, genauso viele regel-konforme Lösungen produzieren wie Probanden, denen kausale Informationen summarisch in einem Stapeldiagramm dargeboten wurden.

Wie für die Experimente 1 und 2 werden nachfolgend Untersuchungsdesign, Variablen und Forschungshypothesen aus Experiment 3 vorgestellt. Im Anschluss an die Vorstellung der teilnehmenden Probanden und Methoden, sowie der Untersuchungsdurchführung werden dann die zentralen Ergebnisse aus Experiment 3 näher beschrieben und diskutiert.

## 10.1 Untersuchungsdesign

### Versuchsplan und unabhängige Variablen

In Experiment 3 wurde der Einfluss des vierstufigen Faktors *Notizen anfertigen* untersucht. Dabei wurden die beiden Einzelfallbedingungen *Notizen frei* (kurz *N frei*) und *Notizen strukturiert* (kurz *N strukturiert*) untereinander und mit den beiden schon aus Experiment 2 bekannten Bedingungen "*visuelle Strukturierung hoch bei Einzelfalldarbietung*" (Vierfeldertafel – kurz *VT tbt*) und "*visuelle Strukturierung hoch bei summarischer Darbietung*" (Stapeldiagramm – kurz *SD sum*) verglichen.

Die Planung von Experiment 3 erfolgte zeitgleich mit der Planung von Experiment 2. Für alle drei Einzelfallbedingungen wurde wie für Bedingung *VT tbt* aus Experiment 2 eine Vierfeldertafel mit fallweise wechselndem Positionskreuz eingesetzt (siehe Abbildung 21 in der Beschreibung der unabhängigen Variablen zu Experiment 2). Durch die Position des Kreuzes auf der Vierfeldertafel wird angezeigt, welches Ereignis gerade eingetreten ist (links unten: geimpfte Wolke, die anfangt zu regnen; links oben: geimpfte Wolke, die nicht anfängt zu regnen; rechts unten: nicht geimpfte Wolke, die anfängt zu regnen; rechts oben: nicht geimpfte Wolke, die nicht anfängt zu regnen). Mit der Wahl eines Stapeldiagramms mit Häufigkeiten (siehe Abbildung 20 in der Beschreibung der unabhängigen Variablen zu Experiment 2) als summarischer Darbietungsform sollte wie in Experiment 2 die Vergleichbarkeit mit den Einzelfallbedingungen gewahrt bleiben.

Unter der Bedingung *N frei* wurde den Probanden mitgeteilt, dass sie während des Versuchs auf den beiliegenden leeren Blättern Notizen anfertigen können. Unter der Bedingung *N strukturiert* bekamen die Probanden auf einem separaten Blatt eine Anleitung, in der beschrieben wurde, wie sie die Notizen erstellen sollten. Die für den Versuchsablauf beigelegten Notizblätter unter dieser Bedingung waren nicht leer, sondern für jede zu beurteilende Substanz mit einer leeren Vierfeldertafel bedruckt, die während des Versuchs mit Strichen ausgefüllt werden sollte, wobei jeder Strich für ein Beobachtungsereignis stehen sollte. Abbildung 37 zeigt beispielhaft für Substanz 2 eine solche leere Vierfeldertafel, wie sie die Probanden unter der Bedingung *N strukturiert* in siebenfacher Ausfertigung vorgelegt bekamen (6 Testsubstanzen plus eine Probesubstanz):



**Notizblatt für Substanz 2**

Kein Regen		
Regen		

geimpfte Wolken
nicht geimpfte  
Wolken

**Abbildung 37:** Experiment 3: Leere Vierfeldertafel, wie sie die Probanden zum Anfertigen der Strichlisten unter der Bedingung *N strukturiert* vorgelegt bekamen.

Die Bedingung *VT tbt* aus Experiment 2 entsprach in Experiment 3 der Bedingung *ohne Notizen einzelfallweise* (kurz *ohne N tbt*). In dieser Bedingung wurde den Probanden explizit mitgeteilt, dass sie während des Versuchs keine Notizen anfertigen sollten. Die Bedingung *SD sum* aus Experiment 2 entsprach in Experiment 3 der Bedingung *ohne Notizen summarisch* (kurz *ohne N sum*), auch hier durften keine Notizen anfertigt werden.

Insgesamt wurde in Experiment 3 der Einfluss des vierstufigen Faktors *Notizen anfertigen* untersucht. Das zugrunde liegende 1 x 4 – Design ist in Tabelle 18 abgebildet:

**Tabelle 18:** Experiment 3: 1 x 4 – Design.

Faktor Notizen anfertigen			
ohne Notizen summarisch ( <i>ohne N sum</i> )	Notizen strukturiert einzelfallweise ( <i>N strukturiert</i> )	Notizen frei einzelfallweise ( <i>N frei</i> )	ohne Notizen einzelfallweise ( <i>ohne N tbt</i> )

Wie in Experiment 2 wurde als Kovariate das numerische Denken einschließlich der numerischen Verarbeitungskapazität eines jeden Probanden mit neun dafür ausgewählten Items aus dem Berliner Intelligenzstruktur-Test (BIS-Test, Jäger, Süß & Beauducel, 1997; Süß, 1996) mit erhoben.

### **Abhängige Variablen**

Es wurden dieselben abhängigen Variablen erhoben wie in Experiment 2.

## **10.2 Forschungshypothesen**

### **10.2.1 Einfluss des Faktors Notizen anfertigen**

#### *Einfluss von Notizen auf die Anzahl regel-konformer Lösungen*

Es wird erwartet, dass Personen, die die Möglichkeit haben, bei Einzelfalldarbietung Notizen anzufertigen, signifikant häufiger regel-konforme Lösungen bei der Ermittlung der Wirksamkeit einer Ursache produzieren als Personen, die keine Möglichkeit haben, während der Einzelfalldarbietung Notizen zu erstellen. Bei der Einzelfalldarbietung ohne Notizen sind die Probanden in einem nur sehr geringen Ausmaß dazu in der Lage, die vier verschiedenen Arten von Ursache-Wirkungs-Zusammenhängen zu ordnen, aufzuzaddieren und regel-konform miteinander zu verrechnen. Ohne die Anfertigung von Notizen fehlt den Probanden eine externale Repräsentation zur Unterstützung des Arbeitsgedächtnisses bei der Aufgabenbearbeitung. Dagegen können bei Einzelfalldarbietung mit Notizen interne Gedächtnisprozesse externalisiert werden können. Durch das Anfertigen der Notizen erstellen die Probanden eine externale Repräsentation, in welcher die Informationen über Ursache-Wirkungs-Zusammenhänge in summarischer Form zugänglich gemacht sind. Das Arbeitsgedächtnis wird dadurch entlastet und die Probanden haben mehr mentale Kapazität zur Verfügung, um komplexere, regel-konforme Lösungen zu generieren oder aus dem Langzeitgedächtnis abzurufen.

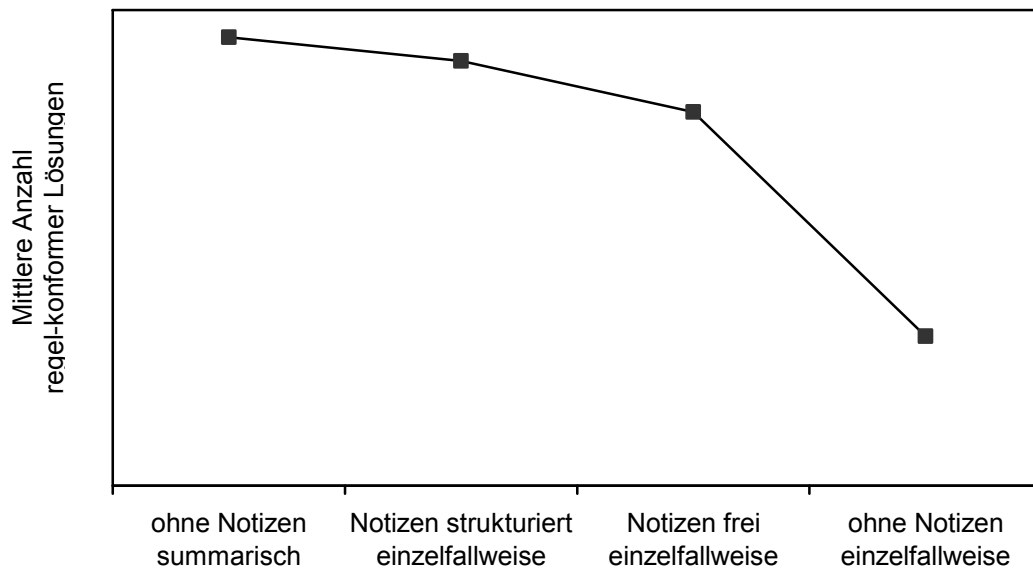
Weiter wird angenommen, dass die Probanden unter der Bedingung *N strukturiert* zwar mehr regel-konforme Lösungen produzieren als unter der Bedingung *N frei*, dass aber dieser Unterschied nicht signifikant wird. Unter der Bedingung *N strukturiert* wird nicht nur das Arbeitsgedächtnis entlastet, sondern es kommt zusätzlich zu einer Fokussierung der Aufmerksamkeit und zu einer Strukturierung der Informationsverarbeitung. Unter der Bedingung *N frei* wird das Gedächtnis der Probanden zwar entlastet, doch bleibt es ihnen dabei selbst überlassen, wie und in welcher Form die Notizen angefertigt werden. Anders als unter der Bedingung *N strukturiert* wird keine die Informationsverarbeitung strukturierende

Vierfeldertafel dargeboten, in die die einzelnen Striche in geordneter Form eingetragen werden können. Auf der Basis der Ergebnisse aus Experiment 1 (Kontrast zwischen tabellarischer Vierfeldertafel mit Häufigkeiten und der Kontrollbedingung Protokollliste wurde nicht signifikant) wird nicht angenommen, dass die Unterschiede zwischen den beiden Bedingungen mit Notizen signifikant werden.

Schließlich wird vermutet, dass unter der Bedingung *ohne N sum* zwar mehr regel-konforme Lösungen produziert werden als unter den drei anderen Bedingungen, dass aber nur der Unterschied zwischen den beiden Bedingungen *ohne N sum* und *ohne N tbt* signifikant wird. Bei summarischer Informationsdarbietung im Stapeldiagramm werden den Probanden regel-konforme Schlussfolgerungen und Vergleichsprozesse nahe gelegt und sie können ihre Aufmerksamkeit ganz auf das richtige Zueinander-in-Beziehung-setzen und Verrechnen der kausalen Informationen konzentrieren. Anders als in Bedingung *ohne N tbt* kann das Gedächtnis entlastet werden; anders als in den beiden Bedingungen *N strukturiert* und *N frei*, in denen den Probanden zumindest unter der Bedingung *N frei* das richtige Zueinander-in-Beziehung-setzen und Verrechnen der kausalen Informationen nicht besonders nahe gelegt wird, geht keine Aufmerksamkeit durch das Führen von Strichlisten oder –tabellen verloren. Die Annahme, dass der Unterschied zwischen der Bedingung *ohne N sum* und den beiden Bedingungen mit Notizen nicht signifikant wird, beruht darauf, dass auch die Probanden unter den beiden Bedingungen *N strukturiert* und *N frei* bei der Wirksamkeitsbeurteilung auf externale Repräsentationen zurückgreifen können, unter der Bedingung *N strukturiert* sogar auf eine Vierfeldertafel mit Häufigkeiten. In Kurzschreibweise zusammengefasst ergibt sich folgende Hypothese:

H9a: # regel-konformer Lösungen *ohne N sum* / *N strukturiert* / *N frei* > # regel-konformer Lösungen *ohne N tbt*

In Abbildung 38 wird der vorhergesagte Einfluss des Faktors *Notizen anfertigen* auf die *Anzahl regel-konformer Lösungen* grafisch veranschaulicht:



**Abbildung 38:** Experiment 3: Vorhergesagter Einfluss von *Notizen anfertigen* auf die abhängige Variable *Anzahl regel-konformer Lösungen*.

#### *Einfluss von Notizen auf die Bearbeitungszeit*

Es wird erwartet, dass die mittlere Gesamtbearbeitungszeit bei Personen, die keine Notizen anfertigen müssen, geringer ausfällt als bei Personen, die Notizen anfertigen müssen. Durch das Anfertigen von Notizen bei Einzelfalldarbietung wird zwar das Gedächtnis im Vergleich zur Bedingung *ohne N tbt* entlastet, doch durch das Anfertigen der Strichlisten entsteht für die Probanden ein zusätzlicher Bearbeitungsaufwand. Weiter wird wie schon in Experiment 2 erwartet, dass die Probanden unter der Bedingung *ohne N sum* signifikant schneller arbeiten als unter der Bedingung *ohne N tbt*. Zwischen den beiden Bedingungen mit Notizen, *N strukturiert* und *N frei*, werden keine signifikanten Unterschiede erwartet.

$$H9b: \text{Gesamtbearbeitungszeit}_{\text{ohne N sum}} < \text{Gesamtbearbeitungszeit}_{\text{ohne N tbt}} < \text{Gesamtbearbeitungszeit}_{\text{N frei / N strukturiert}}$$

#### *Einfluss von Notizen auf die Verarbeitungseffizienz*

Die Verarbeitungseffizienz pro Person wird aus den beiden Quotienten *Anzahl regel-konformer Lösungen / Gesamtbearbeitungszeit* ermittelt. An Hand der Hypothesen H9a und H9b wird unter der Bedingung *ohne N sum* eine höhere Verarbeitungseffizienz erwartet als unter den drei Bedingungen mit Einzelfalldarbietung. Im Zusammenhang mit den drei Einzelfallbedingungen vermuten wir einen *speed-accuracy trade off*: Unter der Bedingung *ohne N tbt* arbeiten die Probanden zwar schneller, aber weniger genau als unter den beiden Einzelfallbedingungen *N strukturiert* und *N frei*. Die größere "Genauigkeit" in den beiden

Einzelfallbedingungen mit Notizen resultiert aus einer größeren Anzahl regel-konformer Lösungen infolge des Anfertigen von Notizen.

H9c: Verarbeitungseffizienz<sub>ohne N sum</sub> > Verarbeitungseffizienz<sub>ohne N tbt / N frei / N strukturiert</sub>

### *Einfluss von Notizen auf die Strategiewahl*

Es wird erwartet, dass die Zahl der Probanden, die entweder von der  $\Delta P$ -Regel oder der Power PC Theorie Gebrauch machen, unter der Bedingung *ohne N sum* sowie unter den beiden Einzelfallbedingungen *N strukturiert* und *N frei* höher ist als unter der Bedingung *ohne N tbt*. Es wird weiter vermutet, dass auch zwischen den drei Bedingungen *ohne N sum*, *N strukturiert* und *N frei* Häufigkeitsunterschiede auftreten, die aber nicht signifikant werden. Demnach werden unter der Bedingung *ohne N sum* mehr Probanden von regel-basierten Strategien Gebrauch machen als unter der Bedingung *N strukturiert*, in der die Anzahl der Probanden, die regel-basierte Strategien einsetzen, wiederum höher liegen wird als unter der Bedingung *N frei*. Die Begründung für den Einfluss der unabhängigen Variable *Notizen anfertigen* auf die Strategiewahl entspricht derjenigen für den Einfluss von *Notizen anfertigen* auf die abhängige Variable *Anzahl regel-konformer Lösungen*.

H9d: # Pbn mit  $\Delta P$  oder PPC<sub>ohne N sum / N strukturiert / N frei</sub> > # Pbn mit  $\Delta P$  oder PPC<sub>ohne N tbt</sub>

## 10.2.2 Einfluss der Kovariate numerisches Denken

Wie schon in Experiment 2 wird ein signifikanter Einfluss der Kovariate *numerisches Denken* auf alle abhängigen Variablen außer der *Bearbeitungszeit* erwartet. Damit können die für Experiment 2 abgeleiteten Hypothesen H7a und H7b hier wiederholt werden.

H7a: # regel-konformer Lösungen<sub>ND hoch</sub> > # regel-konformer Lösungen<sub>ND gering</sub>

H7b: Verarbeitungseffizienz<sub>ND hoch</sub> > Verarbeitungseffizienz<sub>ND gering</sub>

## 10.2.3 Einfluss des Faktors Studienausrichtung

Für Experiment 3 gelten die schon aus Experiment 2 bekannten Hypothesen H8a und H8b. Es wird wie in Experiment 2 erwartet, dass der Faktor *Studienausrichtung* bei Nichtberücksichtigung der Kovariate *numerisches Denken* einen signifikanten Einfluss auf die beiden abhängigen Variablen *Anzahl regel-konformer Lösungen* und *Verarbeitungseffizienz* haben wird. Wird die Kovariate *numerisches Denken* hingegen mitberücksichtigt, dann ist zu erwarten, dass der Faktor *Studienausrichtung* nicht signifikant

wird. Für die *Bearbeitungszeit* wird kein signifikanter Effekt in Abhängigkeit von *Studienausrichtung* vorhergesagt.

H8a: # regel-konformer Lösungen<sub>NW</sub> > # regel-konformer Lösungen<sub>GW</sub>

H8b: Verarbeitungseffizienz<sub>NW</sub> > Verarbeitungseffizienz<sub>GW</sub>

Im Zusammenhang mit dem Einfluss des Faktors *Studienausrichtung* auf die Strategiewahl gehen wir im Gegensatz zu Experiment 2 von der Nullhypothese aus, da dieser Einfluss in Experiment 2 nicht signifikant wurde. In Experiment 2 wurde nur für die Probanden, die durchgängig von regel-basierten Strategien Gebrauch gemacht haben, der Trend erzielt, dass die Anzahl der Probanden, die von regel-basierten Strategien Gebrauch macht, in der Probandengruppe mit mathematisch-naturwissenschaftlicher Studienausrichtung höher ist als in der Probandengruppe, die nicht mathematisch-naturwissenschaftlich ausgerichtet studiert.

### 10.3 Methodik und Vorgehensweise

#### Probanden

An Experiment 3 nahmen 80 Versuchspersonen (52 Frauen und 28 Männer) teil. Das Alter variierte zwischen 19 und 41 Jahren ( $M = 24.05$ ;  $s = 4.044$ ). Voraussetzung für die Teilnahme war wiederum eine gute Beherrschung der deutschen Sprache. Alle Teilnehmer waren Studierende der Universität Tübingen oder angehende Studenten am Leibnizkolleg in Tübingen (72 Studierende, 8 Leibniz-Kollegiaten). Die Ergebnisdaten für die Bedingungen *ohne N sum* und *ohne N tbt* wurden schon in Experiment 2 unter den Bedingungen *SD sum* und *VT tbt* erhoben. Wie schon in den Experiment 1 und 2 handelte es sich bei der einen Hälfte um Teilnehmer, die einen Studiengang absolvierten oder absolvieren wollten, in denen mathematische Fächer (wie z.B. Statistik) oder naturwissenschaftliche Methoden gelehrt wurden. Die andere Hälfte der Probanden setzte sich aus Studierenden oder angehenden Studierenden geisteswissenschaftlicher Disziplinen zusammen, in denen mathematische oder naturwissenschaftliche Methoden kaum oder gar nicht vermittelt wurden. Über die Zusammensetzung der Probanden hinsichtlich ihrer Studenausrichtung und sonstiger demografischer Merkmale gibt Tabelle A5 im Anhang A3.1 Aufschluss. Die Versuchsteilnehmer wurden extern durch Aushänge oder Ansprache angeworben und erhielten für die Teilnahme € 16,-. Die Probanden wurden einzeln untersucht und gleichmäßig nach dem Zufallsprinzip auf die vier Experimentalbedingungen verteilt.

#### Material

Wie in den Experimenten 1 und 2 hatten alle Probanden die Aufgabe zu beurteilen, ob und in welchem Ausmaß ein Ereignis A (die Ursache) ein Ereignis B (die Wirkung) hervorruft. Eine oder mehrere konkurrierende Ursachen lagen nicht vor. Wie in den beiden Experimenten zuvor kam das Wolkenimpfparadigma zum Einsatz; entsprechend dem Vorgehen in Experiment 2 wurden die Probanden hinsichtlich ihrer Fähigkeit im numerischen Denken untersucht. Informationen über Ursache-Wirkungs-Zusammenhänge wurden wie in Experiment 2 computergesteuert dargeboten. Wieder wurde die Versuchsumgebung mit Hilfe der Software Marcomedia-Director erstellt. Eingabeinhalte und zeitlicher Ablauf des Hauptversuchs wurden erneut in Log-Dateien festgehalten.

#### Untersuchungsdurchführung

Die Untersuchungsdurchführung war in den Experimenten 2 und 3 identisch. Dabei waren die Instruktionen für die beiden Bedingungen *ohne N sum* und *ohne N tbt* identisch mit den

Instruktionen für die beiden Bedingungen *SD sum* und *VT tbt* aus Experiment 2 (s. Anhang A2.2), d.h. sie enthielten den expliziten Hinweis, dass während des Versuchs keine Notizen gemacht werden sollten. Die Probanden in den beiden Bedingungen *N strukturiert* und *N frei* erhielten dieselben Instruktionen wie die Probanden unter der Bedingung *ohne N tbt*. Der einzige Unterschied bestand darin, dass die Probanden in den beiden Bedingungen mit Notizen darauf hingewiesen wurden, dass sie während des Versuchs auf den beiliegenden Blättern Notizen anfertigen sollten. Die Probanden unter der Bedingung *N strukturiert* erhielten zusätzlich den Hinweis, dass sich auf Seite drei der Instruktion eine Anleitung zum Anfertigen von Notizen befindet. Die Instruktionen zu den Bedingungen *N frei* und *N strukturiert* können im Anhang A3.2 eingesehen werden.

#### 10.4 Explorative Datenanalysen

Die statistischen Analysen wurden erneut mit dem Programmpaket SPSS durchgeführt. Die Voraussetzung der Normalverteilung wurde mit Hilfe des Lilliefors-Tests und des Shapiro-Wilks-Tests, die Voraussetzung der Varianzhomogenität wurde mit dem Levene-Test überprüft. Hinsichtlich der Voraussetzung der Normalverteilung wurden für die abhängige Variable *Anzahl regel-konformer Lösungen* unter allen Bedingungen Verstöße festgestellt, die je nach Bedingung auf Boden- oder Deckeneffekte zurückzuführen sind. Bei der abhängigen Variable *Bearbeitungszeit* wurde die Voraussetzung der Normalverteilung nur bei den Bedingungen *ohne N sum* und *ohne N tbt* verletzt, sonst nicht. Offensichtlich führte also das Notizen anfertigen dazu, dass sich die Zeitwerte gleichmäßiger verteilten. Die Ergebnisse des Levene-Tests weisen auf heterogene Fehlervarianzen aller abhängigen Variablen bis auf die *Bearbeitungszeit* hin. Die fehlende Homogenität der Fehlervarianzen ist vermutlich auf unterschiedliche Einflüsse der Versuchsbedingungen auf die abhängigen Variablen zurückzuführen. Trotz der Verstöße haben wir wegen der Robustheit der Varianzanalyse gegenüber Verletzungen ihrer Voraussetzungen Varianzanalysen durchgeführt, da ausreichende Zellgrößen ( $n = 20$ ) und gleichgroße Stichproben vorlagen.

Die Werte für die Kovariate *numerisches Denken* verteilten sich in allen vier Versuchsbedingungen normal, allerdings waren die Fehlervarianzen inhomogen. Weitere Analysen ergaben, dass sich die vier Mittelwerte im *numerischen Denken* in Abhängigkeit vom Faktor *Notizen anfertigen* nicht signifikant voneinander unterschieden;  $F_{(3, 76)} = 1.341$ ,  $MSE = 38.229$ ;  $p = .267$ .



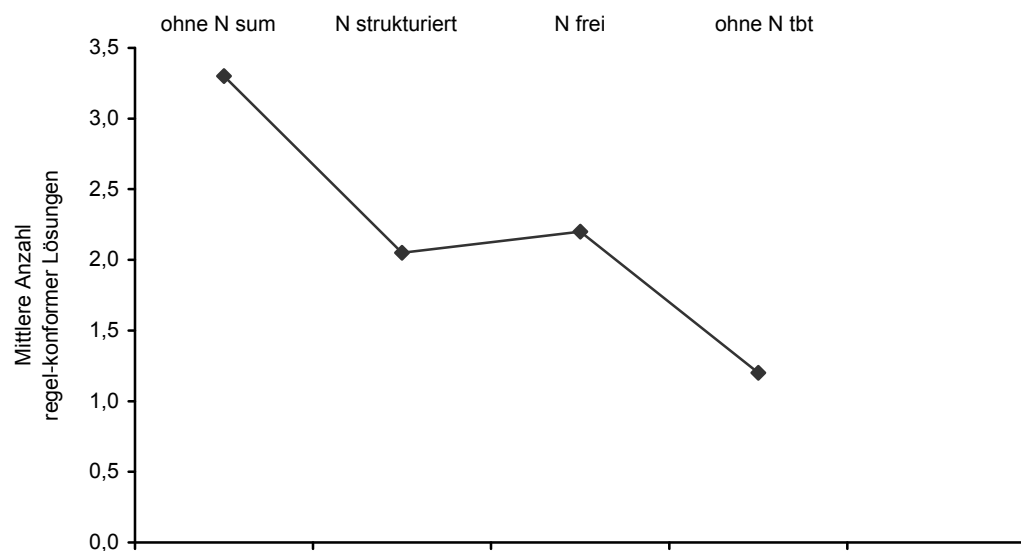
## 10.5 Einfluss des Faktors Notizen anfertigen

Wegen der Komplexität der Diskussionen in Verbindung mit dem Einfluss des Faktors *Notizen anfertigen* auf die abhängigen Variablen, werden Ergebnisse und Diskussion unter die jeweilige abhängige Variable subsumiert.

### 10.5.1 Anzahl regel-konformer Lösungen

#### 10.5.1.1 Ergebnisse

Abbildung 39 zeigt die Maße der zentralen Tendenz für die *Anzahl regel-konformer Lösungen* in Abhängigkeit vom Faktor *Notizen anfertigen*.



	ohne N sum	N strukturiert	N frei	ohne N tbt	gesamt
<b>Anzahl regel-konformer Lösungen</b>					
arithmetisches Mittel	3.3 (2.6)	2.050 (2.35)	2.2 (2.285)	1.2 (1.47)	2.188 (2.30)
Median	4	1	1	1	1
Modalwert	6	1	0	1	0

**Abbildung 39:** Experiment 3: Mittlere Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) in Abhängigkeit von *Notizen anfertigen*.

Die ermittelten Maße für die zentrale Tendenz weisen darauf hin, dass die höchsten Werte von den Probanden unter der Bedingung *ohne N sum*, die niedrigsten unter der Bedingung *ohne N tbt* erreicht werden. Median und Modalwert vergrößern die Kluft zwischen diesen beiden Bedingungen. Die beiden Bedingungen mit Notizen nehmen eine Mittelstellung zwischen den beiden Bedingungen ohne Notizen ein. Für die *Anzahl regel-konformer Lösungen* wurde eine Kovarianzanalyse mit dem Faktor *Notizen anfertigen* und der Kovariate *numerisches Denken* durchgeführt. Tabelle 19 zeigt die Ergebnisse.

**Tabelle 19:** Experiment 3: Ergebnisse aus der univariaten Kovarianzanalyse mit dem Faktor *Notizen anfertigen* für die abhängige Variable *Anzahl regel-konformer Lösungen*. Signifikante P-Werte sind **fett** gedruckt.

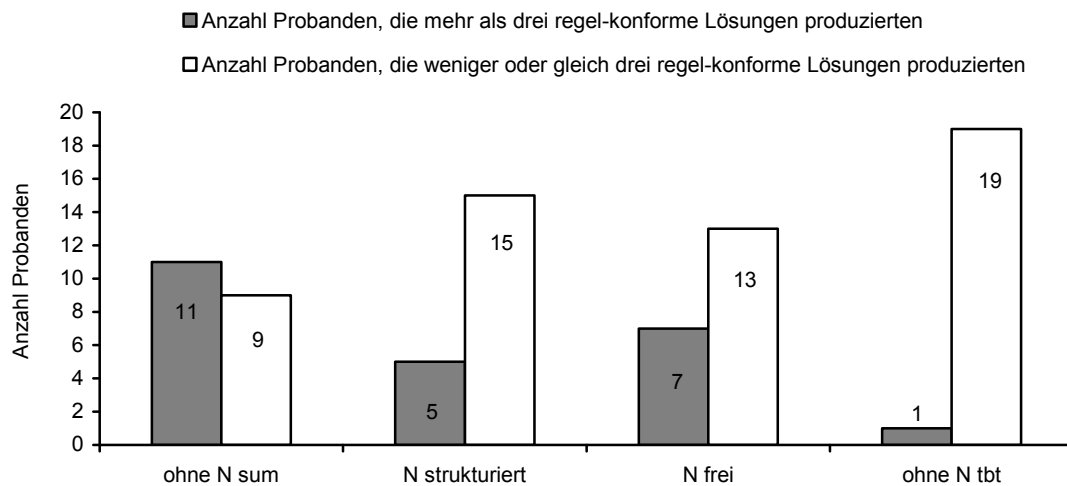
Faktor	F-Wert	MSE	P-Wert
Notizen anfertigen <sub>(3,75)</sub>	2.915	4.379	<b>.040</b>
numerisches Denken <sub>(3,75)</sub>	10.312		<b>&lt; .002</b>

Die Kovarianzanalyse ergab einen signifikanten Haupteffekt von *Notizen anfertigen* auf die *Anzahl regel-konformer Lösungen*. Signifikant wurde auch der Einfluss der Kovariate *numerisches Denken*. Mit den zusätzlich durchgeführten paarweisen Vergleichen zwischen Bonferroni-adjustierten Gruppenmittelwerten, die in Tabelle 20 aufgezeigt sind, konnte gezeigt werden, dass lediglich der Kontrast zwischen den beiden Bedingungen *ohne N sum* und *ohne N tbt* signifikant wurde ( $p = .036$ ). Dagegen wurden die Bonferroni-Kontraste *N strukturiert* und *N frei* nicht signifikant (jeweils  $p = 1.000$ ).

**Tabelle 20:** Experiment 3: Paarweise Vergleiche zwischen den Bonferroni-adjustierten Gruppenmittelwerten für die abhängige Variable *Anzahl regel-konformer Lösungen* in Abhängigkeit von *Notizen anfertigen*. Signifikante P-Werte ( $\alpha < .05$ ) sind **fett** markiert.

	ohne N sum	N strukturiert	N frei	ohne N tbt
ohne N sum	-	.355	.261	<b>.036</b>
N strukturiert		-	1.000	1.000
N frei			-	1.000
ohne N tbt				-

Wie in Experiment 2 wurden im Hinblick auf die abhängige Variable *Anzahl regel-konformer Lösungen* die Voraussetzungen der Varianzanalyse verletzt, außerdem fielen Median und Modalwert unter der Bedingung *ohne N sum* deutlich höher aus als unter den anderen drei Bedingungen. Deshalb wurden erneut alle Probanden dahingehend dichotomisiert, ob sie mehr als drei regel-konforme Lösungen produzierten oder nicht, und daraufhin untersucht, ob sie sich in Abhängigkeit von den vier Versuchsbedingungen unterschiedlich verteilten. Abbildung 40 veranschaulicht, wie viele Probanden pro Versuchsbedingung mehr als drei oder weniger als vier regel-konforme Lösungen entwickelt haben:



**Abbildung 40:** Experiment 3: Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen in Abhängigkeit von *Notizen anfertigen*.

Ein  $\chi^2$  - Test ohne Berücksichtigung der Kovariate *numerisches Denken* führte zu dem hochsignifikanten Ergebnis, dass die Verteilung von Probanden, die mehr als drei regel-konforme Lösungen produzierten oder nicht, in Abhängigkeit von *Notizen anfertigen* nicht zufällig ist ( $\chi^2 = 12.381$ ;  $p = .006$ ; Phi-Wert = .393). Weitere Häufigkeitsanalysen, in denen analog zur Berechnung paarweiser Einzelkontraste bei Varianzanalysen untersucht wurde, ob die Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen von je zwei Versuchsbedingungen unabhängig ist, lassen darauf schließen, dass die beiden Häufigkeitsvergleiche *ohne N sum* versus *ohne N tbt* und *N frei* versus *ohne N tbt* signifikant sind, während für die Paarvergleiche *ohne N sum* versus *N strukturiert* und *N strukturiert* versus *ohne N tbt* lediglich Trends erzielt wurden. Die Vergleiche zwischen den beiden Bedingungen *ohne N sum* und *N frei*, sowie zwischen den beiden Bedingungen mit Notizen verfehlten das Signifikanzniveau von  $\alpha = 5\%$ . Allerdings ist auch hier die Interpretierbarkeit der Vergleiche ohne die Bedingung *ohne N sum* eingeschränkt, da der Anteil an Zellen mit erwarteten Häufigkeiten  $< \text{fünf } 50\%$  betrug. Tabelle 21 zeigt diese Ergebnisse im Überblick.

**Tabelle 21:** Experiment 3: Ergebnisse aus paarweisen Häufigkeitsvergleichen, in denen untersucht wurde, ob die Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen von je zwei Versuchsbedingungen unabhängig ist oder nicht. Signifikante P-Werte sind **fett** markiert, Trends unterstrichen.

Vergleich	$\chi^2$ -Wert	df	Asymptotische Signifikanz (2-seitig)	exakte Signifikanz (exakter Test nach Fisher)
ohne N sum – N strukturiert	3.75	1	<u>.053</u>	<u>.053</u>
ohne N sum – N frei	1.616	1	.204	.170
ohne N sum – ohne N tbt	11.905	1	< <b>.001</b>	< <b>.001</b>
N strukturiert – N frei	.476	1	.490	.366
N strukturiert – ohne N tbt	3.137	1	<u>.077</u>	<u>.091</u>
N frei – ohne N tbt	5.625	1	<b>.018</b>	<b>.022</b>

Im Vergleich zu den paarweisen Vergleichen zwischen Bonferroni-adjustierten Gruppenmittelwerten fallen die Unterschiede zwischen den einzelnen Bedingungen in den Häufigkeitsvergleichen deutlicher aus, beispielsweise befinden sich unter der Bedingung *ohne N tbt* signifikant oder tendenziell weniger Probanden, die mehr als drei regel-konforme Lösungen produzieren als in den drei anderen Bedingungen. Dies liegt im Wesentlichen daran, dass bei einer Dichotomisierung von Probanden in solche mit mehr als drei produzierten regel-konformen Lösungen und solche mit einer Anzahl  $\leq 3$  regel-konformer Lösungen Streuungen nicht ins Gewicht fallen. Das im Vergleich zu den Kovarianzanalysen bessere Abschneiden von Probanden unter der Bedingung *N frei* lässt sich vermutlich dadurch erklären, dass die Probanden unter dieser Bedingung etwas höhere Werte im *numerischen Denken* erzielen ( $MW_{N\text{ frei}} = 102.33$ ) als die Probanden aus den anderen drei Bedingungen, insbesondere der Probanden unter der Bedingung *ohne N tbt* ( $MW_{\text{ohne N tbt}} = 98.42$ ). Bestätigt wird diese Vermutung dadurch, dass die 24 Probanden mit mehr als drei regel-konformen Lösungen in einem T-Test signifikant höhere Werte im *numerischen Denken* aufwiesen als die 56 Probanden, die weniger regel-konforme Lösungen produzierten, ( $MW_{>3\text{ regel-konforme Lösungen}} = 103.48$ ;  $MW_{\leq 3\text{ regel-konforme Lösungen}} = 98.95$ ;  $T_{78} = 3.152$ ,  $p = .002$ ).

Insgesamt kann die Hypothese H9a nur partiell beibehalten werden, da nur der Bonferroni-Kontrast *ohne N sum* versus *ohne N tbt* signifikant wurden, während bei den Bonferroni-Kontrasten *N strukturiert* versus *ohne N tbt* ( $p = 1.000$ ) und *N frei* und *ohne N tbt* ( $p = 1.000$ ) das Signifikanzniveau von  $\alpha = 5\%$  klar verfehlt wurde. Die Häufigkeitsanalysen, in denen die Probanden hinsichtlich der Anzahl regel-konformer Lösungen ( $>$  oder  $\leq 3$  regel-konforme Lösungen) zweigeteilt wurden und in Abhängigkeit von der Versuchsbedingung untersucht wurden, untermauern und verdeutlichen die Ergebnisse aus den Einzelkontrasten.

Offensichtlich wirkt sich eine durch die Möglichkeit des Anfertigen von Notizen reduzierte mentale Beanspruchung bei Einzelfalldarbietung schwächer aus als erwartet, der Unterschied zwischen den beiden Bedingungen *ohne N sum* und *ohne N tbt* kann also allenfalls teilweise auf eine Entlastung des Arbeitsgedächtnisses durch das Anfertigen von Notizen zurückgeführt werden. Kein nennenswerter Unterschied konnte zwischen den Mittelwerten in Abhängigkeit von den beiden Bedingungen *N strukturiert* und *N frei* gefunden werden.

Wie schon in Experiment 2, so kann auch in Experiment 3 die Hypothese H7a klar bestätigt werden. Probanden mit höheren Fertigkeiten im *numerischen Denken* produzieren eine höhere Anzahl regel-konformer Lösungen. Die höchsten Werte im *numerischen Denken* wurden von den Probanden unter der Bedingung *N frei* (MW = 102.33) erzielt, gefolgt von den Probanden unter den Bedingungen *N strukturiert* (MW = 100.31), *ohne N sum* (MW = 100.17) und *ohne N tbt* (MW = 98.42).

#### 10.5.1.2 Diskussion

In Experiment 3 sollte versucht werden, eine Antwort auf die Frage zu finden, warum Probanden bei summarischer Darbietung mehr regel-konforme Lösungen hervorbringen als bei Einzelfalldarbietung. Die Ergebnisse legen nahe, dass die in den Experimenten 2 und 3 zu beobachtenden Unterschiede zwischen einzelfallweiser und summarischer Darbietung nur teilweise auf eine Entlastung des Arbeitsgedächtnisses zurückgeführt werden können, da sonst in Experiment 3 die Unterschiede zwischen den Bedingungen mit Notizen und der Einzelfallbedingung ohne Notizen größer, und die Unterschiede zwischen der summarischen Bedingung ohne Notizen und den beiden Bedingungen mit Notizen geringer ausgefallen sein müssten. Die Annahme einer Entlastung des Arbeitsgedächtnisses ist zwar wesentlich, aber nicht hinreichend und kann nur den nicht signifikanten Anstieg der Anzahl regel-konformer Lösungen zwischen der Einzelfalldarbietung ohne Notizen und den beiden Einzelfalldarbietungen mit Notizen erklären. Anders formuliert ist es in Experiment 3 bei Einzelfalldarbietung trotz einer Entlastung des Arbeitsgedächtnisses durch das Anfertigen von Notizen nicht gelungen, dass die Probanden auch bei Einzelfalldarbietung eine (fast) genau so hohe Anzahl regel-konformer Lösungen produzieren wie bei summarischer Informationsdarbietung. Aus unserer Sicht müssen – anders als bei Shaklee & Mims (1982) - zusätzliche Interpretationsmöglichkeiten in Betracht gezogen werden, um den nicht signifikanten Anstieg der Anzahl regel-konformer Lösungen zwischen den beiden Einzelfalldarbietungen mit Notizen und der summarischen Darbietung im Stapeldiagramm erklären zu können.

Eine erste Erklärungsmöglichkeit geht von der Annahme aus, dass das summarisch dargebotene Stapeldiagramm im Vergleich zu den freien oder strukturierten Notizen bei einzelfallweiser Darbietung eine höhere visuelle Strukturierung aufweist. Diese Interpretationsmöglichkeit bezieht ihr theoretisches Fundament aus den Arbeiten von Zhang und anderen zu den unterschiedlichen Auswirkungen von externalen Repräsentationen (Larkin & Simon, 1987; Stenning & Oberlander, 1994; Zhang, 2000, 1997; Zhang, Johnson, Malin & Smith, 2002; Zhang & Norman, 1994). Danach würde sich das Stapeldiagramm mit summarischer Informationsdarbietung gegenüber den beiden Bedingungen mit Notizen dadurch auszeichnen, dass durch die im Diagramm enthaltenen Strukturierungsangebote und visuellen Hinweise (wie z.B. Farbgebung, Anordnung und Unterteilung der Säulen, Informationsgruppierung) ein regel-konformes Zueinander-in-Beziehung-setzen von Erfolgs- und Basisrate salienter gemacht werden kann als durch ein strukturiertes oder freies Anfertigen von Notizen bei einzelfallweiser Darbietung.

Auch wenn diese Interpretation auf der Basis der Ergebnisse aus Experiment 2 plausibel erscheint, so lassen die bisherigen Ergebnissen aus den Experimenten 1, 2 und 3 auch Zweifel an der Annahme berechtigt erscheinen, dass sich die beobachteten Unterschiede zwischen den Einzelfalldarbietungen mit Notizen und der summarischen Darbietungsform auf Unterschiede der verwendeten Präsentationsformen hinsichtlich ihrer visuellen Struktur und der in ihnen enthaltenen Strukturierungsangebote zurückführen lassen. So konnten in Experiment 1 hinsichtlich der Anzahl regel-konformer Lösungen keine unterschiedlichen Auswirkungen der Präsentationsform einer tabellarischen Vierfeldertafel, die ähnlich aufgebaut war wie die tabellarische Vierfeldertafel in der Bedingung *N strukturiert*, und der Präsentationsform eines Stapeldiagramms festgestellt werden. Dies spricht gegen die Annahme, dass sich das Stapeldiagramm in der Bedingung *ohne N sum* und eine mit Notizen vervollständigte tabellarische Vierfeldertafel in der Bedingung *N strukturiert* hinlänglich hinsichtlich ihrer visuellen Struktur bzw. der Qualität der Informationsanordnung unterscheiden. Auch das Ergebnis aus Experiment 3, dass die Probanden, die ihre Notizen in der übersichtlich strukturierten Vierfeldertafel zusammenfassen konnten, geringfügig weniger regel-konforme Lösungen gebildet haben als die Probanden, die ihre Notizen frei von Strukturvorgaben anfertigen konnten, spricht in unseren Augen eher dafür, dass der Faktor *visuelle Strukturierung* keinen Einfluss auf die Anzahl regel-konformer Lösungen in den beiden Einzelfallbedingungen mit Notizen hatte. Wäre der Anstieg der Anzahl regel-konformer Lösungen zwischen der Bedingung *ohne N sum* und den beiden Einzelfallbedingungen mit Notizen ausschließlich auf eine höhere visuelle Struktur oder ein besseres

Strukturierungsangebot der summarischen Präsentationsform zurückzuführen, so hätten wir damit gerechnet, dass auch unter der Bedingung *N strukturiert* mehr regel-konforme Lösungen gebildet werden als unter der Bedingung *N frei*. Zudem haben die Probanden unter der Bedingung *N strukturiert* aus Experiment 3 und der Bedingung *FT sum* aus Experiment 2 (visuell gering strukturierter Fliesstext mit summarischer Darbietung) eine fast identische mittlere Anzahl regel-konformer Lösungen produziert, wie ein Vergleich von Abbildung 25 mit Abbildung 39 zeigt.

Insgesamt sprechen diese Ergebnisse nicht dafür, dass die in Experiment 3 zwischen der Bedingung *ohne N sum* und den beiden Bedingungen mit Notizen bei einzelfallweiser Informationsdarbietung dadurch zu Stande gekommen wären, dass das summarisch dargebotene Stapeldiagramm im Vergleich zu den freien oder strukturierten Notizen in einer tabellarischen Vierfeldertafel eine höhere visuelle Strukturierung aufweist.

Zutreffender scheint es uns zu sein, von der schon in Experiment 2 formulierten Vermutung auszugehen, dass der Faktor *visuelle Strukturierung des darstellenden Mediums* bei einzelfallweiser Darbietung keinen Einfluss auf die Anzahl regel-konformer Lösungen hat, weil die Probanden bei einzelfallweiser Darbietung nicht dazu in der Lage sind, von einer übersichtlichen Informationsanordnung zu profitieren. In Experiment 2 haben wir gesehen, dass sich die beiden Einzelfallbedingungen *VT tbt* und *FT tbt* nicht hinsichtlich der Anzahl regel-konformer Lösungen unterscheiden, obwohl den Probanden unter der Bedingung *VT tbt* eine visuell hoch strukturierte Vierfeldertafel vorgelegt wurde und unter der Bedingung *FT tbt* nur ein einfacher Fliesstext. Dieser nur bei einer einzelfallweisen Darbietung nachgewiesene fehlende Einfluss des Faktors *visuelle Strukturierung des darstellenden Mediums* auf die Anzahl regel-konformer Lösungen wurde von uns auf eine hohe mentale Beanspruchung der Probanden in den Einzelfallbedingungen zurückgeführt.

Die Erklärung, dass eine hohe mentale Beanspruchung in den Einzelfallbedingungen den Einfluss einer hoch strukturierten und übersichtlichen Informationsanordnung verhindert, kann nicht uneingeschränkt für Experiment 3 übernommen werden, denn schließlich konnte ja das Arbeitsgedächtnis in den beiden Bedingungen *N strukturiert* und *N frei* durch das Anfertigen von Notizen und das Erstellen einer zumindest in der Bedingung *N strukturiert* übersichtlichen summarischen Ergebniszusammenfassung entlastet werden, wie der nicht signifikante Anstieg der Anzahl regel-konformer Lösungen zwischen der Einzelfallbedingung ohne Notizen und den beiden Einzelfallbedingungen mit Notizen vermuten lässt.

Auch Ward & Jenkins (1965; s.a. Wassermann & Shaklee, 1984) kamen in ihrer bereits mehrfach erwähnten Untersuchung zu dem Ergebnis, dass eine Gedächtnisentlastung durch

eine tabellarische Ergebniszusammenfassung, die den Probanden im Anschluss an eine vollständige einzelfallweise Darbietung präsentiert wurde, nicht dazu führte, dass unter dieser Bedingung die Probanden genau so häufig von der  $\Delta P$ -Regel Gebrauch machten wie bei einer ausschließlichen summarischen Darbietung. Stattdessen fielen die Ergebnisse von Ward & Jenkins (1965) ähnlich aus wie die Ergebnisse in Experiment 3; die kombinierte Bedingung (einzelfallweise Darbietung plus nachfolgende Ergebniszusammenfassung) nahm hinsichtlich der Anzahl an Probanden, die von der  $\Delta P$ -Regel Gebrauch machten, eine Zwischenstellung zwischen einer ausschließlich einzelfallweisen und der ausschließlich summarischen Ergebnisdarbietung ein. Während der Unterschied zwischen den Bedingungen mit kombinierter (# Pbn mit  $\Delta P$ : 9 von 24) und einzelfallweiser Darbietung (# Pbn mit  $\Delta P$ : 4 von 24) nicht signifikant wurde, wurde der Unterschied zwischen den Bedingungen mit kombinierter und summarischer Darbietung (# Pbn mit  $\Delta P$ : 18 von 24) sogar signifikant. Bedeutsam an diesem Ergebnis von Ward & Jenkins (1965) ist, dass sich der häufigere Gebrauch regel-basierter Strategien bei einer rein summarischen Darbietung nicht durch eine unterschiedlich übersichtliche Informationsanordnung erklären lässt, da die in der kombinierten und der rein summarischen Bedingung verwendeten Präsentationsformen (tabellarische Ergebniszusammenfassung) identisch waren.

Ward & Jenkins (1965) erklärten den Unterschied zwischen den Bedingungen mit kombinierter und summarischer Darbietung damit, dass es bei summarischer Darbietung nicht entscheidend sei, dass die Informationen gleichzeitig verfügbar seien (denn das sind sie in der kombinierten Bedingung auch). Stattdessen könne die Anordnung, Zusammenstellung oder Struktur (englisch: *organization*), welche der Information für die Bereitstellung der summarischen Zusammenfassung auferlegt wurde, nur dann "effektiv" sein, wenn sie den Probanden bereits vorliegt, *bevor* sie damit beginnen, sich die (summarisch dargebotenen) Informationen anzueignen. Warum dies so ist, diskutierten Ward & Jenkins (1965) nicht.

Nach unserer Einschätzung kommt der leichte Anstieg der mittleren Anzahl regel-konformer Lösungen zwischen der Einzelfallbedingung ohne Notizen und den beiden Einzelfallbedingungen mit Notizen durch eine reduzierte Beanspruchung des Arbeitsgedächtnisses in den beiden Bedingungen mit Notizen zu Stande. Der Anstieg zwischen den beiden Einzelfallbedingungen mit Notizen und der Bedingung *ohne N sum* lässt sich besser dadurch erklären, dass die meisten Probanden von einer einzelfallweisen Darbietung mit anschließender Ergebniszusammenfassung mental immer noch so stark beansprucht werden, dass sie nicht in der Lage sind, von einer hohen visuellen Struktur der



vorliegenden Informationsanordnung zu profitieren und eine gleich hohe Anzahl regelkonformer Lösungen zu produzieren wie die Probanden unter der Bedingung *ohne N sum*.

Über die Gründe, warum die mentale Beanspruchung der Probanden trotz einer Entlastung des Arbeitsgedächtnisses in den beiden Bedingungen *N strukturiert* und *N frei* immer noch so hoch ist, kann im Rahmen dieser Arbeit nur spekuliert werden:

Ein Grund hängt vermutlich mit der unterschiedlichen *Aufgabenbeschaffenheit* (s.a. Kapitel 9.8) bei summarischer und einzelfallweiser Darbietung zusammen. Trotz Gedächtnisentlastung verbringen die Probanden bei einzelfallweiser Darbietung viel Zeit mit dem Sammeln und Ordnen der Daten, sowie mit dem Anfertigen von Notizen (s. die Ergebnisse in Kapitel 10.5 *Einfluss des Anfertigens von Notizen auf die Bearbeitungszeit*). Wir nehmen an, dass bei einer einzelfallweisen Darbietung die Aufmerksamkeit während der Aufgabenbearbeitung auf diese Handlungsrouninen gerichtet ist und nicht wie in der Bedingung *ohne N sum*, bei der den Probanden das Sammeln und Ordnen der Daten erspart bleibt, auf das Erfassen von Zusammenhängen zwischen verschiedenen Häufigkeiten. Durch die Fokussierung auf einfache, aber zeitaufwendige Handlungsrouninen können die Probanden in der Zeit der eigentlichen Aufgabenbearbeitung (*time on task*) die Zusammenhänge und Beziehungen zwischen den einzelnen Häufigkeiten nicht unmittelbar erfassen, es fehlt ihnen die Gesamtschau, der Überblick, der notwendig wäre, um Vergleiche vorzunehmen, Berechnungen anzustellen oder Schlussfolgerungen zu ziehen. Eine Gesamtschau der Informationen, die für eine abschließende Beurteilung der Stärke eines Ursache-Wirkungs-Zusammenhangs notwendig sind, ist während der *time on task* nicht möglich, sondern erst danach. Möglicherweise bauen sich die Probanden während der *time on task* eine Repräsentation der Daten auf, die von den Probanden nur schwer mit den Eindrücken aus einer anschließend präsentierten summarischen Ergebniszusammenfassung in Einklang zu bringen ist (Wasserman & Shaklee, 1984). Im Gegensatz dazu kann sich unter der Bedingung *ohne N sum* die Aufmerksamkeit der Probanden schon während der *time on task* ganz auf eine abschließende Beurteilung der Stärke eines Ursache-Wirkungs-Zusammenhangs richten.

Ein Überblick über die Gesamtheit der Daten ist eine notwendige Bedingung dafür, dass eine abschließende Urteilsbildung vorgenommen werden kann. Erst dann, wenn alle relevanten Informationen überblickt werden können, können die Probanden von Strukturierungsangeboten und visuellen Hinweisen, die in einer externalen Repräsentation enthalten sind, profitieren und regelkonforme Vergleiche, Differenzbildungen und Schlussfolgerungen vornehmen. Dennoch bleibt es eine offene Frage, warum eine Ergebniszusammenfassung im Anschluss an eine einzelfallweise Darbietung mit Notizen, in

denen die Beobachtungsergebnisse schriftlich zusammengefasst sind, nicht dazu ausreicht, dass die Probanden ebenso viele regel-konforme Lösungen produzieren wie bei einer summarischen Darbietung. Hier scheinen weitere Untersuchungen notwendig zu sein, die die in Experiment 3 vermuteten Teileinflüsse einer Übersichtsdarstellung und des Zeitpunktes der Präsentation einer Gesamtdarstellung vor oder nach einer einzelfallweisen Darbietung aller relevanten Informationen genauer unter die Lupe nehmen und mit dem Einfluss einer rein summarischen Darbietung vergleichen.

Während der einzelfallweisen Darbietung mit oder ohne Notizen werden die Zusammenhänge zwischen den verschiedenen Zellohäufigkeiten wegen des fehlenden Gesamtüberblicks und der Fokussierung auf die Einzelfälle – unabhängig von der eingesetzten Präsentationsform - nicht salient gemacht. Möglich ist, dass die Probanden nach der längeren Beschäftigung mit einzelfallweisen Informationen weniger als bei einer rein summarischen Darbietung dazu in der Lage sind, visuelle Hinweise und Strukturierungsangebote einer im Anschluss an die dargebotenen Einzelfälle präsentierten Ergebniszusammenfassung zu nutzen; bildhaft gesprochen können die Probanden aufgrund der langen Beschäftigung mit Einzelfällen vor lauter Bäumen den Wald nicht mehr sehen.

Die Auswirkung eines fehlenden Gesamtüberblicks während der *time on task* bei einzelfallweiser Darbietung lässt sich auch mit Hilfe der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) plausibel erklären: Danach wirkt sich eine fehlende Gesamtschau von relevanten Informationen vor und während der Aufeinanderfolge von Einzelfällen so aus, dass die mentale Beanspruchung zwar teilweise durch das Anfertigen von Notizen reduziert wird, aber immer noch im Vergleich zu einer rein summarischen Darbietung erhöht ist, weil zusätzliche Aufmerksamkeit für den Einsatz von einfachen, aber zeitaufwendigen Handlungsabläufen aufgebracht werden muss. Der zusätzliche kognitive Aufwand, der dabei betrieben werden muss, führt dazu, dass die Probanden auf einfachere Strategien bei der Urteilsbildung zurückgreifen und dadurch weniger regel-konforme Lösungen produzieren.

Bei der Interpretation der Ergebnisse aus Experiment 3 kann nicht ausgeschlossen werden, dass bei einzelfallweiser und summarischer Darbietung unterschiedliche Informationsverarbeitungsmechanismen aktiviert werden. Die Ergebnisse zu den Bearbeitungszeiten (s. den nachfolgenden Teilabschnitt *Einfluss des Anfertigens von Notizen auf die Bearbeitungszeit*) zeigten, dass die Prozesse bei einzelfallweiser Darbietung mit Notizen wesentlich zeitaufwendiger sind als bei einzelfallweiser Darbietung ohne Notizen. Speziell das eigenständige Anfertigen von Notizen kann zu einer intensiveren, eingehenden Beschäftigung mit den selbst erstellten Notizen führen. Zhang (1997, in Anlehnung an

Goody, 1977; Norman, 1993; und Ong, 1982) hebt hervor, dass das Schreiben Sprache nicht nur von einer auditiven Form in eine sichtbare Schriftform umwandelt, sondern dass das Denken dadurch selbst verändert wird. Nach Norman (1993) unterstützt Schreiben reflektiertes Denken. In Verbindung mit der Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen könnten die Notizen die Probanden dazu anregen oder auffordern, dass sie sich intensiver und eingehender mit der Urteilsbildung beschäftigen. Dies könnte dazu geführt haben, dass der Anteil von Probanden, die andere Strategien einsetzen, welche möglicherweise von regel-basierten Strategien abweichen und mehr idiosynkratische Züge haben, zunimmt. Idiosynkratische Strategien, die nicht mit den Vorhersagen einer regel-basierten Strategie übereinstimmten, wurden von uns in der Regel als einfache Heuristik eingestuft. Mit der Rolle idiosynkratischer Strategien werden wir uns eingehender in der Diskussion der Ergebnisse aus den Häufigkeitsanalysen in Experiment 3 beschäftigen.

Eine andere Begründung des Anstiegs der Anzahl regel-konformer Lösungen zwischen den beiden Bedingungen mit Notizen und der Bedingung *ohne N sum* geht von der Vermutung aus, dass die Probanden aufgrund der höheren mentalen Beanspruchung in den beiden Bedingungen mit Notizen beim Anfertigen der Notizen Fehler produziert hätten. Diese Fehler würden dann eine im Vergleich zur Bedingung *ohne N sum* geringere Anzahl regel-konformer Lösungen nach sich ziehen. Für die Richtigkeit dieser Erklärung des Anstiegs der Anzahl regel-konformer Lösungen zwischen den beiden Bedingungen mit Notizen und der Bedingung *ohne N sum* haben wir keine Anzeichen gefunden, da eine Überprüfung der angefertigten Notizen und Summenbildungen ergab, dass diese meist korrekt waren. Auch Ward & Jenkins (1965), sowie Smedslund (1963) konnten in ihren Untersuchungen feststellen, dass die dabei gefundenen Unterschiede nicht vollständig auf die Unfähigkeit der Probanden zurückzuführen sind, die vier Zellhäufigkeiten über die Einzelfalldarbietungen hinweg aufzuaddieren.

Zu Abschluss der Diskussion über den Einfluss des Faktors *Notizen anfertigen* auf die Anzahl regel-konformer Lösungen soll noch erörtert werden, warum die Probanden unerwarteter Weise unter der Bedingung *N strukturiert* keine höhere Anzahl regel-konformer Lösungen erzielten als unter der Bedingung *N frei*. Eventuelle Unterschiede im numerischen Denken scheiden als mögliche Ursache aus, da Mittelwertsvergleiche ergaben, dass sich die Probanden in den beiden Bedingungen mit Notizen nicht signifikant hinsichtlich ihrer Fertigkeiten im numerischen Denken voneinander unterschieden, auch wenn die Probanden unter der Bedingung *N frei* etwas höhere Werte erzielten als die Probanden unter der der Bedingung *N strukturiert*.

Im Einklang mit unserer bisherigen Argumentation zur Begründung des Unterschiedes zwischen den beiden Bedingungen mit Notizen und der Bedingung *ohne N sum*, kann angenommen werden, dass die Probanden aufgrund der hohen mentalen Beanspruchung in den beiden Bedingungen mit Notizen von einer im Vergleich zum freien Anfertigen von Notizen vermeintlich größeren Übersichtlichkeit beim strukturierten Anfertigen von Notizen nicht zusätzlich profitieren konnten.

Eine weitere Erklärungsalternative geht von der Annahme aus, dass die Probanden unter der Bedingung *N frei* für sich selbst im Vergleich zur Bedingung *N strukturiert* keine weniger übersichtlichen Notizen angefertigt haben. Dafür spricht, dass immerhin 13 von 20 Probanden (65 %) unter der Bedingung *N frei* eigenständig eine Vierfeldertafel mit Strichen oder anderen Symbolen angefertigt haben. Weitere 3 Probanden (15 %) haben die Einzelfälle entweder in vier Zeilen oder vier Spalten eingetragen. Von den übrigen vier Probanden haben jeweils zwei (je 10 %) eine Einzelfallliste erstellt oder lediglich Erfolgsfälle aufgezeichnet. Ein möglicher Grund für dieses relativ strukturierte Anfertigen von Notizen der Probanden in der Bedingung *N frei* könnte darin liegen, dass die Mehrheit der Probanden durch die Präsentation der Einzelfallinformationen in einer Vierfeldertafel mit Positionskreuz dazu angeregt wurden, ihre Aufzeichnungen in der gleichen Weise anzuordnen. Demzufolge würde die Präsentation der einzelfallweise dargebotenen Informationen in einer Vierfeldertafel mit Positionskreuz eine Triggerfunktion für ein strukturiertes Anordnen der Informationen in einer Vierfeldertafel beinhalten. Um herauszufinden, welchen Einfluss die Strukturierung der Anordnung einzelfallweise dargebotener Informationen auf die Anordnung von Informationen beim Anfertigen von Notizen hat, könnte der Einfluss des Faktors *Strukturierung der Anordnung einzelfallweiser Informationen* auf die Anordnung von Informationen beim Anfertigen von Notizen untersucht werden, indem die beiden Bedingungen *Fliesstext mit Notizen frei* und *Vierfeldertafel mit Positionskreuz und Notizen frei* miteinander verglichen werden.

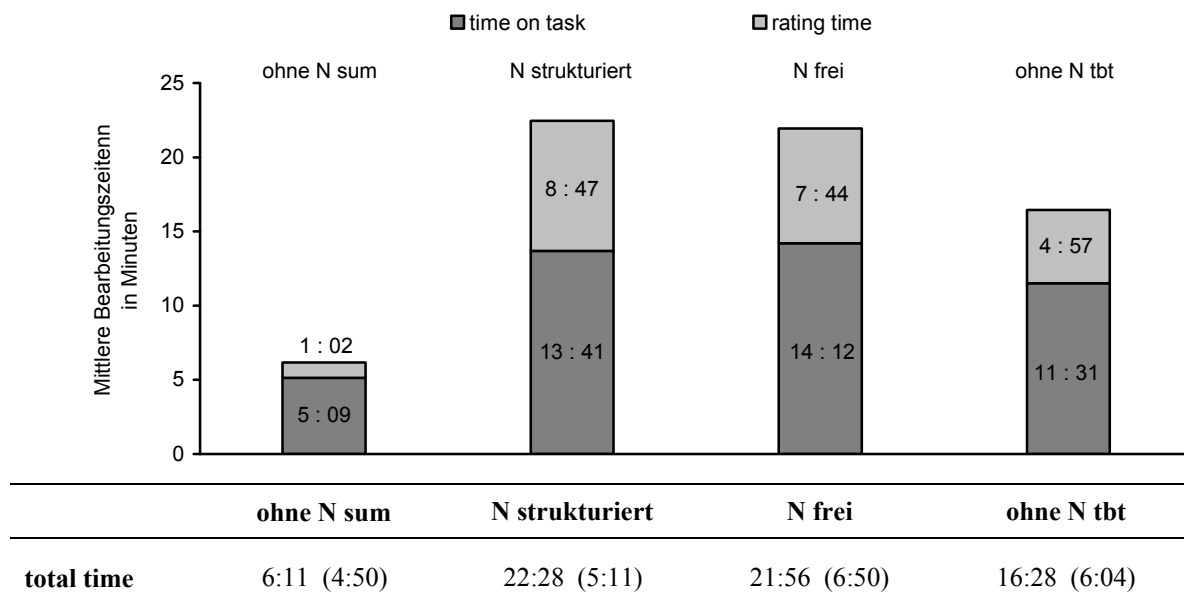
Der Einfluss der Kovariate *numerisches Denken* auf die Anzahl regel-konformer Lösungen wurde wie in Experiment 2 hochsignifikant. Eine eigene Diskussion erübrigt sich hier jedoch.

## 10.5.2 Bearbeitungszeit

### 10.5.2.1 Ergebnisse

Die *mittleren Bearbeitungszeiten* in Abhängigkeit vom Faktor *Notizen anfertigen* sind in Abbildung 41 eingetragen. Wie schon in Experiment 2 steht jede Säule für die

Gesamtbearbeitungszeit ("total time"), die sich aus den beiden Komponenten "time on task" und "rating time" zusammensetzt.



**Abbildung 41:** Experiment 3: Mittlere Bearbeitungszeiten in Minuten : Sekunden (in Klammern Standardabweichungen) in Abhängigkeit von *Notizen anfertigen*.

Die Mittelwerte weisen darauf hin, dass die beiden Bedingungen *N strukturiert* und *N frei* sich nicht signifikant voneinander unterscheiden. Zudem waren die Bearbeitungszeiten unter der Bedingung *ohne N sum* kürzer als diejenigen unter der Bedingung *ohne N tbt*, welche wiederum kürzer ausgefallen ist als diejenigen unter den beiden Bedingungen mit Notizen. Zur Überprüfung von Hypothese H9b wurden Kovarianzanalysen gerechnet, in denen der Einfluss von *Notizen anfertigen* und der Kovariate *numerisches Denken* auf die *total time* und ihre beiden Teilkomponenten *rating time* und *time on task* untersucht wurde. Die Ergebnisse dieser Analysen finden sich in Tabelle 22:

**Tabelle 22:** Experiment 3: Ergebnisse aus den univariaten Kovarianzanalysen, in denen der Einfluss von *Notizen anfertigen* und *numerisches Denken* auf die drei Zeitmaße *total time*, *time on task* und *rating time* untersucht wurde. Signifikante P-Werte sind **fett** gedruckt.

	Faktor	F-Wert	MSE	P-Wert
<b>total time</b>	Notizen anfertigen (3,75)	36.939	31.81	< <b>.001</b>
	numerisches Denken (3,75)	4.909		<b>.030</b>
<b>time on task</b>	Notizen anfertigen (3,75)	19.412	18.579	< <b>.001</b>
	numerisches Denken (3,75)	4.647		<b>.034</b>
<b>rating time</b>	Notizen anfertigen (3,75)	13.949	17.367	< <b>.001</b>
	numerisches Denken (3,75)	.591		.445

Die Ergebnisse zeigen, dass der Faktor *Notizen anfertigen* einen hochsignifikanten Einfluss auf alle drei Zeitmaße hat. Die Kovariate *numerisches Denken* hat einen signifikanten Einfluss auf die *Gesamtbearbeitungszeit*, der im Wesentlichen auf Unterschiede hinsichtlich der *time on task* zurückzuführen ist. Auf die *rating time* hat das *numerische Denken* hingegen keinen signifikanten Einfluss. Mit der Hilfe von paarweisen Vergleichen zwischen Bonferroni adjustierten Gruppenmittelwerten sollte der Frage nachgegangen werden, welche der vier Bedingungen sich signifikant im Hinblick auf die drei Zeitmaße unterscheiden. Tabelle 23 zeigt, welche Einzelkontraste signifikant wurden und welche nicht.

**Tabelle 23:** Experiment 3: Paarweise Vergleiche zwischen den Bonferroni-adjustierten Gruppenmittelwerten für die Variablen *total time*, *time on task* und *rating time*. Signifikante P-Werte sind **fett** gedruckt.

		ohne N sum	N strukturiert	N frei	ohne N tbt
<b>total time</b>	ohne N sum	-	< .001	< .001	< .001
	N strukturiert		-	1.000	.004
	N frei			-	.005
	ohne N tbt				-
<b>time on task</b>	ohne N sum	-	< .001	< .001	< .001
	N strukturiert		-	1.000	.437
	N frei			-	.113
	ohne N tbt				-
<b>rating time</b>	ohne N sum	-	< .001	< .001	.032
	N strukturiert		-	1.000	.023
	N frei			-	.172
	ohne N tbt				-

Für alle drei Zeitmaße gilt, dass die Probanden unter der Bedingung *ohne N sum* signifikant schneller arbeiteten als unter den drei anderen Bedingungen. Probanden in den Bedingungen *N strukturiert* und *N frei* benötigten hingegen sowohl bei der eigentlichen Aufgabenbearbeitung (*time on task*) als auch während der Wirksamkeitsbeurteilung (*rating time*) in etwa gleich viel Zeit. Probanden unter der Bedingung *ohne N tbt* arbeiteten zwar schneller als die Probanden in den beiden Bedingungen mit Notizen, doch wurden nur die Unterschiede hinsichtlich der *Gesamtbearbeitungszeit* sowie der Kontrast *ohne N tbt* versus *N strukturiert* im Hinblick auf die *rating time* signifikant.

#### 10.5.2.2 Diskussion

Die Ergebnisse aus den paarweisen Vergleichen mit adjustierten Bonferroni Gruppenmittelwerten bestätigen die Hypothese H9b. Probanden, die ohne Notizen arbeiten,

benötigen weniger Zeit als solche, die mit Notizen arbeiten. Das Ergebnis, dass Probanden bei summarischer Informationsdarbietung ohne Notizen schneller sind als Probanden bei Einzelfalldarbietung ohne Notizen ist schon aus Experiment 2 bekannt und wurde dort diskutiert.

Die langsameren Bearbeitungszeiten in den beiden Bedingungen mit Notizen lassen sich dadurch erklären, dass die Probanden insbesondere im Vergleich zur Bedingung *ohne N tbt* während der *time on task* Zeit für das zusätzliche Anfertigen von Notizen verloren haben. Während der *rating time* haben sich die Probanden vermutlich in den Bedingungen mit Notizen eingehender mit den selbst erstellten Notizen beschäftigt, welche sie dann bei der Urteilsbildung zu Rate zogen. Möglich ist, dass das Anfertigen von Notizen die Probanden im Sinne von Norman (1993) dazu angeregt hat, intensiver über ihre Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen nachzudenken.

Für die Urteilsbildung (*rating time*) haben sich die Probanden in den beiden Bedingungen mit Notizen sehr viel mehr Zeit gelassen als in der Bedingung *ohne N sum*. Auf den ersten Blick mag dieses Ergebnis überraschend erscheinen, da doch in allen drei Bedingungen summarische Ergebniszusammenfassungen vorlagen. Bei genauerer Betrachtung zeigen sich hier aber zwei Dinge: Zum einen bestätigen die sehr kurzen *rating times* unter der Bedingung *ohne N sum* die bereits im Zusammenhang mit der Diskussion zum Einfluss des Faktors Darbietungsart auf die Bearbeitungszeiten in Experiment 2 geäußerte Vermutung, dass bei einer rein summarischen Darbietung die Prozesse für die Urteilsbildung von der *rating time* auf die *time on task* verlagert werden, d.h. während der *rating time* wird die Urteilsbildung im Wesentlichen nur noch vollzogen und die für die Urteilsbildung erforderlichen mentalen Operationen laufen schon in der Zeit für die Informationsaufnahme ab. Zum anderen weisen die sehr langen *rating times* unter den beiden Bedingungen mit Notizen darauf hin, dass während der *time on task* nur einfache Handlungsroutinen und Prozesse der Informationsaufnahme und –speicherung ausgeführt werden. Dagegen laufen die komplexen Denkoperationen, die im Zusammenhang mit der Urteilsbildung stehen, ausschließlich während der *rating time* ab. Zudem wurden die Probanden während der *rating time* im Vergleich zur Bedingung *ohne N tbt* durch die vorliegenden und eigenständig erstellten Notizen dazu angeregt, sich intensiver mit der Urteilsbildung auseinanderzusetzen.

Unerwartet, wenn auch nicht ganz überraschend ist der signifikante Einfluss der Kovariate *numerisches Denken* auf die beiden Zeitmaße *total time* und *time on task*. Auf die *rating time* hatte das numerische Denken dagegen keinen signifikanten Einfluss. Weitere Analysen zeigten, dass Probanden mit hohen Werten im numerischen Denken für die eigentliche

Aufgabenbearbeitung (*time on task*) signifikant weniger Zeit benötigten, als Probanden mit geringen Werten im numerischen Denken, was sich auch auf die Gesamtbearbeitungszeit auswirkte. Im Zusammenhang mit dem nicht signifikanten Einfluss des numerischen Denkens auf die *rating time* ist zu vermuten, dass für die *rating time* motivationale Faktoren bedeutsamer sind.

Die Annahme, dass die Kovariate *numerisches Denken* einen signifikanten Einfluss auf die Gesamtbearbeitungszeit, insbesondere auf die *time on task*, hat, ist nicht vorhergesagt worden. In Experiment 2 hatte numerisches Denken auf keines der drei Zeitmaße einen signifikanten Einfluss. Die Signifikanz von numerischem Denken in Experiment 3 ist umso erstaunlicher, wenn berücksichtigt wird, dass die Experimente 2 und 3 die zwei Bedingungen *ohne N sum* bzw. *SD sum* und *ohne N tbt* bzw. *VT tbt* gemeinsam haben. Worin liegen also die Unterschiede zwischen beiden Experimenten begründet?

Weitere Analysen der Zeitmaße aus Experiment 2 in Abhängigkeit vom Faktor *numerisches Denken* enthüllten hinsichtlich der *time on task* eine interessante Wechselwirkung zwischen den Faktoren *visuelle Strukturierung* und *numerisches Denken*: Während Probanden mit hohen Werten im numerischen Denken bei hoher visueller Strukturierung des darstellenden Mediums (Bedingungen *SD sum* und *VT tbt*) für die eigentliche Aufgabenbearbeitung (*time on task*) im Mittel 2 Minuten 41 Sekunden weniger benötigten als die Probanden mit geringen Werten im numerischen Denken ( $MW_{ND\text{ gering}} = 9\text{ m }40\text{ s}$ ;  $MW_{ND\text{ hoch}} = 6\text{ m }59\text{ s}$ ), verhielt es sich bei geringer visueller Strukturierung des darstellenden Mediums (Bedingungen *FT sum* und *FT tbt*) fast genau umgekehrt: Hier waren die Probanden mit geringen Werten im numerischen Denken durchschnittlich 1 Minute und 18 Sekunden schneller ( $MW_{ND\text{ gering}} = 10\text{ m }11\text{ s}$ ;  $MW_{ND\text{ hoch}} = 11\text{ m }29\text{ s}$ ). Über alle Probanden hinweg führt die Wechselwirkung zwischen *numerischem Denken* und *visueller Strukturierung* also dazu, dass der Einfluss des numerischen Denkens auf die *time on task* – und damit auch auf die Gesamtbearbeitungszeit – durch die Variation der *visuellen Strukturierung des darstellenden Mediums* neutralisiert wird. Daraus kann die Schlussfolgerung gezogen werden, dass der Einfluss des numerischen Denkens auf die Bearbeitungszeit entscheidend durch mediale Rahmenbedingungen verändert werden kann.

In Experiment 3 wurden den Probanden anders als in Experiment 2 keine Fliesstexte vorgelegt. In den drei Einzelfallbedingungen *N strukturiert*, *N frei* und *ohne N tbt* hatten sie es immer mit einer Vierfeldertafel mit Positionskreuz zu tun (Bedingung *VT tbt* aus Experiment 2; s.a. Abbildung 21), es ist deshalb zu vermuten, dass die Probanden mit hohen Werten im numerischen Denken auch in den beiden Bedingungen *N strukturiert* und *N frei* weniger Zeit



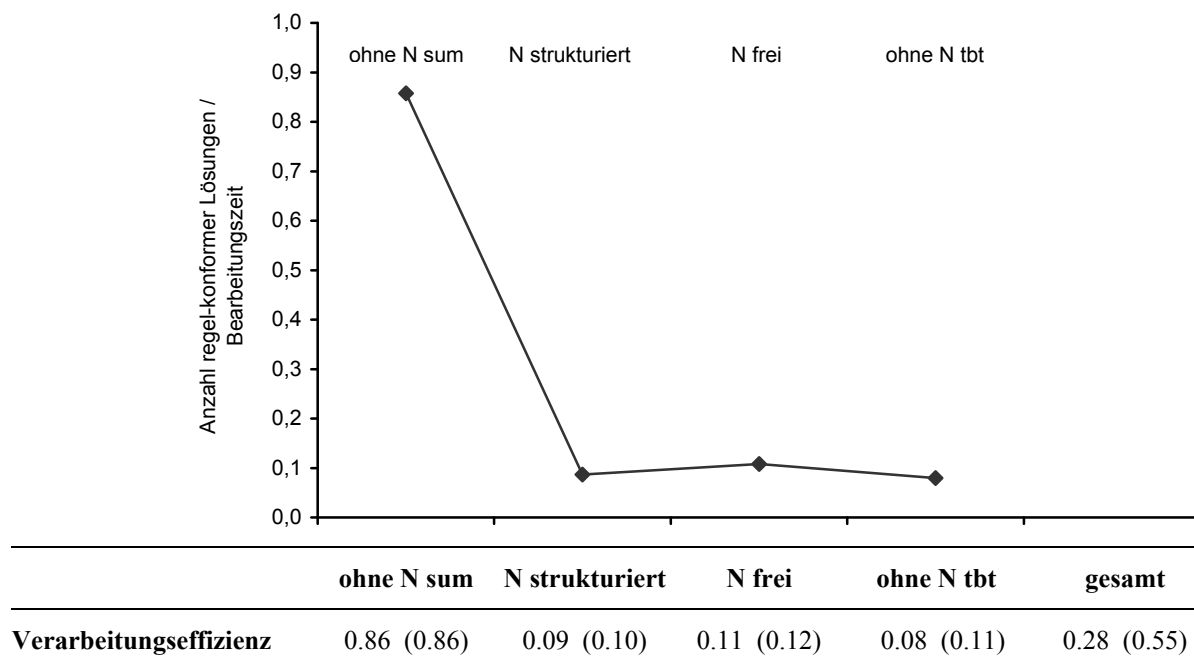
für die eigentliche Aufgabenbearbeitung benötigt haben. Diese Annahme, die schon durch die Signifikanz der Kovariate *numerisches Denken* für die beiden Zeitmaße *total time* und *time on task* nahe gelegt wird, wurde zusätzlich durch einen Vergleich der Bearbeitungszeiten von Probanden mit hohen und mit geringen Werten im numerischen Denken bestätigt (s. Tabelle 35; S. 313).

Im Zusammenhang mit der Diskussion hier erscheint aus theoretischer Sicht die Annahme nahe liegend, dass die Bearbeitungszeit von Fliesstexten in Experiment 2 nicht signifikant mit den Fähigkeiten der Probanden im numerischen Denken korreliert. Im Zusammenhang mit den Ergebnissen aus Experiment 3 ist es im Nachhinein ebenfalls plausibel davon auszugehen, dass Probanden mit hohen Werten im numerischen Denken komplexe Grafiken schneller erfassen und schneller Notizen erstellen.

### 10.5.3 Verarbeitungseffizienz

#### 10.5.3.1 Ergebnisse

Abbildung 42 veranschaulicht die Ergebnisse für das Effizienzmaß *Anzahl regel-konformer Lösungen / Bearbeitungszeit* in Abhängigkeit vom Faktor *Notizen anfertigen*.



**Abbildung 42:** Experiment 3: Verarbeitungseffizienz in Abhängigkeit von *Notizen anfertigen* (in Klammern Standardabweichungen).

Die Mittelwerte weisen darauf hin, dass eine Kovarianzanalyse mit dem Faktor *Notizen anfertigen* und der Kovariate *numerisches Denken* aufgrund der im Vergleich zu den anderen

drei Bedingungen höheren Verarbeitungseffizienz unter der Bedingung *ohne N sum* signifikant werden könnte. Die Ergebnisse der beiden Kovarianzanalysen mit der abhängigen Variable *Anzahl regel-konformer Lösungen / Bearbeitungszeit* werden in Tabelle 24 gezeigt.

**Tabelle 24:** Experiment 3: Ergebnisse aus univariaten Kovarianzanalysen, in denen der Einfluss von *Notizen anfertigen* und *numerisches Denken* auf die Verarbeitungseffizienz untersucht wurde. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen.

Faktor	F-Wert	MSE	P-Wert
Notizen anfertigen <sub>(3,75)</sub>	18.295	.164	< <b>.001</b>
numerisches Denken <sub>(3,75)</sub>	14.945		< <b>.001</b>

Sowohl der Faktor *Notizen anfertigen* als auch die Kovariate *numerisches Denken* hatten einen hochsignifikanten Einfluss auf die *Verarbeitungseffizienz*. Wie die in Tabelle 25 veranschaulichten Ergebnisse aus den paarweisen Vergleichen zwischen Bonferroni-adjustierten Gruppenmittelwerten zeigen, arbeiteten die Probanden unter der Bedingung *ohne N sum* hochsignifikant effizienter als in den drei Einzelfallbedingungen *ohne N tbt*, *N strukturiert* und *N frei*, wo sich die Mittelwerte faktisch nicht unterscheiden. Zwar generierten die Probanden unter der Bedingung *ohne N tbt* etwas weniger regel-konforme Lösungen als die Probanden unter den beiden Bedingungen mit Notizen, doch arbeiteten sie dafür vergleichsweise schneller und erreichten so im Endeffekt fast identische Effizienzmaße wie die Probanden unter den Bedingungen *N strukturiert* und *N frei*.

**Tabelle 25:** Experiment 3: Paarweise Vergleiche zwischen den Bonferroni-adjustierten Gruppenmittelwerten für die abhängige Variable *Verarbeitungseffizienz* in Abhängigkeit von *Notizen anfertigen*. Signifikante P-Werte ( $\alpha < .05$ ) sind **fett** markiert.

	ohne N sum	N strukturiert	N frei	ohne N tbt
ohne N sum	-	< <b>.001</b>	< <b>.001</b>	< <b>.001</b>
N strukturiert		-	1.000	1.000
N frei			-	1.000
ohne N tbt				-

### 10.5.3.2 Diskussion

Die Hypothese H9c wird durch die Ergebnisse bestätigt. Probanden, denen die Informationen summarisch in einem Stapeldiagramm dargeboten wurden, arbeiten auch ohne die Möglichkeit, sich Notizen zu machen, wesentlich effizienter als Probanden, denen die Informationen über Ursache-Wirkungs-Zusammenhänge einzelfallweise dargeboten werden, selbst wenn ihr Arbeitsgedächtnis bei der Wirksamkeitsbeurteilung durch die Möglichkeit,

sich Notizen zu machen, entlastet wird. Die Ergebnisse für die Verarbeitungseffizienz lassen sich klar aus den bisherigen Ergebnissen zum Einfluss des Faktors *Notizen anfertigen* auf die Anzahl regel-konformer Lösungen und die Bearbeitungszeit ableiten. Zum einen erzielten die Probanden unter den drei Bedingungen *N strukturiert*, *N frei* und *ohne N tbt* nicht dieselbe Anzahl regel-konformer Lösungen wie die Probanden unter der Bedingung *ohne N sum*, zum anderen arbeiteten sie auch bei der Beurteilung der Wirksamkeiten (*rating time*) und nicht nur bei der eigentlichen Aufgabenbearbeitung (*time on task*) bedeutend langsamer als die Probanden unter der Bedingung *ohne N sum*.

Der Einfluss des Faktors *Notizen anfertigen* auf die Verarbeitungseffizienz lässt sich sehr gut mit Hilfe der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) interpretieren. Bei einer rein summarischen Darbietungsform arbeiteten die Probanden am effizientesten, weil die mentale Beanspruchung (die *Kosten*) durch die Reduktion auszuführender Aufgabeninhalte (keine Summenbildung, kein Ordnen einzelfallweise dargebotener Informationen erforderlich), die Möglichkeit zur Gesamtschau auf alle dargebotenen Informationen von Anfang an und die hohe visuelle Strukturierung des Stapeldiagramms erheblich reduziert wurde. Dadurch werden die Probanden in die Lage versetzt, eine höhere Anzahl regel-konformer Lösungen zu produzieren (der *Nutzen*).

In den drei Bedingungen mit einzelfallweiser Informationsdarbietung, wo die Verarbeitungseffizienzen signifikant geringer ausfielen als unter der Bedingung *ohne N sum*, zeigt sich der erwartete *speed-accuracy trade-off*: In den beiden Bedingungen mit Notizen wird zwar das Arbeitsgedächtnis durch das Anfertigen von Notizen entlastet. Nach der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) kann davon ausgegangen werden, dass die Probanden in den beiden Bedingungen mit Notizen in Verbindung mit der Entlastung des Arbeitsgedächtnisses antizipieren, dass komplexere Strategien eingesetzt werden können. Sie sind bereit, mit der Hilfe von Notizen einen höheren Aufwand bei der Urteilsbildung zu betreiben, was mit längeren Bearbeitungszeiten einhergeht. In der Bedingung *ohne N tbt* sehen die Probanden keinen Nutzen darin, eine komplexere Strategie einzusetzen und greifen lieber auf einfachere Strategien zurück, für die sie aber im Gegenzug vor allem während der Urteilsbildung (*rating time*) weniger Zeit investieren (bei der *time on task* entsteht der zeitliche Mehraufwand im Vergleich zur Bedingung *ohne N tbt* nicht durch den Einsatz einer komplexeren Strategie zur Urteilsbildung, sondern durch das zusätzliche Anfertigen von Notizen). Letztlich ergibt sich daraus, dass die Probanden in den drei Einzelfallbedingungen in etwa gleich effizient gearbeitet haben.

Der signifikante Einfluss der Kovariate *numerisches Denken* bestätigt die Hypothese H7b. Das Ergebnis, dass Probanden mit hohen Fertigkeiten im numerischen Denken effizienter arbeiten als Probanden mit geringen Fertigkeiten, reproduziert entsprechende Ergebnisse aus Experiment 2 und kann ebenfalls aus bisherigen Ergebnissen zum Einfluss des Faktors *numerisches Denken* auf die Anzahl regel-konformer Lösungen und die Bearbeitungszeit hergeleitet werden.

## 10.6 Häufigkeitsanalysen zum Einfluss von Notizen auf die Strategiewahl

### 10.6.1 Ergebnisse

Wie in den Experimenten 1 und 2 wurde für jeden Probanden an Hand der schriftlichen und mündlichen Nachbefragung, sowie der abgegebenen Wirksamkeitseinschätzungen ermittelt, welche Strategie(n) jeweils eingesetzt wurden. Die Kategorien für Strategietypen, denen die einzelnen Probanden zugeordnet wurden, entsprechen denen aus den Experimenten 1 und 2.

1. Strategie unklar
2. Wechsler unklar – ab-Regel
3. ab-Regel oder eine andere Heuristik
4. Wechsler zu einer regel-basierten Strategie
5. durchgängig regel-basierte Strategie

Tabelle 26 zeigt, wie sich die 80 Probanden insgesamt und in Abhängigkeit von den vier Experimentalbedingungen auf die fünf Kategorien verteilen:

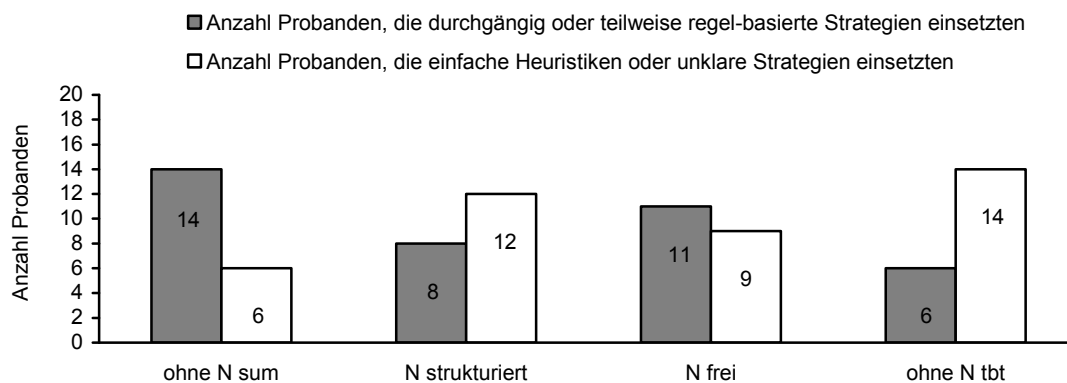
**Tabelle 26:** Experiment 3: Anzahl der Probanden in Abhängigkeit von Strategietyp und Versuchsbedingung (in Klammern Prozentangaben). Zelhäufigkeiten > 5 sind **fett** markiert.

	Strategietyp					Gesamt
	unklar	Wechsler ab-Regel	ab-Regel	Wechsler regel-basiert	regel-basiert	
<b>Versuchsbedingung</b>						
ohne N sum	3	2	1	<b>7</b>	<b>7</b>	20
N strukturiert	4	<b>6</b>	2	4	4	20
N frei	4	3	2	<b>9</b>	2	20
ohne N tbt	5	5	4	5	1	20
<b>Gesamt</b>	16 (20 %)	16 (20 %)	9 (11,25 %)	25 (31,25 %)	14 (17,5 %)	80

Insgesamt haben 39 von 80 Probanden (48,75 %) ausschließlich oder teilweise eine regel-basierte Strategie eingesetzt, davon 14 (17,5 %) ausschließlich. Entsprechend den Erwartungen ist der Anteil der Probanden, die ausschließlich oder teilweise von regel-

basierten Strategien Gebrauch gemacht haben, unter der Bedingung *ohne N sum* am höchsten (70 %) und unter der Bedingung *ohne N tbt* am geringsten (30 %). In den beiden Bedingungen mit Notizen liegt der Anteil der Probanden, die ausschließlich oder teilweise von regel-basierte Strategien eingesetzt haben, bei 55 % (*N frei*) und 40 % (*N strukturiert*). Auffallend ist, dass unter der Bedingung *N frei* relativ viele Probanden zu einer regel-basierten Strategie überwechseln (9 von 20 Probanden).

In Abhängigkeit davon, ob regel-basierte Strategien zum Einsatz kamen oder nicht, wurden die Probanden wie schon in den Experimenten 1 und 2 in zwei Gruppen unterteilt. Gruppe 1 setzte sich aus Probanden zusammen, die durchgängig oder teilweise regel-basierte Strategien angewendet haben, Gruppe 2 aus Probanden, die nur von einfachen Heuristiken wie der ab-Regel oder unklaren Strategien Gebrauch gemacht haben. Abbildung 43 veranschaulicht diese binäre Unterteilung der Probanden hinsichtlich der eingesetzten Strategie.



**Abbildung 43:** Experiment 3: Verteilung von Probanden nach Strategietyp (durchgängig oder teilweise regel-basiert versus nicht regel-basiert) in Abhängigkeit von *Notizen anfertigen*.

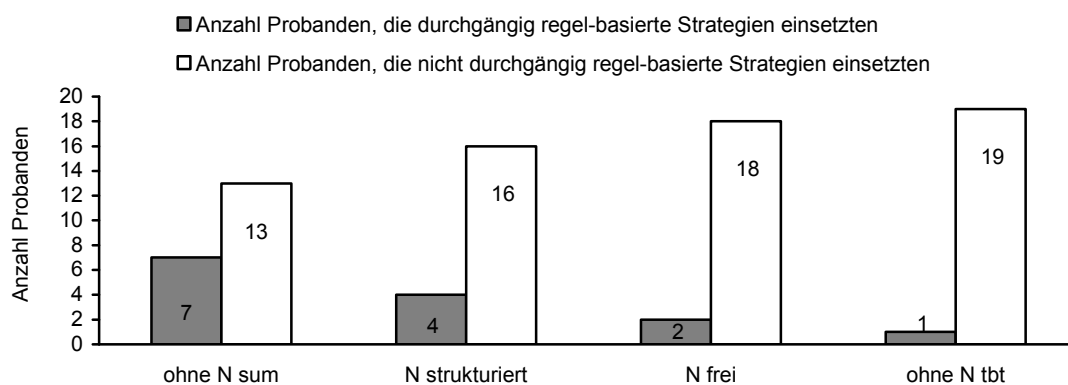
Wie schon in Experiment 2 ergab ein  $\chi^2$  - Test den Trend, dass die Verteilung von regel-basiert und nicht regel-basiert vorgehenden Strategieanwendern in Abhängigkeit von der Versuchsbedingung nicht zufällig ist ( $\chi^2_{(3, 95\%)} = 7,355$ ;  $p = .061$ ; Phi-Wert = .303).

Anhand der Daten von Abbildung 43 konnte auch die Hypothese H9d überprüft werden. Hypothese H9d sagt vorher, dass die Anzahl der Probanden, die teilweise oder durchgängig regel-basierte Strategien einsetzen, unter den Bedingungen *ohne N sum*, *N strukturiert* und *N frei* jeweils höher ist als unter der Bedingung *ohne N tbt*. Zur Überprüfung der Hypothese H9d wurden  $\chi^2$ - Tests durchgeführt, in denen untersucht wurde, ob die Verteilung der Strategiewahl jeweils unabhängig von den paarweisen Bedingungsvergleichen *ohne N sum* vs. *ohne N tbt*; *N strukturiert* vs. *ohne N tbt* und *N frei* vs. *ohne N tbt* ist oder nicht. Die Ergebnisse finden sich in Tabelle 27.

**Tabelle 27:** Experiment 3: Ergebnisse von Häufigkeitsanalysen zur Verteilung der Strategiewahl (durchgängig oder teilweise regel-basiert versus nicht regel-basiert) in Abhängigkeit von den Bedingungsvergleichen *ohne N sum* vs. *ohne N tbt*; *N strukturiert* vs. *ohne N tbt* und *N frei* vs. *ohne N tbt*. Signifikante P-Werte ( $\alpha < .05$ ) sind **fett** markiert.

Vergleich	$\chi^2$ -Wert	df	Asymptotische Signifikanz (2-seitig)	exakte Signifikanz (exakter Test nach Fisher)
ohne N sum – ohne N tbt	6.400	1	<b>.011</b>	<b>.013</b>
N strukturiert – ohne N tbt	.440	1	.507	.371
N frei – ohne N tbt	2.558	1	.110	.100

Tabelle 27 zeigt, dass die Hypothese H9d nur für den Vergleich *N sum* vs. *ohne N tbt* Gültigkeit hat, nicht aber für die Vergleiche *N strukturiert* vs. *ohne N tbt* und *N frei* vs. *ohne N tbt*. Ein sehr ähnliches Bild ergibt sich, wenn man die Anzahl der Probanden, die durchgängig regel-basierte Strategien eingesetzt haben, mit der Anzahl von Probanden vergleicht, die regel-basierte Strategien nur teilweise oder gar nicht eingesetzt haben. Abbildung 44 veranschaulicht die Verteilung von Probanden nach Strategietyp (durchgängig regel-basiert versus teilweise oder gar nicht regel-basiert) in Abhängigkeit von *Notizen anfertigen*.



**Abbildung 44.** Experiment 3: Verteilung von Probanden nach Strategietyp (teilweise oder gar nicht regel-basiert vs. durchgängig regel-basiert) in Abhängigkeit von *Notizen anfertigen*.

Wie schon in Experiment 2 ergab ein  $\chi^2$ -Test den Trend, dass die Verteilung von regel-basiert und nicht regel-basiert vorgehenden Strategiewendern in Abhängigkeit von der Versuchsbedingung nicht zufällig ist ( $\chi^2_{(3, 95\%)} = 7,273$ ;  $p = .064$ ; Phi-Wert = .302). Weitere  $\chi^2$ -Tests zum Vergleich der Bedingung *ohne N tbt* mit den drei anderen Bedingungen führten zu dem in Tabelle 28 eingetragenen Ergebnis, dass nur der Vergleich *ohne N tbt* versus *ohne N sum* signifikant wurde. Dies bedeutet, dass die Anzahl der Probanden, die durchgängig von regel-basierten Strategien Gebrauch machen, nur unter der Bedingung *ohne*

*N sum*, nicht aber unter den beiden Bedingungen mit Notizen, signifikant höher ist als unter der Bedingung *ohne N tbt*.

**Tabelle 28:** Experiment 3: Ergebnisse von Häufigkeitsanalysen zur Verteilung der Strategiewahl (durchgängig versus nicht durchgängig regel-basiert) in Abhängigkeit von den drei Bedingungsvergleichen *ohne N sum* vs. *ohne N tbt*; *N strukturiert* vs. *ohne N tbt* und *N frei* vs. *ohne N tbt*. Signifikante P-Werte ( $\alpha < .05$ ) sind **fett** markiert.

Vergleich	$\chi^2$ -Wert	df	Asymptotische Signifikanz (2-seitig)	exakte Signifikanz (exakter Test nach Fisher)
ohne N sum – ohne N tbt	5.625	1	<b>.018</b>	<b>.022</b>
N strukturiert – ohne N tbt	2.057	1	.151	.171
N frei – ohne N tbt	.360	1	.548	.500

In Experiment 3 haben fast ebenso viele Probanden wie in Experiment 2 ihre Strategie bei der Bearbeitung der verschiedenen Aufgaben gewechselt (41 von 80 oder 51,25 %). Tabelle A6 im Anhang A3.3 verschafft einen Eindruck von der tatsächlichen Vielfalt der in Experiment 3 eingesetzten Strategien.

Analog zu den Experimenten 1 und 2 wurden die Probanden aus Experiment 3, die durchgängig regel-basierte Strategien eingesetzt haben, weiter ausdifferenziert und daraufhin untersucht, ob sie von der  $\Delta P$ -Regel, der Power PC-Regel oder einer anderen regel-basierten Strategie Gebrauch gemacht haben. Tabelle 29 zeigt, wie viele Probanden in welcher Versuchsbedingung welche Strategie angewendet haben. Im Einzelnen wurden für Tabelle 29 die folgenden Kategorien, denen die Probanden zugeordnet wurden, gebildet:

1. Delta P-Regel, kurz " $\Delta P$ -Regel"
2. Power PC-Theorie, kurz "Power PC"
3. Wechsler Delta P-Regel - Power PC Theorie, kurz " $\Delta P$  & PPC"
4. Erfolgsrate  $x$  (1 – Basisrate), kurz " $ER * (1 - BR)$ "

**Tabelle 29:** Experiment 3: Anzahl Probanden, die durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von der Art der regel-basierten Strategie und der Versuchsbedingung (in Klammern Prozentangaben)..

	Art der regel-basierten Strategie				Gesamt
	$\Delta P$ -Regel	Power PC	$\Delta P$ & PPC	$ER * (1 - BR)$	
<b>Versuchsbedingung</b>					
ohne N sum	3	3	1	0	7 (50.0 %)
N strukturiert	0	0	3	1	4 (28,6 %)
N frei	0	2	0	0	2 (14.3 %)
ohne N tbt	0	1	0	0	1 (7.1 %)
<b>Gesamt</b>	3 (21,4 %)	6 (42,9 %)	4 (28,6 %)	1 (7,1 %)	14 (100 %)

Von den 14 Probanden, die durchgängig eine regel-basierte Strategie einsetzten, arbeiteten 50 % unter der Bedingung *ohne N sum*. Wie schon in Experiment 2 machten über alle vier Bedingungen hinweg betrachtet durchgängig mehr Probanden von der Power PC-Theorie als von der  $\Delta P$ -Regel Gebrauch (6 : 3). In der Gruppe der Wechsler (siehe Tabelle A6 im Anhang A3.3) hatten insgesamt 14 Probanden zur  $\Delta P$ -Regel (ohne die Kategorie "25"), 8 zur Power PC Theorie und 3 zu zwei regel-basierten Strategien gewechselt. 5 Probanden hatten zwischen zwei regel-basierten Strategien gewechselt.

Zusammengefasst erbrachten die Häufigkeitsanalysen zu Experiment 3 folgende zentrale Ergebnisse: Die Hypothese H9d wird durch die Häufigkeitsanalysen nur teilweise bestätigt. Signifikant wurde lediglich der Unterschied, dass unter der Bedingung *ohne N sum* mehr Probanden regel-basierte Strategien einsetzen als unter der Bedingung *ohne N tbt*. Unter den beiden Bedingungen *N strukturiert* und *N frei* setzten zwar weniger Probanden als unter der Bedingung *ohne N sum* und mehr Probanden als unter der Bedingung *ohne N tbt* durchgängig und teilweise regel-basierte Strategien ein, doch wurden diese Häufigkeitsunterschiede nicht signifikant. Die Zahl der Probanden, die durchgängig von regel-basierten Strategien Gebrauch machten, ist unter der Bedingung *ohne N sum* etwas höher als unter den anderen drei Bedingungen. Damit werden die Ergebnisse aus den Kovarianzanalysen mit der abhängigen Variablen *Anzahl regel-konformer Lösungen* bestätigt, in denen die Probanden unter der Bedingung *ohne N sum* im Durchschnitt eine höhere Anzahl regel-konformer Lösungen erzielten als unter den anderen drei Bedingungen *N strukturiert*, *N frei* und *ohne N tbt*. Von den wenigen Probanden, die durchgängig regel-basierte Strategien einsetzten, berechnen etwas mehr die kausale Stärke  $p$  entsprechend der Power PC Theorie, während bei den Wechslern mehr Probanden zur  $\Delta P$ -Regel übergangen.

### 10.6.2 Diskussion

Erneut bestätigen die Häufigkeitsanalysen die Ergebnisse der Kovarianzanalyse mit der abhängigen Variable *Anzahl regel-konformer Lösungen*. Der Anteil der Probanden, die teilweise oder durchgängig von regel-basierten Strategien Gebrauch gemacht haben, ist unter der Bedingung *ohne N sum* am größten, unter der Bedingung *ohne N tbt* am kleinsten. Die beiden Bedingungen mit Notizen nehmen eine Zwischenstellung ein. Entsprechend ist der Anteil der Probanden, die unklare Strategien oder einfache Heuristiken eingesetzt haben, in der Bedingung *ohne N tbt* am größten und in der Bedingung *ohne N sum* kleinsten. Diese Ergebnisse weisen darauf hin, dass die mentale Beanspruchung von der summarischen Bedingung ohne Notizen über die beiden Einzelfallbedingungen mit Notizen zur



Einzelfallbedingung ohne Notizen kontinuierlich anzusteigen scheint. Erneut kann die Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) zur Interpretation der Daten herangezogen werden, nach welcher die Probanden bei wachsender mentaler Beanspruchung vermehrt einfachere und ungenauere Strategien verwenden.

Der Prozentanteil der Probanden, die durchgängig von regel-basierten Strategien Gebrauch gemacht haben, lag mit 17,5 % nur unwesentlich höher als in Experiment 2, in dem 15 % ausschließlich von regel-basierten Strategien einsetzten. Weitere Teilanalysen ergaben, dass von den 14 Probanden, die durchgängig von einer regel-basierten Strategie Gebrauch machten, 6 die kausale Stärke entsprechend der Power PC Theorie berechnet haben, während nur 3 Probanden die  $\Delta P$ -Regel und immerhin 4 Probanden beide Strategien eingesetzt haben. Bei den Probanden, die nur teilweise von regel-basierter Strategie Gebrauch gemacht haben, ergab sich wie in Experiment 2 ein umgekehrtes Bild, hier wechselten 14 Probanden zur  $\Delta P$ -Regel und nur 8 zu einer Berechnung der kausalen Stärke entsprechend der Power PC Theorie. Auffällig war, dass über ein Viertel der Probanden (5 von 19), die in den beiden Bedingungen mit Notizen zumindest teilweise regel-basiert vorgegangen sind, sowohl entsprechend den Vorhersagen der  $\Delta P$ -Regel, als auch der Power PC Theorie geurteilt haben. Im Zusammenhang mit den Diskussionen in Kapitel 10.5 wurde darüber spekuliert, ob die Probanden in den beiden Bedingungen mit Notizen vermehrt von idiosynkratischen Strategien Gebrauch gemacht haben. Eine Analyse der Ergebnisse in Tabelle A6 in Anhang A3.3 ergab keine Hinweise für die Richtigkeit dieser Vermutung, lediglich unter der Bedingung *N strukturiert* machte eine Probandin von der selbst generierten Strategie "(c + d) / c \* a" Gebrauch. Die Annahme eines vermehrten Einsatzes idiosynkratischer Strategien in den beiden Bedingungen mit Notizen ergab sich aus dem Eindruck, dass sich die Probanden bei der Urteilsbildung sehr eingehend und zeitintensiv mit ihren eigenen Notizen auseinandergesetzt haben. Das Fehlen idiosynkratischer Strategien in den beiden Bedingungen mit Notizen muss diesem Eindruck, der vor allem durch die langen Bearbeitungszeiten (vor allem der *rating times*) bestätigt wird, nicht widersprechen.

Allgemein spricht die Tatsache, dass auch in Experiment 3 die meisten der 80 Probanden (41 plus 5 Probanden, die zwischen zwei regel-basierten Strategien gewechselt haben) von mehr als einer Strategie Gebrauch machten, erneut gegen die Annahme eines einheitlichen Verarbeitungsmechanismus bei der Beurteilung der Stärke von kausalen Zusammenhängen. In den beiden Bedingungen mit Notizen wurden immerhin 25 von 40 Probanden (62,5 %) als Wechsler kategorisiert. Dieser relativ hohe Anteil weist darauf hin, dass auch die Mehrheit der Probanden in den beiden Bedingungen mit Notizen versucht, ihre Strategien von Urteil zu

Urteil zu optimieren oder an die subjektiv wahrgenommene Aufgabenkomplexität anzupassen. Erste Eindrücke deuteten darauf hin, dass ein Teil der Probanden vor allem in den drei Einzelfallbedingungen bei den ersten Urteilen auf einfachere Heuristiken oder unklare Strategien zurückgriff und erst verzögert zur Anwendung einer regel-basierten Strategie gelangt, doch wurde diese Fragestellung anhand der vorliegenden Daten noch nicht untersucht. Die Annahme, dass zahlreiche Probanden darum bemüht sind, ihre Strategie schrittweise zu optimieren oder anzupassen, impliziert aus unserer Sicht, dass die Wahl von Strategien zur Beurteilung der Stärke kausaler Zusammenhänge auf Prozessen oder Mechanismen beruht, die nicht nur rein automatisiert ablaufen, sondern einer höheren kognitiven Ebene entspringen.

Die Häufigkeitsanalysen in Experiment 3 sprechen aus unserer Sicht dafür, dass die in Abhängigkeit von der Versuchsbedingung zu beobachtenden Unterschiede hinsichtlich der Strategiewahl eher kontinuierlicher Natur sind. Je höher die mentale Beanspruchung, desto geringer wird der Anteil an Probanden, die von einer regel-basierten Strategie Gebrauch machen und dadurch zu regel-konformen Lösungen gelangen. Bei einer geringen mentalen Beanspruchung wie z.B. unter der Bedingung *ohne N sum* setzt ein relativ großer Anteil von Probanden regel-basierte Strategien ein, aber es gibt auch bei geringer mentaler Beanspruchung Probanden, die von einfachen Heuristiken, sowie unklaren Strategien oder Schätzungen Gebrauch machen.

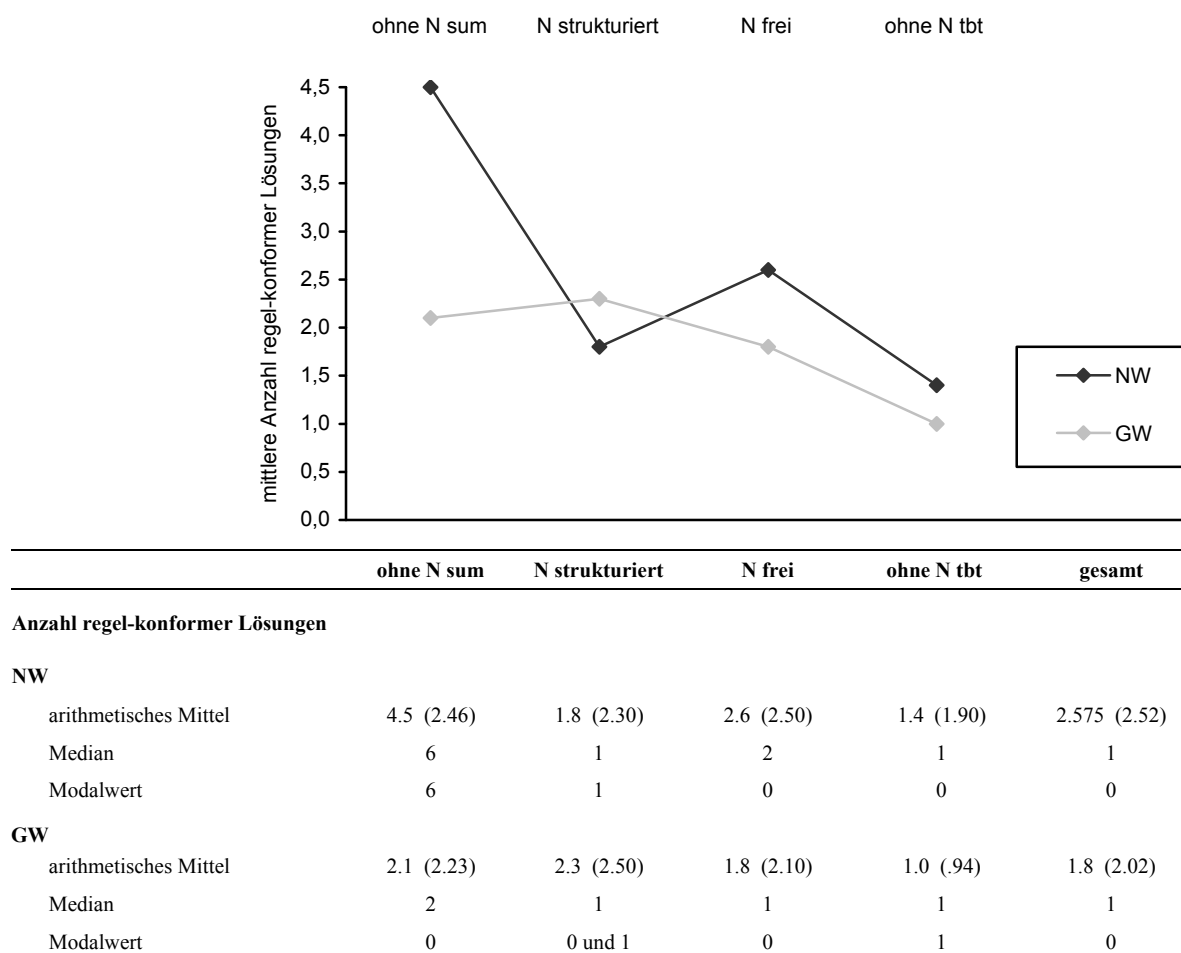
## 10.7 Einfluss von Studieneausrichtung und Unterschieden im numerischem Denken

### 10.7.1 Einfluss der Studieneausrichtung

#### 10.7.1.1 Ergebnisse

##### Anzahl regel-konformer Lösungen

Auch in Experiment 3 wurde vor Durchführung des Experiments erfasst, ob die Probanden in ihrem Studium eher mathematisch-naturwissenschaftlich oder eher geisteswissenschaftlich orientiert sind. Die Probanden wurden je nach Studieneausrichtung zu gleichen Teilen auf die vier Experimentalbedingungen zufallsverteilt. Abbildung 45 zeigt Maße der zentralen Tendenz für die *Anzahl regel-konformer Lösungen* getrennt nach *Studieneausrichtung*.



**Abbildung 45:** Experiment 3: Maße der zentralen Tendenz für die abhängige Variable *Anzahl regel-konformer Lösungen* (in Klammern Standardabweichungen) getrennt nach Studieneausrichtung in Abhängigkeit von *Notizen anfertigen*.

Anders als in Experiment 2 erzielten Probanden mit mathematisch-naturwissenschaftlicher Studiausrichtung nicht in allen Bedingungen höhere arithmetische Mittelwerte. Unter der Bedingung *N strukturiert* schneiden die Probanden ohne mathematisch-naturwissenschaftliche Studiausrichtung im Hinblick auf die abhängige Variable *Anzahl regel-konformer Lösungen* besser ab. Auffällig ist auch, dass nur bei Probanden mit mathematisch-naturwissenschaftlicher Studiausrichtung die Bedingung *ohne N sum* im Vergleich mit den drei anderen Bedingungen besser abschneidet. Unter Nichtberücksichtigung der Kovariate *numerisches Denken* wurde mit den Daten aus Experiment 3 eine zweifaktorielle Varianzanalyse mit den Faktoren *Studiausrichtung* und *Notizen anfertigen* gerechnet. Darüber hinaus wurden getrennt nach Studiausrichtung zwei einfaktorielle Varianzanalyse mit dem Faktor *Notizen anfertigen* durchgeführt. Tabelle 30 gibt einen Überblick über die varianzanalytischen Ergebnisse:

**Tabelle 30:** Experiment 3: Ergebnisse für die abhängige Variable *Anzahl regel-konformer Lösungen* aus der zweifaktoriellen Varianzanalyse mit den beiden Faktoren *Studiausrichtung* und *Notizen anfertigen*, sowie den beiden einfaktoriellen Varianzanalysen mit dem Faktor *Notizen anfertigen* für die beiden nach *Studiausrichtung* getrennten Probandengruppen. Signifikante P-Werte sind **fett** gedruckt.

Faktor	F-Wert	MSE	P-Wert
<b>Alle Probanden</b>			
Studiausrichtung (1,72)	2.548	4.715	.115
Notizen anfertigen (3,72)	3.156		<b>.030</b>
Studiausrichtung x Notizen anfertigen (3,72)	1.558		.207
<b>Pbn mit mathematisch-naturwissenschaftlicher Studiausrichtung</b>			
Notizen anfertigen (3,36)	3.575	5.303	<b>.023</b>
<b>Pbn ohne mathematisch-naturwissenschaftliche Studiausrichtung</b>			
Notizen anfertigen (3,36)	.791	4.128	.507

Im Gegensatz zu Experiment 2 verfehlte der Faktor *Studiausrichtung* das Signifikanzniveau von  $\alpha = 5\%$ , was zumindest teilweise darin begründet liegt, dass die Probanden ohne mathematisch-naturwissenschaftliche Studiausrichtung unter der Bedingung *N strukturiert* höhere arithmetische Mittelwerte erzielten als die Probanden mit mathematisch-naturwissenschaftlicher Studiausrichtung. Der Faktor *Notizen anfertigen* wurde nur über alle Probanden hinweg und für die Gruppe mit mathematisch-naturwissenschaftlicher Studiausrichtung signifikant. Für die Gruppe ohne mathematisch-naturwissenschaftliche Studiausrichtung wurde *Notizen anfertigen* klar nicht signifikant. Der wesentliche Grund für die Unterschiede zwischen den Probandengruppen ist wie schon in Experiment 2 ein stark

abweichendes Abschneiden unter der Bedingung *SD sum* bzw. *ohne N sum*, wo Probanden mit mathematisch-naturwissenschaftlicher Studieneausrichtung wesentlich höhere Werte erzielten als Probanden ohne mathematisch-naturwissenschaftliche Studieneausrichtung.

Für beide Probandengruppen und den Faktor *Notizen anfertigen* wurden a posteriori Einzelkontraste berechnet. Während in der Teilstichprobe der Probanden ohne mathematisch-naturwissenschaftliche Studieneausrichtung aufgrund des geringen Gesamteinflusses des Faktors *Notizen anfertigen* ( $p = .507$ ) erwartungsgemäß keiner der Kontraste signifikant wurde, wurde in der Teilstichprobe der Probanden mit mathematisch-naturwissenschaftlicher Studieneausrichtung der Kontrast *ohne N sum* versus *ohne N tbt* signifikant (gleiche Daten wie in Experiment 2), der Kontrast *ohne N sum* versus *N strukturiert* wurde in liberalen Verfahren (Least significant difference Test, Student-Newman-Keuls-Test) signifikant, in konservativeren Verfahren (Tukey honestly significant difference-Test, Scheffé-Test) wurde ein Trend erzielt. Für die Teilstichprobe der Probanden mit mathematisch-naturwissenschaftlicher Studieneausrichtung sind die Ergebnisse aus den Einzelkontrasten in den Tabellen 31 (LSD, Tukey HSD, Scheffé) und 32 (SNK-Test) zusammengefasst.

**Tabelle 31:** Experiment 3: Einzelkontraste (Least significant difference, Tukey HSD; Scheffé-Test) in der Teilstichprobe Probanden mit mathematisch-naturwissenschaftlicher Studieneausrichtung für die abhängige Variable *Anzahl regel-konformer Lösungen* in Abhängigkeit von *Notizen anfertigen*. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen.

Kontrast	LSD	Tukey HSD	Scheffé
ohne N sum vs. N strukturiert	<b>p = .013</b>	<u>p = .059</u>	<u>p = .095</u>
ohne N sum vs. N frei	<u>p = .073</u>	p = .270	p = .348
ohne N sum vs. ohne N tbt	<b>p = .005</b>	<b>p = .023</b>	<b>p = .042</b>
N strukturiert vs. N frei	p = .442	p = .864	p = .895
N strukturiert vs. ohne N tbt	p = .700	p = .980	p = .985
N frei vs. ohne N tbt	p = .252	p = .652	p = .717

**Tabelle 32:** Experiment 3: Einzelkontraste (Student-Newman-Keuls-Test) in der Teilstichprobe Probanden mit mathematisch-naturwissenschaftlicher Studieneausrichtung für die abhängige Variable *Anzahl regel-konformer Lösungen* in Abhängigkeit von *Notizen anfertigen*. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen.

	Bedingung	Zellgröße	Untergruppe für $\alpha = .05$ .	
			1	2
Student-Newman-Keuls-Test	ohne N tbt	N = 10	1.400	
	N strukturiert	N = 10	1.800	
	N frei	N = 10	2.600	2.600
	ohne N sum	N = 10		4.500
		<b>Signifikanz</b>	.481	<u>.073</u>

*Bearbeitungszeit und Verarbeitungseffizienz*

In gleicher Weise wie für die *Anzahl regel-konformer Lösungen* wurden Analysen mit den abhängigen Variablen *Bearbeitungszeit* und *Verarbeitungseffizienz* durchgeführt. Die Mittelwerte für die restlichen abhängigen Variablen, die in Tabelle 33 eingetragen sind, legen erneut nahe, dass naturwissenschaftlich-mathematisch ausgerichtete Probanden schneller und effizienter arbeiten.

**Tabelle 33:** Experiment 3: Bearbeitungszeit und Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von *Studienausrichtung* und *Notizen anfertigen*.

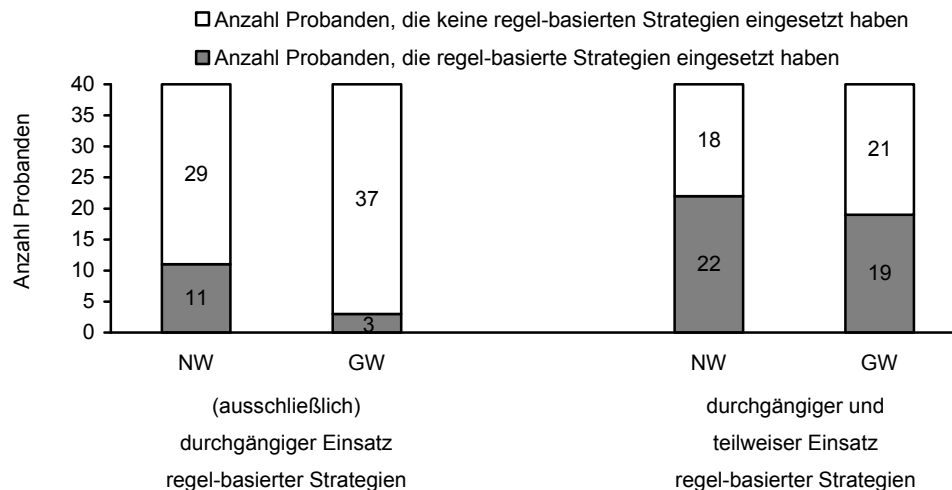
	ohne N sum		N strukturiert		N frei		ohne N tbt		Gesamt	
	NW	GW	NW	GW	NW	GW	NW	GW	NW	GW
<b>Bearbeitungszeit</b>	3 : 42 (1 : 14)	8 : 40 (5 : 51)	20 : 23 (3 : 39)	24 : 31 (5 : 49)	22 : 32 (7 : 08)	21 : 19 (6 : 52)	16 : 23 (7 : 38)	16 : 33 (4 : 23)	15 : 46 (9 : 08)	17 : 46 (8 : 13)
<b>Verarbeitungseffizienz</b>	1.32 (.89)	.40 (.55)	.08 (.10)	.09 (.10)	.14 (.14)	.08 (.08)	.09 (.13)	.07 (.08)	.41 (.69)	.16 (.31)

Die varianzanalytischen Vergleiche über alle Probanden hinweg (auf deren ausführliche Darstellung hier aus Platzgründen verzichtet wird, siehe Tabelle A7 im Anhang A3.3 mit einer Auflistung signifikanter und nicht signifikanter Faktoren und Interaktionen) zeigten, dass der Faktor *Notizen anfertigen* für beide abhängigen Variablen, der Faktor *Studienausrichtung* aber wie erwartet nur für die *Verarbeitungseffizienz* und nicht für die *Bearbeitungszeit* signifikant wurde. Zudem wurde der Einfluss der Interaktion *Studienausrichtung* x *Notizen anfertigen* auf die *Verarbeitungseffizienz* signifikant. Die signifikante Interaktion für die *Verarbeitungseffizienz* lässt sich bei näherer Betrachtung von Tabelle 33 damit erklären, dass sich beide Probandengruppen nur unter der Bedingung *ohne N sum* deutlich unterscheiden, nicht aber unter den anderen drei Bedingungen.

Weiter wurden auch für die *Bearbeitungszeit* und die *Verarbeitungseffizienz* getrennt nach Studienausrichtung und in Abhängigkeit von *Notizen anfertigen* Varianzanalysen ohne Hinzunahme der Kovariate *numerisches Denken* durchgeführt. Eine Auflistung der signifikant gewordenen Faktoren und Interaktionen befindet sich ebenfalls in Tabelle A7 im Anhang A3.3. Für beide Probandengruppen hatte dabei der Faktor *Notizen anfertigen* sowohl auf die *Bearbeitungszeit* als auch auf die *Verarbeitungseffizienz* einen signifikanten Einfluss, weil unabhängig von der Studienausrichtung alle Probanden unter der Bedingung *ohne N sum* wesentlich schneller arbeiteten als die Probanden in den drei Einzelfallbedingungen *N strukturiert*, *N frei* und *ohne N tbt*.

### Strategienwahl

Die varianzanalytischen Ergebnisse zur *Anzahl regel-konformer Lösungen* werden im Wesentlichen durch Häufigkeitsanalysen bestätigt: Abbildung 46 zeigt, dass die 40 mathematisch-naturwissenschaftlich ausgerichteten Probanden häufiger durchgängig oder zumindest teilweise regel-basierte Strategien eingesetzt haben als die 40 Probanden mit geisteswissenschaftlicher Studienausrichtung.



**Abbildung 46:** Experiment 3: Anzahl Probanden, die durchgängig oder zumindest teilweise regel-basierte Strategien in Abhängigkeit von *Studienausrichtung* eingesetzt haben.

Ein durchgeführter  $\chi^2$  – Test zeigt, dass mathematisch-naturwissenschaftlich ausgerichtete Probanden signifikant häufiger durchgängig von regel-basierten Strategien Gebrauch machten als Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung ( $\chi^2 = 5.541$ ,  $p = .019$ ; einseitiger P-Wert nach Fisher's exaktem Test = .018). Dieser signifikante Unterschied zwischen den verschiedenen Probandengruppen verschwindet allerdings vollständig, wenn zusätzlich diejenigen Probanden mit einbezogen wurden, die während der Aufgabenbearbeitung zu einer regel-basierten Strategie gewechselt haben. Hier ergaben weder ein  $\chi^2$ -Test, noch Fisher's exakter Test einen signifikanten Zusammenhang zwischen den beiden Variablen *teilweiser Einsatz regel-basierter Strategien* und *Strategietyp* ( $\chi^2 = .450$ ,  $p = .502$ ; einseitiger P-Wert nach Fisher's exaktem Test = .327).

#### 10.7.1.2 Diskussion

Die Ergebnisse aus den Varianzanalysen bestätigen für Experiment 3 nur die Hypothese H8b, die Hypothese H8a kann nicht uneingeschränkt beibehalten werden, auch wenn für die

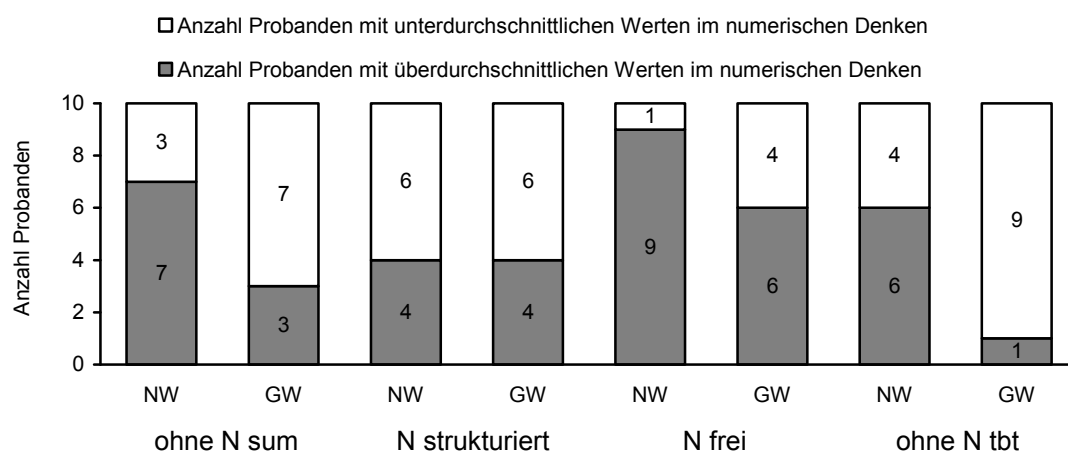
Mittelwertunterschiede hinsichtlich der Anzahl regel-konformer Lösungen immerhin Trendnähe erreicht wurde ( $p = .115$ ).

*Anzahl regel-konformer Lösungen:* Die ausbleibende Bestätigung für die Hypothese H8a ist vermutlich auf das unerwartet schlechte Abschneiden mathematisch-naturwissenschaftlich ausgerichteter Probanden unter der Bedingung *N strukturiert* zurückzuführen.

Unklar ist nun aber, warum die Probandengruppe mit mathematisch-naturwissenschaftlicher Studiausrichtung im Gegensatz zu den anderen drei Bedingungen unter der Bedingung *N strukturiert* schlechter abschneidet als die Probandengruppe ohne mathematisch-naturwissenschaftliche Studiausrichtung. In Experiment 2 konnte die Annahme bestätigt werden, dass die Unterschiede in Abhängigkeit von der *Studiausrichtung* wesentlich auf Unterschiede im numerischen Denken zurückgeführt werden können.

Eine Erklärung für das geringfügig schlechtere Abschneiden der Probandengruppe ohne mathematisch-naturwissenschaftliche Studiausrichtung unter der Bedingung *N strukturiert* könnte sein, dass das numerische Denken gerade unter dieser Bedingung keinen bedeutenden Einfluss auf die Anzahl regel-konformer Lösungen hat. Durch einen T-Test, der nur mit den Probanden aus der Bedingung *N strukturiert* durchgeführt wurde ( $MW_{NW} = 101.69$ ;  $MW_{GW} = 98.93$ ;  $T_{18} = 1.103$ ,  $p = .284$ ), konnte die Annahme bestätigt werden, dass zwischen den beiden nach Studiausrichtung getrennten Probandengruppen unter der Bedingung *N strukturiert* keine signifikanten Unterschiede im numerischen Denken bestehen.

Abbildung 47 veranschaulicht zudem, dass unter der Bedingung *N strukturiert* ebenso viele nicht mathematisch-naturwissenschaftlich wie mathematisch-naturwissenschaftlich ausgerichtete Probanden überdurchschnittliche Werte im numerischen Denken erreichten.



**Abbildung 47:** Experiment 3: Anzahl Probanden mit über- oder unterdurchschnittlichen Werten im numerischen Denken in Abhängigkeit von *Studiausrichtung* und *Notizen anfertigen*.



Abbildung 47 macht deutlich, dass in allen Bedingungen mit Ausnahme von *N strukturiert* der Anteil von Probanden mit überdurchschnittlichen Werten im numerischen Denken in der Fraktion der mathematisch-naturwissenschaftlich ausgerichteten Probanden höher ist als in der Gruppe von Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung. Darin kann ein weiterer Beleg für die in Experiment 2 vertretene Annahme gesehen werden, dass sich signifikante Unterschiede hinsichtlich der Anzahl regel-konformer Lösungen in Abhängigkeit von *Studienausrichtung* wesentlich auf Unterschiede im numerischen Denken zurückführen lassen. In Experiment 3 verfehlte der Faktor *Studienausrichtung* das Signifikanzniveau von  $\alpha < 5\%$  zwar relativ knapp, doch wurde wie in Experiment 2 für die Anzahl regel-konformer Lösungen eine zweifaktorielle Kovarianzanalyse durchgeführt. Wenn sich die etwas schwächeren Auswirkungen des Faktors *Studienausrichtung* auf die Anzahl regel-konformer Lösungen auch in Experiment 3 substantiell auf Unterschiede im numerischen Denken zurückführen lassen, dann sollte in einer zweifaktoriellen Kovarianzanalyse mit der Kovariate *numerisches Denken* und den beiden Faktoren *Studienausrichtung* und *Notizen anfertigen* der Einfluss des Faktors *Studienausrichtung* zusätzlich minimiert werden. Diese Annahme konnte wie schon in Experiment 2 bestätigt werden, der F-Wert für den Einfluss von *Studienausrichtung* auf die Anzahl regel-konformer Lösungen betrug lediglich .211 ( $p = .647$ ).

*Bearbeitungszeit und Verarbeitungseffizienz:* Der Einfluss des Faktors *Studienausrichtung* auf die Bearbeitungszeit wurde wie erwartet und bereits in Experiment 2 zu beobachten nicht signifikant.

Hinsichtlich der Verarbeitungseffizienz ergab sich eine signifikante Interaktion zwischen den Faktoren *Studienausrichtung* und *Notizen anfertigen*. Diese lässt sich darauf zurückführen, dass die Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung nur unter der Bedingung *ohne N sum* signifikant effizienter als die die Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung arbeiteten. In den beiden Bedingungen ohne Notizen und in der Bedingung *ohne N tbt* unterschieden sich die beiden Probandengruppen dagegen kaum. In Experiment 2 haben wir dieses Ergebnis dahingehend interpretiert, dass nur Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung von einer Informationsanordnung im Stapeldiagramm stark profitieren, während Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung dies nicht tun. Unter der Bedingung *N strukturiert* arbeiteten die Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung zwar schneller, aber nicht effizienter als die Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung, weil letztere eine etwas höhere

Anzahl regel-konformer Lösungen produzieren. Unter der Bedingung *N frei* ergaben sich ebenfalls keine wesentlichen Effizienz-Unterschiede zwischen den beiden Probandengruppen; die Probanden mit mathematisch-naturwissenschaftlicher Studiaausrichtung waren zwar etwas langsamer als die Probanden ohne mathematisch-naturwissenschaftliche Studiaausrichtung, aber sie erzielten eine etwas höhere Anzahl regel-konformer Lösungen. Die Frage, warum die Probanden mit mathematisch-naturwissenschaftlicher Studiaausrichtung unter den beiden Bedingungen mit Notizen unerwarteter Weise nicht effizienter arbeiteten als die Probanden ohne mathematisch-naturwissenschaftliche Studiaausrichtung kann nicht abschließend beantwortet werden, eine Rolle spielt sicherlich, dass die Probanden mit mathematisch-naturwissenschaftlicher Studiaausrichtung unter der Bedingung *N strukturiert* etwas weniger regel-konforme Lösungen als die Probanden ohne mathematisch-naturwissenschaftliche Studiaausrichtung produziert haben und sich beide Probandengruppen nicht wesentlich hinsichtlich der Anzahl ihrer Werte im numerischen Denken unterschieden haben.

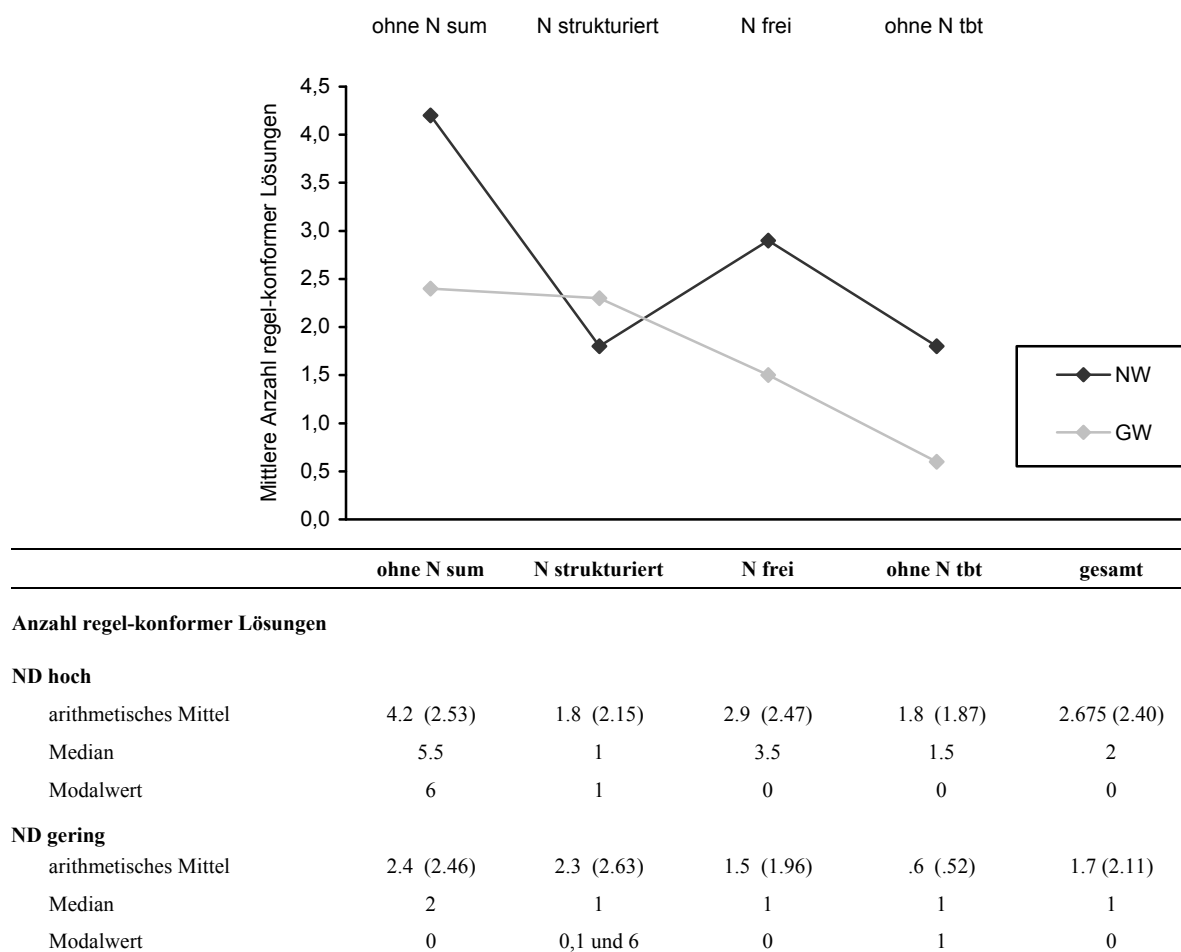
*Strategienwahl:* Im Gegensatz zu Experiment 2 war in der Gruppe mit mathematisch-naturwissenschaftlicher Studiaausrichtung die Anzahl von Probanden, die durchgängig von regel-basierten Strategien Gebrauch machte, signifikant höher als in der Gruppe der Probanden ohne mathematisch-naturwissenschaftliche Studiaausrichtung. Somit wäre die für Experiment 2 formulierte Hypothese H8c im Zusammenhang mit einem durchgängigen Einsatz regel-basierter Strategien gültig gewesen, was möglicherweise mit dem aus Experiment 2 bekannten Ergebnis zusammenhängt, dass naturwissenschaftlich ausgerichtete Studenten besser von visuell hoch strukturierten Grafiken profitierten, als Probanden ohne mathematisch-naturwissenschaftliche Studiaausrichtung. Diese signifikanten Unterschiede verschwanden aber, wenn zusätzlich zur Menge der Probanden, die durchgängig von regel-basierten Strategien Gebrauch machten, diejenigen Probanden mit hinzugenommen wurden, die lediglich teilweise von regel-basierten Strategien Gebrauch machten. Das Verschwinden eines Unterschiedes zwischen den beiden Probandengruppen mit und ohne naturwissenschaftliche Studiaausrichtung bei Hinzunahme von Probanden die nur teilweise regel-basiert vorgehen, spricht dafür, dass Probanden ohne mathematisch-naturwissenschaftliche Studiaausrichtung entweder verzögert zum Gebrauch einer regel-basierten Strategie gelangen, oder dass sie häufiger von einer regel-basierten zu einer einfacheren oder unklaren Strategie wechseln.

## 10.7.2 Einfluss von Unterschieden im numerischen Denken

### 10.7.2.1 Ergebnisse

#### Anzahl regel-konformer Lösungen

Wie in Experiment 2, so sollte auch in Experiment 3 die Annahme untersucht werden, ob der Faktor *Notizen anfertigen* bei Probanden mit hohen und geringen Intelligenzwerten im numerischen Denken einen unterschiedlichen Einfluss auf die *Anzahl regel-konformer Lösungen* hat. Zu diesem Zweck wurden die 20 Probanden in jeder der vier Experimentalbedingungen *ohne N sum*, *N strukturiert*, *N frei* und *ohne N tbt* in je 10 Probanden mit hohen und je 10 Probanden mit geringen Intelligenzwerten im numerischen Denken unterteilt. Über alle Experimentalbedingungen hinweg ergaben sich also wie in Experiment 2 zwei Teilstichproben á 40 Probanden, die sich hinsichtlich ihrer erreichten Werte im numerischen Denken unterscheiden (*hoch* vs. *gering*). Getrennt für die beiden Teilstichproben wurden Maße der zentralen Tendenz für die *Anzahl regel-konformer Lösungen* bestimmt. Sie sind in Abbildung 48 eingetragen.



**Abbildung 48:** Experiment 3: Maße der zentralen Tendenz für *Anzahl regel-konformer Lösungen* bei Probanden mit hohen und geringen Werten im *numerischen Denken* (ND hoch vs. ND gering) in Abhängigkeit von *Notizen anfertigen*.

Die Ergebnisse in Abbildung 48 und Abbildung 45 weisen darauf hin, dass Intelligenzunterschiede im numerischen Denken und der Faktor *Studienausrichtung* einen sehr ähnlichen Einfluss auf die *Anzahl regel-konformer Lösungen* haben. So wie Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung in allen Bedingungen mit Ausnahme der Bedingung *N strukturiert* besser abgeschnitten haben als Probanden ohne mathematisch-naturwissenschaftliche Studienausrichtung, so erzielten auch Probanden mit hohem im Vergleich zu Probanden mit geringen Testwerten im numerischen Denken in allen Bedingungen mit Ausnahme der Bedingung *N strukturiert* durchschnittlich eine höhere *Anzahl regel-konformer Lösungen*. Darüber hinaus legen die Mittelwerte für die *Anzahl regel-konformer Lösungen* nahe, dass der Einfluss des Faktors *Notizen anfertigen* in den beiden Gruppen *ND hoch* und *ND gering* etwas ähnlicher ausfallen dürfte als noch bei den beiden Probandengruppen mit unterschiedlicher Studienausrichtung. Zur Überprüfung des Einflusses beider Probandengruppen *ND hoch* und *ND gering* auf die *Anzahl regel-konformer Lösungen* wurde unter Nichtberücksichtigung der Kovariate *numerisches Denken* mit den Daten aus Experiment 3 eine zweifaktorielle Varianzanalyse mit den Faktoren *numerisches Denken* (*ND hoch* versus *ND gering*) und *Notizen anfertigen* durchgeführt. Darüber hinaus wurden für jede Probandengruppe getrennt zwei einfaktorische Varianzanalysen mit dem Faktor *Notizen anfertigen* durchgeführt. Tabelle 34 gibt einen Überblick über die Ergebnisse:

**Tabelle 34:** Experiment 3: Ergebnisse für die abhängige Variable *Anzahl regel-konformer Lösungen* aus der 2 (*numerisches Denken*) x 4 (*Notizen anfertigen*) Varianzanalyse, sowie den beiden einfaktorischen Varianzanalysen für die Probandengruppen *ND hoch* und *ND gering* mit dem Faktor *Notizen anfertigen*. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen.

Faktor	F-Wert	MSE	P-Wert
<b>Alle Probanden</b>			
numerisches Denken (1,72)	4.037	4.710	<b>.048</b>
Notizen anfertigen (3,72)	3.159		<b>.030</b>
numerisches Denken x Notizen anfertigen (3,72)	1.093		.358
<b>Pbn mit hohen Werten im numerischen Denken</b>			
Notizen anfertigen (3,36)	2.525	5.158	<u>.073</u>
<b>Pbn mit geringen Werten im numerischen Denken</b>			
Notizen anfertigen (3,36)	1.643	4.261	.197

Tabelle 34 zeigt zwei wesentliche Ergebnisse: Zum einen wurde der Faktor *numerisches Denken* trotz gegenläufiger Ergebnisse unter der Bedingung *N strukturiert* signifikant, zum anderen wurde der Faktor *Notizen anfertigen* wegen größerer Zellbesetzungen nur in der Gesamtstichprobe, nicht aber in den beiden Teilstichproben signifikant. In der Gruppe *ND*

hoch erreichte der Faktor *Notizen anfertigen* aber immerhin einen Trend. Da sowohl die Interaktion *Teilstichprobe x Notizen anfertigen* in der Gesamtstichprobe, als auch der Faktor *Notizen anfertigen* in einer der beiden Teilstichproben das Signifikanzniveau von  $\alpha = 5\%$  verfehlten, wurde auf die Berechnung von a posteriori Einzelkontrasten verzichtet.

#### *Bearbeitungszeit und Verarbeitungseffizienz*

Die Auswirkungen des zweistufigen Faktors *numerisches Denken* wurden auch für die *Bearbeitungszeit* und die *Verarbeitungseffizienz*. Die Mittelwerte, die in Tabelle 35 eingetragen sind, zeigen, dass Probanden mit höheren Werten im numerischen Denken die Aufgaben schneller und effizienter bearbeiten.

**Tabelle 35:** Experiment 3: Bearbeitungszeit und Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von *numerisches Denken* und *Notizen anfertigen*.

	ohne N sum		N strukturiert		N frei		ohne N tbt		Gesamt	
	ND hoch	ND gering	ND hoch	ND gering	ND hoch	ND gering	ND hoch	ND gering	ND hoch	ND gering
<b>Bearbeitungszeit</b>	4 : 23 (2 : 35)	7 : 58 (5 : 58)	20 : 32 (4 : 05)	24 : 24 (5 : 37)	18 : 48 (5 : 05)	25 : 03 (7 : 10)	14 : 52 (2 : 50)	18 : 04 (7 : 59)	<b>14 : 39</b> <b>(7 : 19)</b>	<b>18 : 52</b> <b>(9 : 30)</b>
<b>Verarbeitungseffizienz</b>	1.21 (.97)	.51 (.60)	.08 (.09)	.09 (.11)	.16 (.14)	.06 (.07)	.12 (.13)	.04 (.04)	<b>.39</b> <b>(.67)</b>	<b>.17</b> <b>(.36)</b>

In den durchgeführten varianzanalytischen Vergleichen über alle Probanden hinweg (ausführliche Darstellung siehe Tabelle A8 im Anhang A3.3) hatten sowohl *Notizen anfertigen* als auch *numerisches Denken* einen signifikanten Einfluss auf beide abhängigen Variablen. Für die *Verarbeitungseffizienz* wurde zudem noch die Interaktion *numerisches Denken x Notizen anfertigen* signifikant. Getrennte Varianzanalysen für Probanden mit hohen oder geringen Testwerten im numerischen Denken ergaben, dass der Faktor *Notizen anfertigen* in beiden Probandengruppen einen signifikanten Einfluss auf die *Bearbeitungszeit* und die *Verarbeitungseffizienz* hat.

#### 10.7.2.2 Diskussion

*Anzahl regel-konformer Lösungen:* Insgesamt wird für Experiment 3 die Hypothese H7a also nicht nur durch die Kovarianzanalysen mit der signifikanten Kovariate *numerisches Denken* bestätigt, sondern auch durch Varianzanalysen mit einem signifikanten Faktor *numerisches Denken*. In allen Bedingungen mit Ausnahme der Bedingung *N strukturiert* erzielten Probanden mit hohen Werten im numerischen Denken eine höhere mittlere Anzahl regel-konformer Lösungen als Probanden mit geringen Werten im numerischen Denken.

Zusätzlich durchgeführte T-Tests belegen aber, dass der Einfluss des Faktors *numerisches Denken* auf die Anzahl regel-konformer Lösungen nur unter den beiden Bedingungen *ohne N sum* und *ohne N tbt* signifikant wird ( $MW_{ND\ hoch} = 3.0$ ;  $MW_{ND\ gering} = 1.5$ ,  $T_{38} = 2.126$ ,  $p = .041$ ), nicht aber in den beiden Bedingungen *N strukturiert* und *N frei* ( $MW_{ND\ hoch} = 2.35$ ;  $MW_{ND\ gering} = 1.9$ ,  $T_{38} = .617$ ,  $p = .541$ ).

Weshalb aber hat der Faktor *numerisches Denken* in den beiden Bedingungen mit Notizen keinen signifikanten Einfluss? Eine Erklärung liegt darin, dass der Einfluss der numerischen Verarbeitungskapazität durch die Notizen und die damit verbundene Gedächtnisentlastung zurückgeht. Einzelne Probanden mit geringerer numerischer Verarbeitungskapazität könnten durch das Anfertigen von Notizen weniger überfordert gewesen sein und vermehrt von regelbasierten Strategien Gebrauch gemacht haben. Zudem besteht die Möglichkeit, dass unter den beiden Bedingungen mit Notizen andere Faktoren größeres Gewicht erhalten. Ein möglicher Faktor könnte die Motivation sein, sich mit der Aufgabe und insbesondere den eigenen Notizen auseinanderzusetzen. Diese Motivation wurde von uns nicht unmittelbar erfasst. Sie kann aber indirekt über die abhängige Variable *rating time* erschlossen werden. Unter *rating time* ist die Zeitdauer zu verstehen, die die Probanden nach der Datensammlung zum eigentlichen Einschätzen der Wirksamkeit benötigten. Was durch die *rating time* angezeigt wird, ist unklar, neben motivationalen Faktoren wie z.B. Ausdauer, Gewissenhaftigkeit, Sorgfalt, Interesse an der Aufgabenstellung, Anerkennungs- und Leistungsmotive spielen sicherlich auch kognitive Voraussetzungen (Aufmerksamkeitsspanne, Merkfähigkeit etc.) eine Rolle.

Benötigen die Probanden zum Einschätzen der Wirksamkeit an Hand der angefertigten Notizen viel Zeit, so kann davon ausgegangen werden, dass sie sich intensiv oder hoch motiviert mit der Beurteilung der Wirksamkeit auseinander gesetzt haben. Nehmen sich die Probanden hingegen wenig Zeit für die Beurteilung der Wirksamkeiten, so kann zumindest teilweise angenommen werden, dass sie dabei weniger motiviert vorgegangen sind. Wenn zudem davon ausgegangen wird, dass hoch motivierte Probanden eine höhere Anzahl regelkonformer Lösungen erzielen als gering motivierte Probanden, dann sollten unter den beiden Bedingungen *N strukturiert* und *N frei* Probanden mit hoher *rating time* eine höhere mittlere Anzahl regelkonformer Lösungen erzielen als Probanden mit kurzer *rating time*. Um diese Annahme zu untersuchen wurden die 20 Probanden in jeder der vier Experimentalbedingungen *ohne N sum*, *N strukturiert*, *N frei* und *ohne N tbt* in je zwei gleich große Gruppen aufgeteilt. In der einen Gruppe befanden sich jeweils 10 Probanden mit kurzer, in der anderen Gruppe jeweils 10 Probanden mit langer *rating time*. Über alle

Experimentalbedingungen hinweg ergaben sich in Abhängigkeit vom Ausmaß der rating time (Variablenname: *rating time*) zwei Gruppen á 40 Probanden (*lang* versus *kurz*). Auch wenn der Einfluss des zweistufigen Faktors *rating time* über alle Bedingungen hinweg nicht signifikant wurde ( $MW_{RT\ lang} = 2.475$ ;  $MW_{RT\ kurz} = 1.9$ ,  $T_{78} = 1.119$ ,  $p = .266$ ), so konnte das Signifikanzniveau von  $\alpha = 2.5\%$  hinsichtlich des Einflusses von *rating time* klar unterschritten werden, wenn nur die beiden Bedingungen *N strukturiert* und *N frei* berücksichtigt wurden ( $MW_{RT\ lang} = 3.15$ ;  $MW_{RT\ kurz} = 1.1$ ,  $T_{38} = 3.136$ ,  $p < .004$ ). Offensichtlich ist der Faktor *rating time* für die beiden Bedingungen mit Notizen bedeutsamer als der Faktor *numerisches Denken*.

Zusätzlich durchgeführte Kovarianzanalysen, in denen der Einfluss der beiden Bedingungen mit Notizen (*N strukturiert* und *N frei*) und der beiden Bedingungen ohne Notizen (*ohne N sum* und *ohne N tbt*) auf die abhängige Variable *Anzahl regel-konformer Lösungen* getrennt untersucht wurde, zeigten ebenfalls, dass die Kovariate *rating time* nur in den beiden Bedingungen mit Notizen einen hochsignifikanten Einfluss auf die Anzahl regel-konformer Lösungen hat ( $F_{1,37} = 9.296$ ,  $MSE = 4.410$ ,  $p = .004$ ). In den beiden Bedingungen ohne Notizen hatte *rating time* hingegen keinen Einfluss auf die Anzahl regel-konformer Lösungen ( $F_{1,37} = .038$ ,  $MSE = 6.614$ ,  $p = .847$ ). Für die Kovariate *numerisches Denken* verhielt es sich dagegen genau umgekehrt, sie hatte nur in den beiden Bedingungen ohne Notizen einen hochsignifikanten Einfluss auf die Anzahl regel-konformer Lösungen ( $F_{1,37} = 18.833$ ,  $MSE = 3.034$ ,  $p < .001$ ). Wenn die Probanden wie in den Bedingungen *N strukturiert* und *N frei* Notizen anfertigen, so verschwindet der Einfluss der Kovariate *numerisches Denken* fast völlig ( $F_{1,37} = .123$ ,  $MSE = 5.499$ ,  $p < .728$ ).

Inhaltlich sprechen die Ergebnisse mit *rating time* als unabhängiger Variable und *rating time* als Kovariate dafür, dass unter den beiden Einzelfallbedingungen mit Notizen diejenigen Probanden eine höhere Anzahl regel-konformer Lösungen erreichen, die sich intensiver und länger mit der Einschätzung oder Berechnung der Wirksamkeit einer Substanz auf der Basis der gesammelten Informationen über Ursache-Wirkungs-Zusammenhänge befassen. Unterschiedliche Beschäftigungszeiten mit dem Einschätzen oder Berechnen der Wirksamkeiten an Hand eigener Notizen können unspezifisch als Indikatoren für motivationale Unterschiede aufgefasst werden, die von uns nicht direkt via Fragebögen erfasst wurden. Somit steht eine Validierung der Annahme, dass die *rating time* primär durch motivationale Faktoren bestimmt wird, noch aus.

Die Fertigkeiten der Probanden im numerischen Denken spielten in den beiden Bedingungen *N strukturiert* und *N frei* keine bedeutende Rolle, sie sind im Hinblick auf die Anzahl regel-

konformer Lösungen nur dann vorhersagekräftig, wenn die Probanden wie in Experiment 2 oder unter den Bedingungen *ohne N sum* und *ohne N tbt* aus Experiment 3 die Wirksamkeiten im Kopf und nicht mit Hilfe eigener Notizen ermitteln müssen.

Der unterschiedliche Einfluss der Kovariaten *rating time* und *numerisches Denken* veranschaulicht auch, dass individuelle Unterschiede in Abhängigkeit von der Aufgabenstellung unterschiedliche Auswirkungen haben können. Sollen die Probanden zur Unterstützung bei der Beurteilung der Stärke von einzelfallweise dargebotenen kausalen Zusammenhängen Notizen anfertigen, so geht der Einfluss des Faktors *numerisches Denken* zurück, während motivationale Faktoren, mutmaßlich erfasst durch die *rating time*, an Bedeutung zu gewinnen scheinen. Stehen die Probanden dagegen vor der Aufgabe, die Stärke von einzelfallweise und summarisch dargebotenen Zusammenhängen ohne Notizen im Kopf zu beurteilen, so spielt die Intensität, mit der sich die Probanden zeitlich mit der Urteilsbildung auseinandersetzen im Vergleich zu den individuellen Voraussetzungen im numerischen Denken eine wesentlich geringere Rolle.

Insgesamt konnte die Hypothese H7a über alle Probanden hinweg bestätigt werden. Probanden, die höhere Testwerte im numerischen Denken erzielten, produzierten durchschnittlich eine höhere mittlere Anzahl regel-konformer Lösungen. Bei näherer Betrachtung der Ergebnisse zeigte sich aber, dass die Hypothese H7a nur für Probanden in den beiden Bedingungen ohne Notizen (*ohne N sum* und *ohne N tbt*) Gültigkeit hat. In den beiden Bedingungen mit Notizen (*N strukturiert* und *N frei*) unterschieden sich die beiden Probandengruppen mit hohen und geringen Werten im numerischen Denken nicht signifikant hinsichtlich der mittleren Anzahl regel-konformer Lösungen voneinander. Stattdessen konnten die Ergebnisse der Probanden unter den beiden Bedingungen *N strukturiert* und *N frei* besser durch die Kovariate *rating time* vorhergesagt werden.

*Bearbeitungszeit und Verarbeitungseffizienz:* Tabelle 35 veranschaulicht, dass die Probanden mit hohen Werten im numerischen Denken konstant in allen vier Bedingungen schneller waren als die Probanden mit geringen Werten. Dieses Ergebnis war von uns nicht vorhergesagt worden, bestätigt aber die in Kapitel 10.5 auf Seite 293 formulierte Vermutung, dass Probanden mit hohen Werten im numerischen Denken auch in den beiden Bedingungen *N strukturiert* und *N frei* weniger Zeit für die eigentliche Aufgabenbearbeitung benötigt haben.

Fasst man die Ergebnisse aus den Experimenten 2 und 3 zum Einfluss des *numerischen Denkens* auf die Bearbeitungszeit zusammen, dann kann festgestellt werden, dass Probanden mit hohen Werten im numerischen Denken von einer hohen visuellen Strukturierung des



darstellenden Mediums zeitlich mehr profitieren als Probanden mit geringen Werten im numerischen Denken. Bei geringer visueller Strukturierung des darstellenden Mediums (Fliesstexte, "Textaufgaben") sind die Probanden mit hohen Werten im numerischen Denken sogar etwas langsamer als die Probanden mit geringen Werten im numerischen Denken.

Die hier beschriebene Wechselwirkung zwischen intellektuellen Voraussetzungen (Werte im *numerischen Denken*) auf der Individuenseite und medialen Rahmenbedingungen (*visuelle Strukturierung des darstellenden Mediums*) bei der Aufgabenbearbeitung verdeutlicht erneut, dass es nicht genügt, den Einfluss medialer Rahmenbedingungen unabhängig von individuellen Voraussetzungen untersuchen zu wollen. In Verbindung mit der abhängigen Variable *Anzahl regel-konformer Lösungen* wurden auch schon in Experiment 1 (z.B. signifikanter Einfluss des Faktors *Informationsart* nur bei Leibniz-Kollegiaten, nicht bei Studenten), in Experiment 2 (z.B. signifikanter Einfluss des Faktors *visuelle Strukturierung* nur bei Probanden mit hohen Werten im numerischen Denken, nicht bei Probanden mit geringen Werten) und in Experiment 3 (signifikanter Einfluss der Kovariate *rating time* nur in den beiden Bedingungen mit Notizen, signifikanter Einfluss der Kovariate *numerisches Denken* nur in den beiden Bedingungen ohne Notizen) ähnliche Entdeckungen gemacht.

Der signifikante Einfluss der zweistufigen Variable *numerisches Denken* auf die Verarbeitungseffizienz entsprach den Vorhersagen und bestätigt damit die Hypothese H7b. Probanden mit hohen Werten arbeiteten effizienter als Probanden mit geringen Werten im numerischen Denken, weil sie mehr regel-konforme Lösungen produzierten und dafür weniger Zeit benötigten. Die im Zusammenhang mit der Verarbeitungseffizienz signifikante Interaktion *Notizen anfertigen* x *numerisches Denken* beruht vermutlich darauf, dass die Unterschiede zwischen den Probandengruppen *ND gering* und *ND hoch* in der summarischen Bedingung *ohne N sum* besonders stark ausgeprägt waren und in den drei Einzelfallbedingungen *N strukturiert*, *N frei* und *ohne N tbt* dagegen kaum.

## 10.8 Abschließende Diskussion von Experiment 3

In Experiment 3 sollte primär die Fragestellung untersucht werden, warum die Probanden bei summarischer Darbietung mehr regel-konforme Lösungen hervorbringen als bei einzelfallweiser Darbietung. Insbesondere sollte die Rolle, welche die Beanspruchung des Arbeitsgedächtnisses dabei spielt, näher untersucht werden. Um das Arbeitsgedächtnis auch bei einer einzelfallweisen Darbietung entlasten zu können, wurde einem Teil der Probanden die Aufgabe erteilt, Notizen anzufertigen. Der andere Teil der Probanden fungierte als Kontrollgruppe, hier erhielten die Probanden keine Möglichkeit, Notizen anzufertigen und bekamen die Informationen entweder einzelfallweise in einer Vierfeldertafel mit Positionskreuz oder summarisch in einem Stapeldiagramm dargeboten.

Die Ergebnisse zeigten, dass das Gefälle zwischen einzelfallweiser und summarischer Darbietung hinsichtlich der Anzahl regel-konformer Lösungen nur teilweise durch das Anfertigen von Notizen bei einzelfallweiser Darbietung und die damit verbundene Entlastung des Arbeitsgedächtnisses reduziert werden kann. Eine Entlastung des Arbeitsgedächtnisses durch das Anfertigen von Notizen kann lediglich als Erklärung für den Anstieg der Anzahl regel-konformer Lösungen zwischen der Bedingung *ohne N tbt* und den beiden Einzelfallbedingungen mit Notizen heran gezogen werden. Um den Anstieg zwischen den beiden Einzelfallbedingungen mit Notizen und der Bedingung *ohne N sum* erklären zu können, sind wir, ausgehend von der Diskussion in der Untersuchung von Ward & Jenkins (1965), zu der Schlussfolgerung gelangt, dass die meisten Probanden von einer einzelfallweisen Darbietung mit anschließender Ergebniszusammenfassung mental immer noch so stark beansprucht werden, dass sie nicht in der Lage sind, von einer hohen visuellen Strukturierung der vorliegenden Informationsanordnung zu profitieren und eine gleich hohe Anzahl regel-konformer Lösungen zu produzieren wie die Probanden unter der Bedingung *ohne N sum*.

Einen möglichen Grund für die immer noch hohe mentale Beanspruchung der Probanden in den Bedingungen mit Notizen und einzelfallweiser Darbietung haben wir darin gesehen, dass die Aufgabenbeschaffenheit bei einzelfallweiser Darbietung mit Notizen immer noch eine andere ist als bei einer rein summarischen Darbietung. Im Gegensatz zur summarischen Darbietung müssen die Probanden bei einer einzelfallweisen Darbietung mit Notizen während der eigentlichen Aufgabenbearbeitung (*time on task*) einfache, aber zeitaufwendige Handlungsrountinen (Ordnen und Sammeln von Daten, Notizen anfertigen) ausführen, welche die Aufmerksamkeit beanspruchen. Durch die Fokussierung auf einfache Handlungsrountinen

fehlt den Probanden der Überblick, um Vergleiche vorzunehmen, Berechnungen anzustellen und Schlussfolgerungen zu ziehen. Wegen des fehlenden Gesamtüberblicks und der Fokussierung auf die Einzelfälle ist die mentale Beanspruchung bei einer einzelfallweisen Darbietung mit oder ohne Notizen immer noch so hoch, dass die Zusammenhänge zwischen den verschiedenen Zellohäufigkeiten nicht so salient gemacht werden können wie bei einer rein summarischen Informationsdarbietung, bei welcher sich die Probanden schon während der *time on task* ganz auf die eigentliche Urteilsbildung konzentrieren können. Möglich ist auch, dass die Probanden auf der Basis von Kosten-Nutzen-Erwägungen auf weniger komplexe Strategien bei der kausalen Urteilsbildung zurückgreifen (Kleinmuntz & Schkade, 1993). Letzteres könnte als ein Indiz dafür verstanden werden, dass bei der Beurteilung der Stärke von monokausalen Zusammenhängen Entscheidungsprozesse mitbeteiligt sind.

Aus den Ergebnissen von Experiment 3 ergibt sich die Schlussfolgerung, dass eine rein summarische Darbietung von Informationen über die Stärke von Ursache-Wirkungs-Zusammenhängen nicht mit einer summarischen Ergebniszusammenfassung im Anschluss an eine einzelfallweise Informationsdarbietung gleichgesetzt werden kann. Wir gehen im Einklang mit der Forschung zu den Auswirkungen externaler Repräsentationen (Larkin & Simon, 1987; Stenning & Oberlander, 1994; Zhang, 2000, 1997; Zhang, Johnson, Malin & Smith, 2002; Zhang & Norman, 1994) davon aus, dass eine Variation medialer Bedingungen (hier: Variation der Darbietungsart) aufgabenbezogene Veränderungen nach sich ziehen können (z.B. den Einsatz zusätzlicher Handlungsroutinen bei einzelfallweiser Darbietung), die wiederum Einfluss auf die Aufgabenkomplexität und damit auch auf die mentale Beanspruchung der Probanden haben.

Die *Häufigkeitsanalysen* haben gezeigt, dass der Anteil von Probanden, die von regelbasierten Strategien teilweise oder durchgängig Gebrauch machen, bei zunehmender mentaler Beanspruchung kontinuierlich abnimmt und bestätigen damit die Ergebnisse im Zusammenhang mit der Anzahl regel-konformer Lösungen. Erneut konnten keine Belege dafür gefunden werden, dass die Probanden eine einheitliche Strategiewahl vornehmen. Stattdessen wechselt über die Hälfte der Probanden über alle Bedingungen hinweg mehrfach die Strategie, vermutlich um ihre Strategie von Urteil zu Urteil zu optimieren oder um sie an die subjektiv wahrgenommene Aufgabenkomplexität anzupassen (Kleinmuntz & Schkade, 1993).

Die Fähigkeiten der Probanden im *numerischen Denken* hatten über alle Bedingungen hinweg erneut einen signifikanten Einfluss auf die Anzahl regel-konformer Strategien, die Bearbeitungszeit und die Verarbeitungseffizienz. Der von uns nicht vorhergesagte

signifikante Einfluss auf die Bearbeitungszeiten ergab sich vor allem aus einer in allen vier Bedingungen schnelleren Aufgabenbearbeitung der Probanden mit hohen Werten im numerischen Denken während der *time on task*. Dieser Zeitvorsprung kam vermutlich dadurch zu Stande, dass die Probanden mit hohen Werten im numerischen Denken die Daten aus Stapeldiagramm oder Vierfeldertafel mit Positionskreuz schneller entnehmen und verarbeiten konnten, eventuell konnten sie ihre Notizen in den beiden Bedingungen mit Notizen auch schneller erstellen und strukturieren. In Experiment 2 hatte das numerische Denken keinen signifikanten Einfluss auf die Bearbeitungszeit, da die Probanden mit hohen Werten im numerischen Denken insbesondere in den beiden Bedingungen mit Fliesstexten zum Teil langsamer gearbeitet haben als die Probanden mit geringen Werten im numerischen Denken. In Verbindung mit der Anzahl regel-konformer Lösungen ergaben weitere Analysen, dass der Einfluss des numerischen Denkens vor allem dann groß ist, wenn die Probanden keine Notizen anfertigen müssen. In den beiden Bedingungen mit Notizen fallen dagegen andere, vermutlich motivationale, Faktoren stärker ins Gewicht.

Die Ergebnisse in Experiment 3 stehen wie schon die Ergebnisse aus Experiment 2 im Einklang mit dem Schaubild in Abbildung 5 (S. 132) und zeigen erneut, dass individuelle Unterschiede bedeutsam sind und mit medien- und aufgabenbezogenen Einflüssen interagieren. Die mittlere Anzahl regel-konformer Lösungen pro Proband und die Anzahl der Probanden, die bei der Beurteilung der Stärke von kausalen Zusammenhängen von regel-basierten Strategien Gebrauch machen, hängen von situativen Gegebenheiten und personenbezogenen Faktoren ab. Personenbezogene, mediale und aufgabenbezogene Einflüsse sind ineinander verflochten, die Auswirkungen dieser Einflüsse auf die Informationsverarbeitung im Allgemeinen und die hier untersuchten Prozesse bei der Beurteilung der Stärke von kausalen Zusammenhängen im Besonderen verändern sich in Abhängigkeit von der jeweiligen Zusammensetzung oder Komposition dieser drei Faktorenkomplexe.

Teil 3

**Schlussfolgerungen**



## 11 Zusammenfassende Diskussion

In Kapitel 3.4 haben wir Kritik an monomechanistischen Theorieansätzen zur Erklärung der Prozesse bei der Beurteilung der Stärke kausaler Zusammenhänge geübt und sind dabei zu dem Ergebnis gelangt, dass assoziative und regel-basierte Mechanismen einander nicht ausschließen müssen. Gemeinsam mit Baker, Murphy & Vallée-Tourangeau, (1996), sowie Catena & Perales (2002) nehmen wir einen evolutionären Standpunkt ein und vermuten, dass bei der Verarbeitung von Informationen zu kovariierenden oder kausal verknüpften Ereignissen verschiedene Systeme mit unterschiedlichen Mechanismen beteiligt sind. Im Rahmen dieser Arbeit haben wir uns eingehend mit verschiedenen regel-basierten Ansätzen wie dem *Probabilistischen Kontrastmodell* (Cheng & Novick, 1990, 1991, 1992) und der *Theorie der kausalen Power* (Cheng, 1997) befasst, bei denen es sich um Weiterentwicklungen von *kontingenzttheoretischen Ansätzen* handelt. Bei der Durchsicht der Literatur zum Erwerb von Kausalwissen sind wir zu dem Ergebnis gelangt, dass die häufig untersuchten Prozesse bei der Beurteilung der Stärke von kausalen Zusammenhängen meist nur von einer kognitiven Ebene aus betrachtet wurden und dass es unklar ist, unter welchen situativen Umständen die Probanden regel-basiert vorgehen oder nicht. Wir haben dabei die Ansicht vertreten, dass die Prozesse bei der Beurteilung der Stärke kausaler Zusammenhänge besser als eine Interaktion von Individuum und Informationsumgebung (Clark, 1997; Hutchins, 1995a, b; Oestermeier & Hesse, 2000; Zhang, 1997, 2000; Zhang & Norman, 1994) aufgefasst werden können. Im Besonderen sind wir zu dem Schluss gekommen (s.a. das Schaubild in Abbildung 5), dass es neben personenbezogenen Faktoren entscheidend von aufgaben- und medienbezogenen Faktoren abhängt, ob und in welchem Umfang regel-basierte Strategien bei der Beurteilung der Stärke monokausaler Zusammenhänge zum Einsatz kommen. Der Schwerpunkt der vorliegenden Arbeit lag dabei auf der Untersuchung des Einflusses von medialen und personenbezogenen Faktoren auf den Einsatz komplexer regel-basierter Strategien bei der Beurteilung der Stärke monokausaler Zusammenhänge. Der Einfluss aufgabenbezogener Faktoren sollte dabei so weit wie möglich konstant gehalten werden. In drei Experimenten wurden die folgenden Fragestellungen untersucht:

1. Welchen Einfluss haben die Faktoren *Präsentationsform*, *Informationsart* und *Darbietungsart* auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen und auf die Strategiewahl der Probanden?

2. Unter welchen medialen Rahmenbedingungen gehen Probanden exakt regel-basiert vor?
3. Unterscheiden sich Individuen hinsichtlich Urteilsverhalten und Strategiewahl unter verschiedenen medialen Bedingungen voneinander?
4. Welchen Einfluss haben unterschiedliche individuelle Voraussetzungen wie z.B. die *numerische Verarbeitungskapazität* auf die Beurteilung der Stärke von Ursache-Wirkungs-Zusammenhängen?
5. In welchem Ausmaß lassen sich unterschiedliche Auswirkungen des Faktors *Darbietungsart* auf eine unterschiedliche Gedächtnisbeanspruchung zurückführen?

1. *zum Einfluss der drei Faktoren Präsentationsform, Informationsart und Darbietungsart*

Den deutlichsten Einfluss auf die Beurteilung der Stärke von monokausalen Ursache-Wirkungs-Zusammenhängen und die Strategiewahl hatte der Faktor *Darbietungsart*. Dies konnte in Experiment 2 nachgewiesen werden, in dem die Probanden bei summarischer Darbietungsart eine höhere mittlere Anzahl regel-konformer Lösungen als bei einzelfallweiser Darbietung produzierten und mehr Probanden regel-basierte Strategien einsetzten. Zudem erzielten die Probanden bei summarischer Darbietung schnellere Bearbeitungszeiten und eine höhere Verarbeitungseffizienz als bei einzelfallweiser Darbietung. Worauf der Einfluss des Faktors *Darbietungsart* möglicherweise zurückzuführen ist, soll bei der Beantwortung von Frage 5 diskutiert werden.

Der Einfluss des Faktors *Informationsart* wurde lediglich in Experiment 1 untersucht. Im Zusammenhang mit der Anzahl regel-konformer Lösungen ergaben sich ein trendnaher Einfluss, dass Probanden, die Wahrscheinlichkeiten im Vergleich zu Häufigkeiten zu beurteilen hatten, eine etwas höhere Anzahl regel-konformer Lösungen hervorbrachten und immerhin ein Trend, dass unter Wahrscheinlichkeitsbedingungen mehr Probanden regel-basierte Strategien einsetzten als unter Häufigkeitsbedingungen. Auf die Bearbeitungszeiten und die Verarbeitungseffizienz hatte der Faktor *Informationsart* hingegen keinen Einfluss. Der schwache Einfluss des Faktors *Informationsart* auf die Strategiewahl und die Anzahl regel-konformer Lösungen lässt sich vermutlich darauf zurückführen, dass die mentale Beanspruchung unter den beiden Wahrscheinlichkeitsbedingungen durch eine geringere Komplexität der zu bearbeitenden Aufgabe geringer ausfällt als unter den beiden Häufigkeitsbedingungen. Das Ausbleiben von Effekten des Faktors *Informationsart* auf die Bearbeitungszeit und die Verarbeitungseffizienz war von uns nicht vorhergesagt worden,



möglicherweise waren der Einfluss des Faktors *Präsentationsform* oder von individuellen Unterschieden im Zusammenhang mit der Schnelligkeit bei der Aufgabenbearbeitung stärker. Der Einfluss des Faktors *Präsentationsform* war in Experiment 1 zwar ebenfalls schwach ausgeprägt oder gar nicht vorhanden, aber anders geartet als der Einfluss des Faktors *Informationsart*. Auf die Anzahl regel-konformer Lösungen und die Strategiewahl hatte die *Präsentationsform* wider Erwarten und im Gegensatz zur *Informationsart* keinen Einfluss, was von uns post hoc auf das Auftreten von Deckeneffekten und/oder auf zu geringe Unterschiede zwischen den beiden in Experiment 1 eingesetzten Präsentationsformen Stapeldiagramm und tabellarische Vierfeldertafel zurück geführt wurde. Im Gegensatz zu einer Protokollliste sind die zu verarbeitenden, summarisch dargebotenen Informationen sowohl in einer tabellarischen Vierfeldertafel als auch in einem Stapeldiagramm relativ übersichtlich angeordnet. Beide Medien zeichnen sich durch eine hohe visuelle Strukturierung der darin angeordneten Informationen aus. Dadurch wird die mentale Beanspruchung für die Probanden reduziert und es werden ihnen mehr Strukturierungsangebote gemacht, wie sie die vier Zellinformationen zueinander in Beziehung setzen sollen. In Experiment 2 haben wir uns deshalb für die Frage interessiert, welchen Einfluss eine unterschiedliche *visuelle Strukturierung des darstellenden Mediums* oder, anders formuliert, eine unterschiedliche Übersichtlichkeit der Informationsanordnung auf die Strategiewahl und die Anzahl regel-konformer Lösungen haben. Dabei ergab sich ein Trend, dass visuell hoch strukturierte Stapeldiagramme, in denen die Informationen über Ursache-Wirkungs-Zusammenhänge relativ übersichtlich angeordnet sind, im Vergleich zu visuell gering strukturierten Fliesstexten, in denen die Kausalinformationen relativ unübersichtlich angeordnet sind, mehr Probanden regel-basierte Strategien einsetzten und die Probanden zur Produktion einer höheren Anzahl regel-konformer Lösungen anregten. Wie vorhergesagt, wirkte sich eine höhere *visuelle Strukturierung des darstellenden Mediums* allerdings nur dann auf die Strategiewahl und die Anzahl regel-konformer Lösungen aus, wenn die mentale Beanspruchung durch eine summarische Informationsdarbietung hinreichend reduziert war. Dagegen war bei einzelfallweiser Darbietung die mentale Beanspruchung für die Probanden aufgrund der im Vergleich zur summarischen Darbietung erhöhten Aufgabenkomplexität so hoch, dass die Probanden nicht von der höheren visuellen Strukturierung einer Vierfeldertafel mit Positionskreuz im Vergleich zu einem einfachen Fliesstext mit Aussagen zu Einzelereignissen profitieren konnten.

Auf die Bearbeitungszeit hatte der Faktor *Präsentationsform* in Experiment 1 einen trendnahen Einfluss, da Stapeldiagramme wie vorhergesagt etwas schneller von den

Probanden bearbeitet werden konnten als tabellarische Vierfeldertafeln. Wir haben diesen schwachen Einfluss des Faktors Präsentationsform auf die Bearbeitungszeiten so interpretiert, dass die Verarbeitung von Informationen in einem Stapeldiagramm durch die Möglichkeit der Verlagerung von komplexen Informationsverarbeitungsprozessen auf die Wahrnehmungsebene (Fiedler, 2000; Wickens & Carswell, 1995) oder durch ein Ermöglichen perzeptueller Inferenzen (Larkin & Simon, 1987) stärker beschleunigt werden konnte als in einer tabellarischen Vierfeldertafel. In Experiment 2 beschleunigte eine hohe visuelle Strukturierung des darstellenden Mediums die Informationsverarbeitung nur bei einzelfallweiser, aber nicht bei summarischer Informationsdarbietung. Diese Ergebnisse haben wir darauf zurückgeführt, dass der Faktor *visuelle Strukturierung* in Verbindung mit der Bearbeitungszeit vor allem Prozesse bei der Datenaggregation und -zuordnung, nicht aber bei der eigentlichen Urteilsbildung beeinflusst. Diese Prozesse nehmen bei einzelfallweiser Informationsdarbietung relativ viel Zeit in Anspruch und fallen deshalb stärker ins Gewicht als bei summarischer Informationsdarbietung, bei welcher den Probanden das zeitaufwändige Aggregieren und Zuordnen von Daten erspart bleibt. Bei summarischer Darbietung könnte ein stärkerer Einfluss des Faktors *visuelle Strukturierung* auf die Bearbeitungszeiten auch deshalb ausgeblieben sein, weil die Probanden bei hoher visueller Strukturierung im Stapeldiagramm nachgewiesenermaßen komplexere und vermutlich auch zeitaufwendigere Berechnungen durchgeführt haben als die Probanden, die die summarisch dargebotenen Informationen aus einem Fliesstext mit Häufigkeiten entnommen haben.

## 2. *Unter welchen medialen Rahmenbedingungen gehen Probanden exakt regel-basiert vor?*

Diese Frage lässt sich im Zusammenhang mit der Beurteilung der Stärke von monokausalen Zusammenhängen nicht allgemeingültig beantworten, da der Einfluss medialer Faktoren mit Ausnahme der *Darbietungsart* eher schwach zu sein scheint und das Urteilsverhalten der Probanden zudem von aufgaben- und personenbezogenen Faktoren mit beeinflusst wird.

In Verbindung mit der Anzahl regel-konformer Lösungen und der Strategiewahl haben unsere Untersuchungen das Bild ergeben, dass die Probanden vor allem dann regel-basiert vorgehen, wenn die mentale Beanspruchung für die Probanden möglichst gering gehalten wird. Die mentale Beanspruchung kann vor allem durch eine summarische Darbietung der Informationen über Ursache-Wirkungs-Zusammenhänge reduziert werden.

In subtilerer Weise machen medienbezogene Faktoren den Einsatz regel-basierter Strategien aber nicht nur dadurch wahrscheinlicher, dass sie, wie von einigen Forschergruppen (wie z.B. Arkes & Harkness, 1983; Baker et al., 1996; Kao & Wasserman, 1993; Shaklee & Mims, 1982) angenommen wurde, die Gedächtnisbeanspruchung reduzieren. Vor allem der Einfluss

des Faktors *visuelle Strukturierung* in Experiment 2 hat gezeigt, dass es nicht ausreicht, die Gedächtnisbeanspruchung allein zu reduzieren, um möglichst viele Probanden dazu anzuregen, von regel-basierten Strategien Gebrauch zu machen und regel-konforme Lösungen zu produzieren. Unter der Bedingung *VS hoch* haben in Experiment 2 bei summarischer Darbietung mehr Probanden regel-basierte Strategien eingesetzt als unter der Bedingung *VS gering*, zudem erzielten die Probanden unter der Bedingung *VS hoch* eine höhere mittlere Anzahl regel-konformer Lösungen. In Experiment 1 haben die Probanden unter der Kontrollbedingung *Protokollliste*, unter der die Gedächtnisbeanspruchung auch reduziert war, weniger regel-konforme Lösungen gebildet und seltener regel-basierte Strategien eingesetzt als in den vier Experimentalbedingungen. Die höhere Anzahl regel-basiert vorgehender Probanden in den vier Experimentalbedingungen aus Experiment 1 und in der Bedingung *SD sum* aus Experiment 2 führen wir darauf zurück, dass den Probanden unter diesen Bedingungen mehr Strukturierungsangebote (Larkin & Simon, 1987; Stenning & Oberlander, 1995; Zhang, 2000, 1997; Zhang & Norman, 1994) dazu gemacht wurden, wie sie die vier Zellinformationen weiter verarbeiten und regel-konform zueinander in Beziehung setzen sollen. Zudem ermöglichen eine Protokollliste und ein Fliesstext mit summarischer Darbietung keine perzeptuellen Inferenzen im Sinne von Larkin & Simon (1987) und unterstützen die menschliche Informationsverarbeitung nicht dabei, komplexe mentale Operationen wie z.B. Vergleiche oder Schlussfolgerungen auf die Wahrnehmungsebene zu verlagern (Wickens & Carswell, 1995). Aus einer motivationalen Perspektive heraus antizipieren die Probanden bei summarischer Informationsdarbietung in einem visuell hoch strukturierten Stapeldiagramm einen geringeren Aufwand im Zusammenhang mit dem Einsatz komplexer regel-basierter Strategien und setzen diese deshalb vermehrt ein (Kleinmuntz & Schkade, 1993), während sie bei einzelfallweiser Darbietung oder bei einer Präsentation von visuell gering strukturierten Informationen vermehrt auf einfachere und weniger aufwendige Strategien und Heuristiken zurück greifen.

Der Faktor *Informationsart* hatte in Experiment 1 einen schwachen Einfluss auf die Strategiewahl der Probanden und die Anzahl regel-konformer Lösungen. Das Ergebnis, dass die Probanden bei einer ausschließlich summarischen Darbietung und unter Wahrscheinlichkeitsbedingungen häufiger regel-konforme Lösungen produzieren und vermehrt regel-konforme Lösungen generieren, lässt sich zum einen auf die geringere Gedächtnisbeanspruchung unter Wahrscheinlichkeitsbedingungen zurückführen, zum anderen spielt auch die Wechselwirkung zwischen *Informationsart* und der den Probanden zur Verfügung gestellten Antwortskala eine Rolle. Anhand der vorgegebenen und von 0 bis 100

reichenden Wirksamkeitsskala mussten die Probanden unter Wahrscheinlichkeitsbedingungen weniger Arbeitsschritte ausführen als unter Häufigkeitsbedingungen.

Die durchgeführten Häufigkeitsanalysen in den Experimenten 1 bis 3 haben ergeben, dass die Probanden nicht einheitlich vorgehen und selbst unter gleichen medialen Bedingungen von unterschiedlichen Strategien Gebrauch machen. Dennoch haben die Häufigkeitsanalysen auch nahe gelegt, dass mediale Faktoren nicht nur einen Einfluss darauf haben, in welchem Umfang die Probanden von regel-basierten Strategien Gebrauch machen, sondern auch, von welcher regel-basierten Strategie die Probanden Gebrauch machen. Obwohl in allen drei Experimenten insgesamt mehr Probanden die  $\Delta P$ -Regel eingesetzt zu haben schienen, ergab sich in Experiment 1 ein Trend, dass unter Wahrscheinlichkeitsbedingungen anteilmäßig mehr Probanden die Stärke des Kausalzusammenhangs entsprechend der  $\Delta P$ -Regel ermittelten, während unter Häufigkeitsbedingungen anteilmäßig mehr Probanden von der Power PC-Theorie Gebrauch machten.

Aus unserer Sicht veranschaulichen die Ergebnisse aus den Untersuchungen von Buehner, Cheng & Clifford (2003) besonders gut, unter welchen medialen Einflussbedingungen die Probanden entsprechend den Vorhersagen der PPC-Theorie (Cheng, 1997) vorgehen: Wie in unseren Untersuchungen sollte die Aufgabe der Probanden darin bestehen, die Stärke monokausaler Ursache-Wirkungs-Zusammenhänge zu beurteilen. Zudem sollten die Daten als Häufigkeiten vorliegen. Hilfreich ist es vermutlich auch, wenn die Anzahl beobachteter Einzelfälle wie in unseren Untersuchungen und in den Untersuchungen von Buehner, Cheng & Clifford (2003) bei An- und Abwesenheit der Ursache gleich groß ist. Der Einsatz einer Vorgehensweise entsprechend den Vorhersagen der PPC-Theorie wird auch dann wahrscheinlicher, wenn die zu beurteilenden Einzelinformationen in einer übersichtlich gestalteten Listenform (gruppierte Anhäufungen von Bildertafeln wie in Experiment 2 von Buehner, Cheng & Clifford) oder in einer übersichtlich gestalteten summarischen Darbietungsform (z.B. Stapeldiagramm) angeordnet sind. Gemeinsam ist Experiment 2 von Buehner, Cheng & Clifford (2003) und den Bedingungen mit summarischer Darbietungsart in einem Stapeldiagramm auch, dass die Probanden durch visuelle Hinweisreize, Gruppierungen und die grafische Anordnung der zu beurteilenden Informationen dazu eingeladen werden, Erfolgs- und Basisraten zu ermitteln, miteinander zu vergleichen und zueinander ins Verhältnis zu setzen. Von entscheidender Bedeutung scheint aber die an die Probanden gerichtete Fragestellung zu sein: In unseren Untersuchungen wurden die Probanden danach gefragt, die Wirksamkeit einer Substanz zu beurteilen, in Experiment 1 von Buehner, Cheng & Clifford (2003) danach, wie stark die fragliche Ursache die Wirkung verursacht. Nach

Buehner, Cheng & Clifford (2003) kann eine solche Fragestellung die Probanden sowohl zu einer Vorgehensweise entsprechend den Vorhersagen der Power PC-Theorie anregen, als auch zu einer Vorgehensweise entsprechend den Vorhersagen der  $\Delta P$ -Regel. Indem eine derart ambivalente Fragestellung durch eine eindeutige kontrafaktische Fragestellung ersetzt wird ("in wie vielen von 100 Fällen, in denen eine Wirkung ohne Ursache ausbleibt, würde die Wirkung eintreten, wenn die fragliche Ursache eingeführt werden würde?"), können den Probanden, wie die Untersuchungen von Buehner, Cheng & Clifford (2003, s.a. Collins & Shanks, 2006) zeigen, Berechnungen oder Ermittlungen der kausalen Stärke nach der Power PC-Theorie nahe gelegt werden. Da es sich bei der Fragestellung um einen aufgabenbezogenen Faktor handelt, wurden von uns keine Variationen der Fragestellung untersucht.

### *3. Unterscheiden sich Individuen hinsichtlich Urteilsverhalten und Strategiewahl unter verschiedenen medialen Bedingungen voneinander?*

Die Antwort auf diese Frage lautet ganz klar ja. Wir haben im Theorieteil dieser Arbeit unter anderem herauszuarbeiten versucht, dass unterschiedliche mediale Bedingungen unterschiedliche Anforderungen an die menschliche Informationsverarbeitung stellen. In Abhängigkeit von unterschiedlichen medialen Bedingungen fällt also die mentale Beanspruchung der Probanden unterschiedlich aus. Die Auffassung, den Prozess der kausalen Urteilsbildung als eine Interaktion von Individuum und Informationsumgebung zu begreifen (Clark, 1997; Hutchins, 1995a, b; Oestermeier & Hesse, 2000; Zhang, 1997, 2000; Zhang & Norman, 1994) ist für uns zentral und wurde nach unserem Verständnis innerhalb des Forschungsbereichs zur Kausalkognition an vielen Stellen vernachlässigt. Es genügt nicht, die Prozesse bei der Beurteilung der Stärke kausaler Zusammenhänge nur auf der Ebene der Kognition zu betrachten. Vieles spricht dafür, dass die Prozesse bei der Beurteilung der Stärke von kausalen Zusammenhängen als ein Zusammenspiel von Kognition, sowie medien-, aufgaben- und personenbezogenen Faktoren konzeptualisiert werden können (s. das Schaubild in Abbildung 5 auf S. 132).

Die in den Experimenten 1 bis 3 durchgeführten Häufigkeitsanalysen haben in unseren Augen gezeigt, dass mediale Rahmenbedingungen die Strategiewahl einzelner Individuen zwar nicht determinieren, aber über verschiedene Individuen hinweg den Gebrauch unterschiedlicher Strategien unterschiedlich wahrscheinlich machen. An dieser Stelle soll auch dem Missverständnis vorgebeugt werden, dass wir nicht so weit gehen zu behaupten, dass medienbezogene Faktoren festlegen, welcher Informationsverarbeitungsmechanismus in Gang gesetzt wird oder nicht, obwohl das im Einzelfall zutreffend sein kann (Beispiel: Bei

einer rein summarischen Darbietung sind assoziative Lernmechanismen ausgeschlossen). Im Zusammenspiel mit aufgaben- und personenbezogenen Faktoren legen mediale Faktoren aber die mentale Beanspruchung fest und machen so den Einsatz unterschiedlicher Strategien und Heuristiken unterschiedlich wahrscheinlich. Auch bei der Beurteilung der Stärke von monokausalen Zusammenhängen haben medienbezogene Faktoren vielfältige Auswirkungen, wie sie von Zhang und anderen (Kleinmuntz & Schkade, 1993; Larkin & Simon, 1987; Stenning & Oberlander, 1995; Zhang, 2000, 1997; Zhang, Johnson, Malin & Smith, 2002; Zhang & Norman, 1994) beschrieben worden sind (s.a. Kapitel 6.2).

Traditionelle Fragestellungen der Kausalkognition, wie z.B. die Frage, ob der Erwerb von Kausalwissen regel-basiert abläuft oder auf assoziativen Lernmechanismen beruht, führen nach unserer Auffassung in eine falsche Richtung. Im Hinblick auf die Beurteilung der Stärke von monokausalen Zusammenhängen halten wir es für sinnvoller, von der Annahme vielfältiger Informationsverarbeitungsmöglichkeiten auszugehen und eher danach zu fragen, unter welchen medien- und aufgabenbezogenen Rahmenbedingungen werden in Abhängigkeit von welchen personenbezogenen Faktoren welche Informationsverarbeitungsprozesse in Gang gesetzt oder ausgelöst.

#### *4. Zum Einfluss von unterschiedlichen individuellen Voraussetzungen und der numerischen Verarbeitungskapazität im Besonderen*

Die Experimente 2 und 3 belegen klar, dass Probanden mit hohen Werten im numerischen Denken eine höhere Anzahl regel-konformer Lösungen produzieren und effizienter arbeiten als Probanden mit geringen Werten im numerischen Denken. Zudem konnte herausgefunden werden, dass der Einfluss medienbezogener Faktoren vor allem bei Personen mit hohen Werten im numerischen Denken zu Tage zu treten scheint. In Experiment 1 wurde der Faktor *Informationsart* dann signifikant, wenn nur die Gruppe der Leibniz-Kollegiaten untersucht wurde, von der wir annahmen, dass sie hinsichtlich des numerischen Denkens höhere Werte aufweist als die ebenfalls separat untersuchte studentische Teilpopulation. In Experiment 2 hatten in der Teilstichprobe der Probanden mit hohen Werten im numerischen Denken sowohl der Faktor *Darbietungsart* als auch der Faktor *visuelle Strukturierung* einen signifikanten Einfluss auf die Anzahl regel-konformer Lösungen, während in der Teilstichprobe der Probanden mit geringen Werten im numerischen Denken für den Einfluss des Faktors *Darbietungsart* nur ein Trend erzielt wurde und der Faktor *visuelle Strukturierung* fast gänzlich ohne Einfluss auf die Anzahl regel-konformer Lösungen war. Probanden mit hohen Werten im numerischen Denken scheinen in unseren Augen offensichtlich besser von den Strukturierungsangeboten oder der repräsentationalen Führung (Suthers, 2001) der in unseren

Experimenten eingesetzten visuell hoch strukturierten Stapeldiagramme mit summarischer Informationsdarbietung zu profitieren. Experiment 3 brachte die zusätzliche Erkenntnis, dass der Einfluss der Kovariate *numerisches Denken* auf die Anzahl regel-konformer Lösungen und den Einsatz regel-basierter Strategien bei Probanden, die die Aufgabe hatten, Notizen anzufertigen, zurückgeht. Von größerer Bedeutung war unter diesen Bedingungen die Zeit, welche die Probanden für die eigentliche Urteilsbildung benötigten (*rating time*). Dieses von uns nicht vorhergesehene Ergebnis haben wir dahin gehend interpretiert, dass unter den beiden Bedingungen mit Notizen die Motivation der Probanden, welche sich in der Dauer der eigentlichen Urteilsbildung manifestiert, von größerer Bedeutung ist als ihre numerische Verarbeitungskapazität. Die unterschiedlichen Einflüsse von *rating time* und *numerisches Denken* in Abhängigkeit von der Aufgabenstellung, Notizen anzufertigen oder nicht, zeigt, wie personen- und aufgabenbezogene Faktoren miteinander interagieren können.

Die in allen drei Experimenten durchgeführten Häufigkeitsanalysen offenbarten, dass die Mehrheit der Probanden zwischen verschiedenen Urteilsbildungsprozessen ihre Strategie mindestens einmal wechselt. Dieses Ergebnis, dass eine Mehrheit der Probanden selbst unter den gleichen aufgaben- und medienbezogenen Rahmenbedingungen ihre Vorgehensweise zwischen verschiedenen Urteilsprozessen ändert, lässt es in unseren Augen fragwürdig erscheinen, einen einheitlichen Verarbeitungsmechanismus bei der Beurteilung der Stärke von kausalen Zusammenhängen zu unterstellen. Den hohen Wechsler-Anteil haben wir als Hinweis darauf verstanden, dass eine Mehrheit der Probanden versucht, ihre Strategien von Urteil zu Urteil zu optimieren oder an die subjektiv wahrgenommene Aufgabenkomplexität anzupassen. Ein weiteres, im Zusammenhang mit dem Einfluss personenbezogener Faktoren auf die Strategiewahl stehendes Ergebnis aus Experiment 2 war, dass Probanden mit einer hohen numerischen Verarbeitungskapazität eher dazu neigen, die Stärke eines Kausalzusammenhangs entsprechend den Vorhersagen der PPC-Theorie zu ermitteln, während Probanden mit geringen Werten im numerischen Denken eher entsprechend den Vorhersagen der weniger komplexen  $\Delta P$ -Regel vorgegangen sind. Die durchgeführten Häufigkeitsanalysen in Experiment 2 lassen in gewisser Weise sogar den Schluss zu, dass eine Ermittlung der Stärke des kausalen Zusammenhangs entsprechend der PPC-Theorie bei Probanden mit geringeren Werten im numerischen Denken fast nicht vorkam (mittlere Anzahl PPC-theoriekonformer Lösungen = .10 bei Probanden mit geringen Intelligenzwerten im numerischen Denken).

Die vor allem in Experiment 2 festgestellte Abhängigkeit der Anzahl regel-konformer Lösungen und der Häufigkeit des Einsatzes regel-basierter Strategien von der numerischen

Verarbeitungskapazität der Probanden lässt die bei der Erforschung der Informationsverarbeitungsprozesse bei der Beurteilung der Stärke von Kausalzusammenhängen gängige und auch von uns vollzogene Praxis, ausschließlich studentische Stichproben heranzuziehen, bedenkenswert erscheinen, weil die gefundenen Ergebnisse sich nicht einfach auf andere Populationsgruppierungen übertragen lassen. Personenbezogene Faktoren wie z.B. die numerische Verarbeitungskapazität können bei der Beurteilung der Stärke von Kausalzusammenhängen hoch bedeutsam sein und sollten in zukünftigen Untersuchungen mit erfasst werden.

*5. In wie weit lassen sich unterschiedliche Auswirkungen des Faktors Darbietungsart auf eine unterschiedliche Gedächtnisbeanspruchung zurückführen?*

Im Hinblick auf die Anzahl regel-konformer Lösungen sind wir in Experiment 3 zu dem Ergebnis gelangt, dass sich die höhere Anzahl regel-konformer Lösungen bei einer summarischen im Vergleich zu einer einzelfallweisen Darbietung nicht allein auf eine reduzierte Gedächtnisbeanspruchung zurückführen lässt. In Experiment 3 wurde die Gedächtnisbeanspruchung dadurch reduziert, dass die Probanden bei einzelfallweiser Informationsdarbietung die Aufgabe hatten, Notizen anzufertigen. Diese Gedächtnisentlastung führte im Vergleich zur Bedingung einzelfallweise Darbietung ohne Notizen zu einem nicht signifikanten Anstieg der Anzahl regel-konformer Lösungen unter den beiden Einzelfallbedingungen mit Notizen. Der zu beobachtende Anstieg der Anzahl regel-konformer Lösungen von der Einzelfalldarbietung ohne Notizen zu den beiden Bedingungen mit einzelfallweiser Darbietung und Notizen war allerdings nicht so stark wie der signifikante Anstieg zwischen der Einzelfallbedingung ohne Notizen und der summarischen Bedingung ohne Notizen. Dieses schon aus den Untersuchungen von Ward & Jenkins (1965) bekannte Ergebnis wurde von uns so interpretiert, dass der Anstieg der Anzahl regel-konformer Lösungen zwischen den Bedingungen einzelfallweiser und summarischer Informationsdarbietung ohne Notizen nur teilweise, aber nicht vollständig durch eine reduzierte Gedächtnisbeanspruchung zu erklären ist. Offensichtlich beinhaltet die Bedingung mit rein summarischer Informationsdarbietung Qualitäten, die über die Qualitäten einer reinen Gedächtnisentlastung und auch einer strukturierten Datenaufbereitung hinausgehen.

Wir haben uns deshalb gefragt, wodurch sich der im Zusammenhang mit der Anzahl regel-konformer Lösungen festgestellte Unterschied zwischen den beiden Einzelfallbedingungen mit Notizen und der summarischen Bedingung ohne Notizen erklären lässt. Die Annahme von bedeutsamen Unterschieden hinsichtlich der Informationsanordnung zwischen den beiden Einzelfallbedingungen mit Notizen und der summarischen Bedingung ohne Notizen wurde



von uns ebenso verworfen, wie die Annahme, dass die Notizen der Probanden in den beiden Bedingungen mit Notizen fehlerhaft erstellt worden sein könnten. Schon in Experiment 2 haben wir Hinweise dafür gefunden, dass die Qualität der Informationsanordnung oder das Ausmaß an visueller Strukturierung des darstellenden Mediums wegen der hohen mentalen Beanspruchung bei einzelfallweiser Informationsdarbietung keinen bedeutenden Einfluss auf die Anzahl regel-konformer Lösungen hat. Diese Erklärung lässt aber die Frage offen, warum die mentale Beanspruchung in den beiden Bedingungen mit Notizen trotz reduzierter Anforderungen an das Gedächtnis immer noch höher ist als bei summarischer Darbietung ohne Notizen.

Wir haben vermutet, dass die mentale Beanspruchung in den beiden Bedingungen mit Notizen höher ist als unter der Bedingung mit summarischer Darbietung ohne Notizen, weil die Aufgabenbeschaffenheit in den beiden Bedingungen mit Notizen durch die mit einer einzelfallweisen Informationsdarbietung einhergehenden Fokussierung auf einfache, aber zeitaufwendige Handlungsroutinen und den fehlenden Gesamtüberblick während der eigentlichen Aufgabenbearbeitung (*time on task*) eine andere war als in der Bedingung mit summarischer Darbietung ohne Notizen, in der die Probanden schon während der *time on task* an Hand der vorliegenden Zusammenfassung aller für den Urteilsprozess notwendigen Informationen in die Lage versetzt wurden, Vergleiche vorzunehmen, Berechnungen anzustellen oder Schlussfolgerungen zu ziehen. In den Einzelfallbedingungen mit oder ohne Notizen hat sich die Aufgabenkomplexität für die Probanden erhöht, weil die Zusammenhänge zwischen den verschiedenen Zellhäufigkeiten durch die Fokussierung auf die Einzelfälle und den fehlenden Überblick nicht salient gemacht werden konnten. Obwohl das Gedächtnis durch das Anfertigen von Notizen entlastet wird, erhalten die Probanden während der *time on task* durch den fehlenden Überblick keine Strukturierungsangebote (Larkin & Simon, 1987; Stenning & Oberlander, 1995; Zhang, 2000, 1997; Zhang & Norman, 1994) und können weder perzeptuelle Inferenzen (Larkin & Simon, 1987) vornehmen noch kognitive Operationen auf die Wahrnehmungsebene verlagern (Wickens & Carswell, 1995). Erneut kann die Kosten-Nutzen Theorie von Kleinmuntz & Schkade (1993) als weitere Interpretationsmöglichkeit heran gezogen werden. So war der von den Probanden betriebene Aufwand bei einzelfallweiser Darbietung mit oder ohne Notizen deutlich größer als bei einer summarischen Darbietung ohne Notizen. Dies zeigte sich an den mittleren Bearbeitungszeiten in den drei Einzelfallbedingungen mit oder ohne Notizen, die in Experiment 3 signifikant länger waren als in der Bedingung mit summarischer Informationsanordnung ohne Notizen, in welcher der Aufwand für die Aufgabenbearbeitung im Vergleich zu den drei

Einzelfallbedingungen wegen der geringeren Aufgabenkomplexität deutlich reduziert war. Der Blick auf die Verarbeitungseffizienzen zeigte, dass die mentale Beanspruchung der Probanden in den drei Einzelfallbedingungen wegen der längeren Bearbeitungszeiten (höherer Aufwand) und der geringeren Anzahl regel-konformer Lösungen (geringerer Nutzen) höher ausfiel als in der summarischen Bedingung ohne Notizen. Demgegenüber unterschieden sich die beiden Einzelfallbedingungen mit Notizen und die Einzelfallbedingung ohne Notizen nicht hinsichtlich der Verarbeitungseffizienz, was von uns als Hinweis darauf verstanden wurde, dass die mentale Beanspruchung in den drei Einzelfallbedingungen in etwa gleich hoch war, obwohl die Probanden in den beiden Einzelfallbedingungen mit Notizen einen höheren Aufwand (längere Bearbeitungszeiten) als in der Einzelfallbedingung ohne Notizen betrieben haben. Dieser höhere Aufwand in den beiden Einzelfallbedingungen mit Notizen ergab sich zum einen durch das zusätzliche Anfertigen von Notizen (längere *time on task*), zum anderen durch einen höheren zeitlichen Aufwand bei der Urteilsbildung (längere *rating time*). Dieser höhere Aufwand, der von den Probanden in den beiden Einzelfallbedingungen mit Notizen betrieben wurde, konnte aber durch einen höheren Nutzen (höhere Anzahl regel-konformer Lösungen) kompensiert werden und führte letztlich dazu, dass die Verarbeitungseffizienz in den drei Bedingungen mit einzelfallweiser Darbietung in etwa gleich groß war. Im Sinne der Kosten-Nutzen-Theorie (Kleinmuntz & Schkade, 1993) wäre es vorstellbar, dass die Probanden in den beiden Bedingungen mit Notizen und in Verbindung mit der damit einher gehenden Gedächtnisentlastung die Möglichkeit eines Einsatzes von komplexeren Strategien antizipiert haben und deshalb dazu bereit waren, insbesondere während der *rating time* mehr Aufwand für die Urteilsbildung im Vergleich zur Einzelfallbedingung ohne Notizen zu betreiben.

## 12 Ausblick

Ausgehend von dem Schaubild in Abbildung 5 (S. 132) erscheint es lohnenswert, den Einfluss von weiteren aufgaben-, medien- und personenbezogenen Faktoren auf die kognitiven Prozesse bei der Beurteilung der Stärke von monokausalen Zusammenhängen und die Strategiewahl der Probanden zu untersuchen. Speziell zum Einfluss medialer Faktoren wurde im Rahmen dieser Arbeit ein theoretisches Fundament gelegt, welches für Folgeuntersuchungen nutzbar gemacht werden könnte und steter Erweiterung, Modifizierung und Zuspitzung bedarf. In Verbindung mit den experimentellen Ergebnissen, über die im Rahmen dieser Arbeit berichtet wurde, erschienen uns folgende Ideen verfolgenswert:

*Aufgabenebene:* Abgesehen von dem bereits untersuchten Faktor *Notizen anfertigen* halten wir auf der Aufgabenebene insbesondere Variationen der *Fragestellung* (Instruktion, Wortwahl) und der eingesetzten *Antwortskalen* und *-formate* für interessant. Neuere Untersuchungen zum Einfluss dieser aufgabenbezogenen Faktoren legen deren Bedeutung nahe (Buehner, Cheng & Clifford, 2003; Collins & Shanks, 2006; Perales & Shanks, in press; Shanks, 2004; White, 2003b). Zu vermuten ist, dass unterschiedliche Fragestellungen, eine unterschiedliche Wortwahl oder unterschiedliche Antwortskalen die Aufgabe in einem unterschiedlichen Ausmaß festlegen, strukturieren, eingrenzen oder eindeutig machen. Dadurch könnten unterschiedliche regel-basierte Strategien in einem unterschiedlichen Ausmaß nahe gelegt oder salient gemacht werden. Denkbar wären auch Untersuchungen, in denen mögliche Wechselwirkungen zwischen den bereits untersuchten medien- und personenbezogenen Faktoren und den genannten aufgabenbezogenen Faktoren analysiert werden. Darüber hinaus würden künftige Untersuchungen zum Einfluss bestimmter aufgabenbezogener Faktoren begriffliche Definitionen und eine theoretische Fundierung erforderlich machen.

*Medienebene:* In Experiment 2 haben wir den Einfluss der Faktoren *Darbietungsart* und *visuelle Strukturierung* untersucht. Dabei war der Einfluss des Faktors *visuelle Strukturierung* schwächer als derjenige des Faktors *Darbietungsart*, immerhin konnte aber ein Trend für den Einfluss der *visuellen Strukturierung* auf die *Anzahl regel-konformer Lösungen* erzielt werden. Unklar blieb dabei, worin eigentlich die größere Übersichtlichkeit der Informationsanordnung eines Stapeldiagramms mit Häufigkeiten und summarischer Informationsdarbietung gegenüber einem Fliesstext mit Häufigkeiten und summarischer Informationsdarbietung besteht. Welche Rolle spielen Faktoren wie die räumliche Anordnung

der Informationen oder die Farbgebung dabei? In Experiment 2 wurde versucht, die Informationen über Kausalzusammenhänge so anzuordnen, dass dadurch ein regel-konformes Zueinander-in-Beziehung-setzen der Informationen möglichst nahe gelegt und erleichtert wird. Was geschieht aber zum Beispiel, wenn die Informationen in einem Stapeldiagramm so angeordnet werden, dass sie regel-konforme Vergleiche und Schlussfolgerungen eher erschweren und den Probanden eigene Transformationen der präsentierten Informationen abverlangen, um zu regel-konformen Lösungen zu gelangen? Produzieren sie dann weniger regel-konforme Lösungen und setzen weniger regel-basierte Strategien ein als in einem Stapeldiagramm mit einer erleichternden Informationsanordnung oder gar einem Fliesstext ohne übersichtliche Informationsanordnung? Arbeiten sie dabei langsamer und weniger effizient?

Experiment 3 zielte darauf ab, zu klären, worauf der signifikante Einfluss des Faktors *Darbietungsart* beruht. Dabei wurde die Notwendigkeit von Folgeuntersuchungen hervorgehoben, in denen die vermuteten Einflüsse einer Gesamtüberblicks und des Zeitpunktes der Präsentation einer Gesamtschau vor oder nach einer einzelfallweisen Darbietung aller relevanten Informationen genauer analysiert und mit dem Einfluss einer rein summarischen Darbietung verglichen werden. Offen ist auch die Frage, ob es einen Unterschied macht, wenn die Probanden wie in Experiment 3 eigenständig Notizen anfertigen oder wenn diese Notizen computergestützt, ohne eigenes Zutun der Probanden, aber für sie einsehbar erstellt werden.

Weiter würden wir es als interessant erachten, ikonische Präsentationsformen in summarischer und einzelfallweiser Darbietung, sowie im Listenformat zu entwickeln und diese mit den bereits eingesetzten, eher abstrakten Präsentationsformen im Hinblick auf ihren Einfluss auf die *Anzahl regel-konformer Lösungen*, die *Bearbeitungszeit* oder die *Strategienwahl* zu untersuchen. Ikonische Präsentationsformen wurden schon in mehreren Untersuchungen (z.B. Anderson & Sheu, 1995; Buehner, Cheng & Clifford, 2003; Melchers, Üngör & Lachnit, 2005) eingesetzt, ohne jedoch medienpsychologisch reflektiert zu werden.

Unabhängig von den gemachten Vorschlägen für mögliche Folgeuntersuchungen ist festzuhalten, dass die in unseren Untersuchungen festgestellten medialen Effekte teilweise schwach ausgefallen sind. Das macht weitere Untersuchungen notwendig, in denen die gefundenen schwachen Effekte repliziert oder widerlegt werden können. Daran anschließend könnten in der Hoffnung, die gefundenen Effekte zu akzentuieren oder deutlicher werden zu lassen, Verfeinerungen am eingesetzten Paradigma vorgenommen werden.

*Personenebene:* Nach unserer Auffassung wurde in den Experimenten 2 und 3 mit der *numerischen Verarbeitungskapazität* ein wichtiger personenbezogener Faktor identifiziert, der die Prozesse bei der Beurteilung der Stärke monokausaler Zusammenhänge beeinflusst. Um den gefundenen Einfluss der Kovariate *numerisches Denken* auf nicht studentische Populationen übertragen zu können, wäre es sinnvoll, auch mit nicht studentischen Stichproben ähnliche Untersuchungen durchzuführen wie in den Experimenten 2 und 3. Dabei wäre allerdings zu vermuten, dass die Unterschiede in Abhängigkeit von den Faktoren *Darbietungsart* und *visuelle Strukturierung* geringer ausfallen als in Experiment 2, da Probanden aus nicht studentischen Populationen wahrscheinlich geringere Werte im *numerischen Denken* erzielen als die Probanden aus unseren studentischen Stichproben. Fraglich wäre dabei auch, ob die in den Experimenten 2 und 3 eingesetzten Items aus dem BIS-Test auch für nicht studentische Stichproben geeignet wären, da der BIS-Test ursprünglich für Jugendliche und junge Erwachsene mit Ober- und Mittelschulbildung entwickelt wurde und insofern Bodeneffekte bei Probanden mit niedrigerer Bildung zu befürchten wären.

Abgesehen von der Wahl der Stichprobe und des eingesetzten Messinstruments sollte auch die *numerische Verarbeitungskapazität* in möglichen Folgeuntersuchungen mit erhoben werden, um mögliche Wechselwirkungen zwischen dem *numerischen Denken* und den darin untersuchten aufgaben- und/oder medienbezogenen Faktoren erfassen zu können.

Interessant wäre es auch, den Einfluss *motivationaler Faktoren* mit zu erheben (s.a. De Houwer & Beckers, 2002) und die Fragestellung zu untersuchen, welchen Einfluss *motivationale Faktoren* auf die *Anzahl regel-konformer Lösungen* und die *Strategienwahl* in Untersuchungen zum Einfluss von unterschiedlichen aufgaben- und/oder medienbezogenen Faktoren haben. Zu diesem Zweck müsste erst einmal ein geeignetes Instrument zur Messung der Motivation gefunden oder entwickelt werden. Im Zusammenhang mit den Ergebnissen aus Experiment 3 haben wir die Vermutung geäußert, dass durch die Messung der *rating time* die Motivation eines Probanden mit erfasst wird. Diese vage Hypothese müsste erst einmal durch entsprechende Studien validiert werden. Zu bedenken ist dabei sicher, dass die Dauer der *rating time* nicht nur von motivationalen Faktoren, sondern vor allem von der auszuführenden Aufgabe abhängt. Insofern würden Validierungsstudien zur *rating time* nur unter *ceteris paribus*-Bedingungen in Verbindung mit der Aufgabenstellung Sinn machen.

*Ebene der Kognition:* Die durchgeführten Experimente und Ergebnisse aus anderen Studien (Allan, 1993; Buehner, Cheng & Clifford, 2003; Collins & Shanks, 2006; Perales & Shanks, in press; White 2003b) weisen klar darauf hin, dass aufgaben- und medienbezogene Faktoren

die kognitiven Prozesse beim Erwerb von Wissen über kausale Zusammenhänge und bei der kausalen Urteilsbildung beeinflussen. Solche Einflüsse werden aber weder von den verschiedenen regel-basierten Ansätzen (kontingenztheoretischen Ansätzen, Probabilistisches Kontrastmodell, Power PC Theorie, Theorie der empirischen Evidenzevaluation), noch von assoziationalistischen Ansätzen vorhergesagt. Wir vermuten, dass die bestehenden Ansätze die Auswirkungen medialer und aufgabenbezogener Einflüsse besser vorhersagen könnten, wenn zusätzlich das Ausmaß der Gedächtnisbeanspruchung erfasst und als gewichteter Faktor mit berücksichtigt werden könnte. Bislang existiert im Bereich der Kausalkognition noch kein Maß für die Gedächtnisbeanspruchung. Unsere Studien weisen in Verbindung mit der Beurteilung der Stärke von monokausalen Zusammenhängen darauf hin, dass vor allem die Darbietungsart und die numerische Verarbeitungskapazität die Gedächtnisbeanspruchung beeinflussen. Während nach unserem Kenntnisstand noch kein Modell Unterschiede in Abhängigkeit von der numerischen Verarbeitungskapazität vorhersagen kann, wurde kurz vor Fertigstellung dieser Arbeit ein Modell entwickelt, welches für sich in Anspruch nimmt, die Einflüsse unterschiedlicher Darbietungsarten (einzelfallweise, summarisch und listenweise) vorherzusagen. Es handelt sich dabei um das schon in Kapitel 3.1.3 (*Empirische Befunde und Kritik zur Theorie der kausalen Power*, S. 45) vorgestellte *Causal Support Modell* von Griffiths & Tenenbaum (2005), einer Variante eines Bayesianischen Netzwerkmodells (Gopnik, Glymour, Sobel, Schulz, Kushnir & Danks, 2004; Gopnik, Sobel, Schulz & Glymour, 2001). Griffiths & Tenenbaum (2005) haben unter anderem retrospektiv die Daten einer Reihe von prominenten Studien (Anderson & Sheu, 1995; Buehner & Cheng, 1997; Lober & Shanks, 2000; White, 1998, 2002c, 2003c) analysiert und kamen dabei zu dem Ergebnis, dass das *Causal Support Modell* die Daten im Vergleich zur  $\Delta P$ -Regel und zur Power PC Theorie entweder am besten oder nahezu am besten vorhersagt.

Im Rahmen dieser Arbeit haben wir uns nur mit der Beurteilung monokausaler probabilistischer Zusammenhänge beschäftigt. Wir erheben nicht den Anspruch, dass die durchgeführten Experimente die Fülle der kognitiven Prozesse widerspiegeln, die ablaufen, wenn Menschen Wissen über kausale Zusammenhänge in unserer Alltagswelt erwerben. Beispielsweise beinhaltet der Erwerb von Wissen über kausale Zusammenhänge häufig die das gleichzeitige Berücksichtigen des Zusammenwirkens von mehreren, teilweise verborgenen Ursachen. Außerdem sind Ursache und Wirkung in der Alltagswelt meist nicht zeitgleich zu beobachten, sondern in einem unterschiedlichen Ausmaß zeitlich versetzt. Im Unterschied zu Laborbedingungen, wo die Zahl der zu beobachtenden Ereignisse meist klar begrenzt ist, sind Menschen in der Alltagswelt nicht dazu in der Lage, statistische

Zusammenhänge zwischen Ereignissen zu bestimmen, da ihnen meist nicht klar ist, welche Ereignisse und Informationen für die Beurteilung eines Kausalzusammenhangs relevant sind und für den Urteilsprozess herangezogen werden sollen. In der Alltagswelt sind die Menschen einem kontinuierlichen Strom von Ereignissen ausgesetzt. Um regel-basierte Strategien einsetzen oder statistische Zusammenhänge ermitteln zu können, müssen sie diesen Ereignisstrom erst einmal strukturieren (s.a. Hagmayer, 2001; Hagmayer & Waldmann, 2002), d.h. sie müssen identifizieren, welche Ereignisse für sie Ursache(n) und Wirkung(en) repräsentieren.

Unabhängig von den im Rahmen dieser Arbeit durchgeführten Experimenten vermuten wir, dass sich die Forschung im Bereich Erwerb von Kausalwissen verstärkt dem Erwerb von Wissen über kausale Zusammenhänge in alltagsnahen Umwelten mit kontinuierlichen Ereignisströmen zuwenden wird. Neuere Forschungsergebnisse auf diesem Gebiet weisen darauf hin, dass zeitliche Verzögerungen zwischen den zu beurteilenden Ereignissen dazu führen, dass kausale Zusammenhänge schwächer beurteilt werden. Dieser Effekt kann aber reduziert oder auch teilweise ganz verhindert werden, wenn den Probanden Informationen zum zeitlichen Rahmen des zu beurteilenden Kausalzusammenhangs vorgegeben werden (Buehner & May, 2003b). Darüber hinaus zeichnet sich in der Forschungsliteratur der Trend ab, dass mehr und mehr von einer Interaktion zwischen grundlegenden assoziationalistischen Mechanismen und (vor-)wissensgesteuerten Prozessen, die unter bestimmten und noch weiter zu untersuchenden Umständen regel-basiert ablaufen können, ausgegangen wird (Baker, Murphy, Vallée-Tourangeau, 1996; Catena, Maldonado & Cándido, 1998; De Houwer & Beckers, 2002; Catena & Perales, 2002; Perales, Catena & Maldonado, 2004; Perales, Catena Shanks & González, 2005; Pineño & Miller, 2007; Price & Yates, 1995; Shanks, 2007; Tangen & Allan, 2004; Young, 1995). Von zunehmender Bedeutung werden hierbei Bayesianische Netzwerkmodelle (Glymour, 1998, 2001; Gopnik, Glymour, Sobel, Schulz, Kushnir & Danks, 2004; Gopnik, Sobel, Schulz & Glymour, 2001; Griffiths & Tenenbaum, 2005; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003; Tenenbaum & Griffiths, 2001, 2003) und die Theorie der Kausalmodelle (Hagmayer, 2001; Hagmayer & Waldmann, 2006; Waldmann & Holyoak, 1992) sein.





## Literaturverzeichnis

- Ahn, W.-K., & Bailenson, J. (1996). Causal attribution as a search for underlying mechanisms: An explanation of the conjunction fallacy and the discounting principle, *Cognitive Psychology*, *31*, 82-123.
- Ahn, W., & Kalish, C. W. (2000). The Role of Mechanism Beliefs in Causal Reasoning. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and Cognition*. Cambridge, MA: MIT Press.
- Ahn, W.-K., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, *54*, 299-352.
- Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*, 131-152.
- Ainsworth, S., Bibby, P., & Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *Journal of the Learning Sciences*, *11*, 25-61.
- Ainsworth, S., & Loizou, A.T. (2003). The effects of self explaining when learning with text or diagram. *Cognitive Science*, *27*, 669-681.
- Aitken, M. R. F., Larkin, M. J. W., & Dickinson, A. (2000). Super-learning of causal judgements. *Quarterly Journal of Experimental Psychology*, *53B*, 59-81.
- Aitken, M. R. F., Larkin, M. J. W., & Dickinson, A. (2001). Re-examination of the role of within-compound associations in the retrospective revaluation of causal judgements. *Quarterly Journal of Experimental Psychology*, *54B*, 27-51.
- Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, *15*, 147-149.
- Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin*, *114*, 435-448.
- Allan, L.G. (2003). Assessing power PC. *Learning and Behavior*, *31* (2), 192-204.
- Allan, L. G., & Jenkins, H. M.(1980). The judgment of contingency and the nature of the response alternatives. *Canadian Journal of Psychology*, *34*, 1-11.

- Allan, L. G., & Jenkins, H. M. (1983). The effect of representations of binary variables on judgment of influence. *Learning and Motivation, 14*, 381-405.
- Allan, L. G., Tangen, J. M., Wood, R., & Shah, T. (2003). Temporal contiguity and contingency judgments: A Pavlovian Analogue. *Integrative Physiological and Behavior Science, 38* (3), 214-229.
- Alloy, L. B., & Abramson, L. J. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General, 108*, 441-485.
- Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review, 91*, 112-149.
- Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R., & Sheu, C.-F. (1995). Causal inferences as perceptual judgments. *Memory and Cognition, 23*, 510-524.
- Aristoteles (1970). *Metaphysik*. Übersetzt und herausgegeben von Franz F. Schwarz. Stuttgart: Reclam-Verlag.
- Arkes, H. R., & Harkness, A. R. (1983). Estimates of contingency between two dichotomous variables. *Journal of Experimental Psychology: General, 112*, 117-135.
- Aschenbrenner, K.M. (1978). Single peaked risk preferences and their dependability on the gambles' presentation mode. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 513-520.
- Baddeley, A.D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A.D. (1992). Working memory. *Science, 255*, 556-559.
- Baddeley, A.D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A*, 5-28.
- Baddeley, A. D. (1999). *Human memory*. Boston: Allyn & Bacon.
- Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), *The psychology of learning and motivation, 8*, (pp. 47-89). New York: Academic Press.

- Baddeley, A. & Logie, R.H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). Cambridge: Cambridge University Press.
- Baker, A. G., Berbrier, M., & Vallee-Tourangeau, F. (1989). Judgments of a 2 x 2 contingency table: Sequential processing and the learning curve. *Quarterly Journal of Experimental Psychology, 41B*, 65-97.
- Baker, A. G., Mercier, P., Vallée-Tourangeau, F., Frank, R., & Pan, M. (1993). Selective associations and causality judgments: Presence of a strong causal factor may reduce judgments of a weaker one. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 414-432.
- Baker, A. G., Murphy, R. A., & Vallée-Tourangeau, F. (1996). Associative and normative models of causal induction: Reacting to versus understanding cause. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol. 34: Causal learning* (pp. 1–45). San Diego, CA: Academic Press.
- Baker, A. G., Vallee Tourangeau, F., & Murphy, R. A. (2000). Asymptotic judgment of cause in a relative validity paradigm. *Memory and Cognition, 28(3)*, 466-479.
- Ballstaedt, S.-P. (1997). *Wissensvermittlung. Die Gestaltung von Lernmaterial*. Weinheim: PVU.
- Ballstaedt, St.-P., Mandl, H., Schnotz, W., & Tergan, S.-O. (1981). *Texte verstehen – Texte gestalten*. München: Urban & Schwarzenberg.
- Beach, L.R. & Mitchell, T.R. (1987). A contingency model for the selection of decision strategies. *Academy of Management Review, 3*, 439-449.
- Beckers, T., De Houwer, J., Pineño, O., & Miller, R.R. (2005). Outcome additivity and outcome maximality influence cue competition in human causal learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31 (2)*, 238–249.
- Benbasat, I., Dexter, A. S., & Todd, P. (1986). An experimental program investigating color-enhanced and graphical information presentation: An integration of the findings. *Communications of the ACM, 29*, 1095–1105.
- Bettmann, J.R. (1979). *An information processing theory of customer choice*. London: Addison-Wesley.

- Beyth-Marom, R. (1982). Perception of correlation reexamined. *Memory and Cognition*, 10, 511-519.
- Bortz, J. (1984). *Lehrbuch der empirischen Forschung: für Sozialwissenschaftler*. Springer-Verlag: Berlin Heidelberg New York Tokyo.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Berlin: Springer-Verlag.
- Bröder, A. & Eichler, A. (2001). *Individuelle Unterschiede in bevorzugten Entscheidungsstrategien*. Manuskript, Psychologisches Institut der Universität Bonn, 2001.
- Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.
- Buehner, M. J. (2001). Inducing causation: covariation assessment and the assumption of causal power. In M. May & U. Oestermeier (Eds.), *Interdisciplinary perspectives on causation* (pp. 33-58). Bern.
- Buehner, M. J., & Cheng, P. W. (1997). Causal induction: The power PC theory versus the Rescorla-Wagner theory. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (pp. 55-60). Mahwah, NJ: Erlbaum.
- Buehner, M. J., Cheng, P. W., & Clifford, D. (2003). From covariation to causation: A test of the assumption of causal power. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1119-1140.
- Buehner, M. J., & May, J. (2002). Knowledge mediates the timeframe of covariation assessment in human causal induction. *Thinking and Reasoning*, 8(4), 269-295.
- Buehner, M. J., & May, J. (2003a). Rethinking temporal contiguity and the judgement of causality: Effects of prior knowledge, experience, and reinforcement procedure. *Quarterly Journal of Experimental Psychology*, 56A, 865-890.
- Buehner, M.J. & May, J. (2003b). *Causal Induction from Continuous Event Streams. Implications from Associative Learning Theory*. Associative Learning Symposium VII, Gregynog, Wales, UK, April 15th-17th 2003.
- Buehner, M. J., & May, J. (2004). Abolishing the effect of reinforcement delay on human causal learning. *Quarterly Journal of Experimental Psychology*, 57B, 179-191.

- Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209-254). New York: Academic Press.
- Bunge, M. (1987). *Kausalität, Geschichte und Probleme*. Tübingen, Mohr.
- Busemeyer, J.R. (1991). Intuitive statistical estimation. In N.H. Anderson (Ed.): *Contributions to information integrating theory, 1*, 187-215. Hillsdale, NJ: Erlbaum.
- Bußmann, H. (2002). *Lexikon der Sprachwissenschaft*. Dritte, aktualisierte und erweiterte Auflage. Stuttgart: Kröner-Verlag.
- Carpenter, P.A., & Shah, P. (1998): A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, 4, 75-100.
- Carswell, C.M., Emery, C., & Lonan, A.M. (1993). Stimulus complexity and information integration in the spontaneous interpretation of line graphs. *Applied Cognitive Psychology*, 7, 341-357.
- Carswell, C.M., & Wickens, C.D. (1987). Information integration and the object display: An interaction of task demands and display superiority. *Ergonomics*, 30, 511-527.
- Carter, L.F. (1947). An experiment on the design of tables and graphs used for presenting numerical data. *Journal of Applied Psychology*, 31, 640-650
- Cartwright, N. (1989). *Nature's capacities and their measurement*. Oxford: Clarendon Press.
- Catena, A., Maldonado, A., & Cándido, A. (1998). The effect of the frequency of judgment and the type of trials on covariation learning. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 481-495.
- Catena, A., Maldonado, A., Megías, J.L., & Frese, B (2002). The frequency of judgment effect: Belief revision and serial processing of causal information. *Quarterly Journal of Experimental Psychology*, 55(B), 267-281
- Catena, A. & Perales, J.C. (2002). *The Psychology of Causality: An overview*. Unpublished manuscript. (pdf in English).
- Catena, A., Perales, J.C., & Maldonado, A. (2004). Judgment frequency effects in generative and preventative causal learning. *Psicológica*, 25, 67-85.
- Chambers, D., & Reisberg, D. (1985). Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception and Performance*, 11, 317-328.

- Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 837-854.
- Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition*, *18*, 537-545.
- Cheng, P. W. (1993). Separating causal laws from casual facts: Pressing the limits of statistical relevance. In D. L. Medin (Ed.), *The psychology of learning and motivation*, Vol. 30, pp. 215-264. San Diego: Academic Press.
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367-405.
- Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391-416.
- Cheng, P. W., & Holyoak, K. J. (1995). Complex adaptive systems as intuitive statisticians: Causality, contingency, and prediction. In H. L. Roitblat & J.-A. Meyer (Eds.), *Comparative approaches to cognitive science* (pp. 271-302). Cambridge, MA: MIT Press.
- Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology*, *58*, 545-567.
- Cheng, P. W., & Novick, L. R. (1991). Causes versus enabling conditions. *Cognition*, *40*, 83-120.
- Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, *99*, 365-382.
- Cheng, P.W., & Novick, L.R. (2005). Constraints and nonconstraints in causal learning: Reply to White (2005) and to Luhmann and Ahn (2005). *Psychological Review*, *112*(3), 694-707.
- Cheng, P.W., Park, J., Yarlas, A.S., & Holyoak, K. J. (1996). A causal-power theory of focal sets. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation*, Vol. 34: Causal learning (pp. 47-88). San Diego: Academic Press.
- Clark, A. (1997). *Being there: putting brain, body, and world together again*. Cambridge, MA: MIT Press.
- Clark, R.E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, *53*, 445-459.

- Clark, R.E. (1994). Media will never influence learning. *Educational Technology Research and Development*, 42 (2), 21-30.
- Clark, R.E., & Salomon, G. (1986). Media in teaching. In M.C. Wittrock (Ed.), *Handbook of Research on Teaching* (3<sup>rd</sup> edition, pp. 464-478). New York: Macmillan.
- Cleveland, W.S. (1993). *Visualizing data*. AT&T Bell Laboratories, Murray Hill, NJ.
- Cobos, P. L., López, F. J., Cano, A., Almaraz, J., & Shanks, D. R. (2002). Mechanisms of predictive and diagnostic causal induction. *Journal of Experimental Psychology: Animal Behavior Processes*, 28, 331-346.
- Coll, J. H. (1992). An experimental study of the efficacy of tables versus bar graphs with respect to type of task. *Information Management*, 23, 45-51.
- Coll, J.H., & Coll, R.A. (1993). Tables and graphs: a classification scheme for display presentation variables and a framework for research in this area. *Information Processing & Management*, 29 (6), 745-750.
- Coll, R. A., Coll, J. H., & Thankur G. (1994). Graphs and tables. A four-factor experiment. *Communications of the ACM*, 37 (4), 77-86
- Coll, R., Thyagarajan, A., & Chopra, S. (1991). An experimental study comparing the effectiveness of computer graphics data versus computer tabular data. *IEEE Transactions on Systems, Man and Cybernetics*, 21 (4), 897-900.
- Collins, D. J., & Shanks, D. R. (2002). Momentary and integrative response strategies in causal judgment. *Memory & Cognition*, 30, 1138–1147.
- Collins, D. J., & Shanks, D. R. (2006). Conformity to the power PC theory of causal induction depends on the type of probe question. *The Quarterly Journal of Experimental Psychology*, 59 (2), 225–232.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, 58, 1–73.
- Crocker, J. (1981). Judgment of covariation by social perceivers. *Psychological Bulletin*, 90, 272-292.
- Crocker, J. (1982). Biased questions in judgment of covariation studies. *Personality and Social Psychology Bulletin*, 8, 214-220.

- Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, *19*, 450-466.
- Danks, D. (2005). Causal learning from observations and manipulations. In M. Lovett, & P. Shah (Eds.), *Thinking with data*. Mahwah, N.J.: Lawrence Erlbaum Associates, Inc.
- De Houwer, J. (2002). Forward blocking depends on retrospective inferences about the presence of the blocked cue during the elemental phase. *Memory & Cognition*, *30*, 24-33.
- De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theories on human contingency learning. *Quarterly Journal of Experimental Psychology*, *55B*, 289-310.
- De Houwer, J., & Beckers, T. (2003). Secondary task difficulty modulates forward blocking in human contingency learning. *Quarterly Journal of Experimental Psychology: Comparative & Physiological Psychology*, *56B*, 345-357.
- De Houwer, J., Beckers, T., & Glauthier, S. (2002). Outcome and cue properties modulate blocking. *Quarterly Journal of Experimental Psychology*, *55A*, 965-985.
- De Houwer, J., Beckers, T., & Vandorpe, S. (2005). Evidence for the role of higher order reasoning processes in cue competition and other learning phenomena. *Learning & Behavior*, *33*, 239-249.
- De Jong, T., Ainsworth, S., Dobson, M., van der Hulst, A., Levonen, J., Reimann, P., Sime, J.-A., van Someren, M.W., Spada, H., & Swaak, J. (1998). Acquiring Knowledge in Science and Mathematics: The Use of Multiple Representations in Technology-Based Learning Environments. In M.W. van Someren, P. Reimann, H.P.A. Boshuizen, & T. de Jong (Eds.), *Learning with Multiple Representations. Advances in Learning and Instruction Series* (pp. 9-40). Amsterdam: Pergamon.
- Dennis, M. J., & Ahn, W.-K. (2001). Primacy in causal strength judgments: The effect of initial evidence for generative versus inhibitory relationships. *Memory and Cognition*, *29*, 152-164.
- De Sanctis, G. (1984). Computer graphics as decision aids: directions for research. *Decision Sciences*, *15*, 463 - 487.
- Dickinson, A. (2001). Causal learning: An associative analysis. *Quarterly Journal of Experimental Psychology*, *54B*, 3-25.



- Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective reevaluation of causality judgements. *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology*, 49 (B), 60-80.
- Dickinson, A., Shanks, D. R., & Evenden, J. L. (1984). Judgment of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, 36A, 29-50.
- Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Cambridge, MA: Harvard University Press.
- Downing, C.J., Sternberg, R.J., & Ross, B.H. (1985). Multicausal inference: Evaluation of evidence in causally complex situations. *Journal of Experimental Psychology: General*, 114, 239-263.
- Dunnett, C.W. (1955). A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association*, 50, 1096-1121.
- Eells, E. (1991). *Probabilistic causality*. Cambridge: Cambridge University Press.
- Einhorn, H.J. & Hogarth, R.M. (1986). Judging Probable Cause. *Psychological Bulletin*, 99 (1), 3-19.
- Ericsson, K.A., & Dulaney, P.F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 257-297). Cambridge: Cambridge University Press.
- Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research*, 50, 123-129.
- Fiedler, K. (2000). Beware of samples: A cognitive-ecological sampling approach to judgment biases. *Psychological Review*, 107(4), 659-676.
- Fiedler, K., Brinkmann, B., Betsch, T., & Wild, B. (2000). A sampling approach to biases in conditional probability judgments: Beyond base rate neglect and statistical format. *Journal of Experimental Psychology: General*, 129, 399-418.
- Fischer, K. & Jungermann, H. (1996). Rarely occurring headaches and rarely occurring blindness: Is rarely = rarely? The meaning of verbal frequentistic labels in specific medical contexts. *Journal of Behavioral Decision Making*, 9, 153-172.

- Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Ford, G.T. & Smith, R.A. (1987). Inferential beliefs in consumer evaluations: An assessment of alternative processing strategies. *Journal of Consumer Research*, 14, 363-371.
- Fratianne, A., & Cheng, P. W. (1995). *Assessing causal relations by dynamic hypothesis testing*. Unpublished manuscript.
- Freedman, E.G., & Shah, P. (2002). Toward a model of knowledge-based graph comprehension. In M. Hegarty, B. Meyer, and N. Hari Narayanan (Eds.), *Diagrams 2002, LNAI* (pp. 18-30). Springer-Verlag: Berlin Heidelberg.
- Freedman, E.G., & Smith, L.D. (1996). The role of theory and data in covariation assessment: Implications for the theory-ladenness of observation. *Journal of Mind and Behavior*, 17, 321-343.
- Friege, G. & Lind, G. (2003). Allgemeine und fachspezifische Problemlösekompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, 9, 63-74.
- Fum, D. & Stocco, A. (2003). The role of compound cues in causal judgment: Associative and probabilistic effects. In F. Schmalhofer, R. M. Young & G. Katz (Eds.), *Proceedings of the EuroCogSci 03* (pp. 127-132). Mahwah, NJ: Lawrence Erlbaum.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102, 684-704.
- Glaser, W.R. (1978). *Varianzanalyse*. UTB 584. Fischer, Stuttgart.
- Glymour, C. (1998). Learning causes: Psychological explanations of causal explanation. *Minds and Machines*, 8, 39-60.
- Glymour, C. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology*. Cambridge, MA: MIT Press.
- Glymour, C. (2003). Learning, prediction and causal Bayes nets. *Trends in Cognitive Science*, 7, 43-48.
- Glymour, C., & Cooper, G. F. (1999). *Computation, causation, and discovery*. Menlo Park, MA: MIT Press.

- Goody, J. (1977). *The Domestication of the Savage Mind*. Cambridge: Cambridge University Press.
- Goody, J. (1987). *The interface between the written and the oral*. Cambridge: Cambridge University Press.
- Goolkasian, P., & Foos, P.W. (2002). Presentation format and its effect on working memory. *Memory and Cognition*, 30 (7), 1096-1105.
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, 111, 3-32.
- Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two, three, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, 37, 5, 620–629.
- Greeno, J.G. (1989). Situations, mental models, and generative knowledge. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon*. Hillsdale, NJ: Erlbaum.
- Greeno, J.G. (1994). Gibson's affordances. *Psychological Review*, 101(2), 336-342.
- Griffiths, T.L., & Tenenbaum, J.B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, 51, 334–384.
- Häcker, H., & Stapf, K.-H. (2004). *Dorsch Psychologisches Wörterbuch*. 14. überarbeitete und erweiterte Auflage. Verlag Hans Huber: Bern Göttingen Toronto Seattle.
- Hagmayer, Y. (2001). *Denken mit und über Kausalmodelle*. Dissertationsschrift.
- Hagmayer, Y., & Waldmann, M. R. (2000). Simulating causal models: The way to structural sensitivity. In L. Gleitman & A. Joshi (Eds.), *Proceedings of the Twenty-second Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.
- Hagmayer, Y., & Waldmann, M. R. (2002). How temporal assumptions influence causal judgments. *Memory & Cognition*, 30, 1128-1137.
- Hagmayer, Y. & Waldmann, M.R. (2006). Kausales Denken. In J. Funke (Hrsg.), *Enzyklopädie der Psychologie "Denken und Problemlösen"*, Band C/II/8 (S. 87-166). Göttingen: Hogrefe.
- Hart, H.L., & Honoré, A.M. (1985). *Causation in the law*. 2<sup>nd</sup> edition. Oxford: University Press (original work published in 1959).

- Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review*, 93 (3), 258-268.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Wiley.
- Höffe, O. (1996). *Immanuel Kant*. Vierte durchgesehene Auflage. München: Beck.
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, 24, 1–55.
- Hume, D. (1975). *An Enquiry concerning Human Understanding*. Oxford: Oxford University Press (Original work published 1748).
- Hume, D. (1978). *A Treatise of Human Nature* (2<sup>nd</sup> edn.). Oxford: Clarendon Press. (Original work published 1739).
- Hutchins, E. (1995a). *Cognition in the wild*. Cambridge: MIT Press.
- Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cognitive Science*, 19, 265-288.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. London: Routledge and Kegan Paul.
- Jacobs, B. (1990). Ein Vergleich der Auswirkungen graphischer und tabellarischer Präsentationsformen auf die Schnelligkeit und Genauigkeit beim Erkennen und Interpretieren statistischer Daten. *Arbeitsberichte aus dem Medienzentrum der Philosophischen Fakultät der Universität des Saarlandes*, Nr. 3.
- Jacobs, B. (1994). Graphische vs. tabellarische Präsentation von statistischen Daten. *Zeitschrift für Pädagogische Psychologie*, 8, 73-84.
- Jacobs, B. (1999). *Tabelle oder Graphik – Was ist besser?* Erstellt: 05.01.1999. Verfügbar unter: <http://www.phil.uni-sb.de/~jakobs/wwwartikel/tablgraf/index.htm> [16.04.2007].
- Jäger, A.O. (1984). Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven. *Psychologische Rundschau*, 35, 21-35.
- Jäger, A.O., Süß, H.-M. & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test. BIS-Test. Form 4, Handanweisung*. Göttingen: Hogrefe.
- Jarvenpaa, S.L., & Dickson, G.W. (1988). Graphics and managerial decision making: research based guidelines. *Communications of the ACM*, 31, 764-774.

- Jarvenpaa, S. L., Dickson, G. W. & De Sanctis, G. (1985). Methodological issues in experimental IS research: experiences and recommendations. *MIS Quarterly*, 9, 141-156.
- Jenkins, H. M., & Ward, W. C. (1965). Judgement of contingency between responses and outcomes. *Psychological Monographs*, 79 (Whole No. 10).
- Jennings, D.L., Amabile, T., & Ross, L. (1982). Informal covariation assessment: Data-based versus theory-based judgments. In D. Kahnemann, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Jonassen, D.H., Campbell, J.P., & Davidson, M.E. (1994). Learning with media: Restructuring the debate. *Educational Technology Research and Development*, 42 (2), 31-39.
- Jungermann, H., Pfister, H.-R., & Fischer, K. (2005). *Die Psychologie der Entscheidung. Eine Einführung*. Zweite Auflage. Elsevier Spektrum Akademischer Verlag.
- Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: individual differences in working memory: *Psychological Review*, 99, 122-149.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, 103, 582-591.
- Kaiser, M.K., & Proffitt, D.R. (1984). The development of sensitivity to causally relevant dynamic information. *Child Development*, 55, 1614-1624.
- Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9-33). Miami: University of Miami Press.
- Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 276-296). New York: Appleton-Century-Crofts.
- Kant, I. (1974). *Kritik der reinen Vernunft, Bd 1*. Herausgegeben von Wilhelm Weischedel. Frankfurt am Main: Suhrkamp Taschenbuch Verlag (Original aus dem Jahre 1781).

- Kant, I. (1976). *Prolegomena zu einer jeden künftigen Metaphysik*. Herausgegeben von Karl Vorländer. Hamburg: Felix Meiner Verlag (Original aus dem Jahre 1783).
- Kao, S.-F., & Wasserman, E. A. (1993). Assessment of an information integration account of contingency judgment with examination of subjective cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1363-1386.
- Kelley, H. H. (1967). Attribution in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (Vol.15, S. 192-238). Lincoln, NB: University of Nebraska Press.
- Kelley, H. H. (1972). Causal schemata and the attribution process. In E. E. Jones, D. E. Knaouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution, Perceiving the causes of behavior*. Morristown, NJ: General Learning Press.
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, 28, 107-128.
- Kerzel, D., & Hecht, H. (2001). Visual causality. In M. May & U. Oestermeier (Eds.), *Interdisciplinary perspectives on Causation* (pp. 119-140). Bern.
- Kirk, R.E. (1995). *Experimental Design: Procedures for the Behavioral Sciences* (3<sup>rd</sup> edition). Brooks / Cole Publishing company, Pacific Grove, California.
- Kleinmuntz, D. J., & Schkade, D.A. (1993). Information displays and decision processes. *Psychological Science*, 4, 221-227.
- Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. Cambridge, MA: MIT Press.
- Kosslyn, S.M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology*, 3, 185 - 226.
- Kosslyn, S.M. (1994). *Image and Brain. The resolution of the imagery debate*. Cambridge: The MIT Press.
- Kozma, R.B. (1994). Will media influence learning? Reframing the debate. *Education Technology Research and Development*, 42, 7-19.
- Lagnado, D. A., & Shanks, D.R. (2002). Probability judgment in hierarchical learning: a conflict between predictiveness and coherence. *Cognition*, 83, 81-112.

- Lagnado, D. A., & Sloman, S. (2004). The advantage of timely interventions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30 (4), 856-876.
- Larkin, J. H., & Simon, H.A. (1987). Why a diagram is sometimes worth ten thousand words. *Cognitive Science*, 11, 65-99.
- Larkin, M.J.W., Aitken, M.R.F., & Dickinson, A. (1998). Retrospective reevaluation of causal judgments under positive and negative contingencies. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, 24, 1331-1352.
- Leslie, A.M. (1982). The perception of causality in infants. *Perception*, 11, 173-186.
- Leslie, A.M. (1995). A theory of agency. In D. Sperber, D. Premack & A.J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 121-141). New York, NY: Clarendon.
- Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, 25, 265-288.
- Levin, I. P., Wasserman, E. A., & Kao, S.-F. (1993). Multiple methods for examining biased information use in contingency judgments. *Organizational Behavior and Human Decision Processes*, 55, 228-250.
- Lien, Y., & Cheng, P. W. (2000). Distinguishing genuine from spurious causes: A coherence hypothesis. *Cognitive Psychology*, 40, 87-137.
- Lipe, M. G. (1991). Counterfactual Reasoning as a Framework for Attribution Theories. *Psychological Bulletin*, 109(3), 456-471.
- Lober, K., & Shanks, D. R. (2000). Is causal induction based on causal power? Critique of Cheng (1997). *Psychological Review*, 107, 195-212.
- Lohse, G.L. (1993a). A cognitive model for understanding graphical perception. *Human-Computer Interaction*, 8, 353-388.
- Lohse, G.L. (1993b). Eye movement-based analyses of graphs and tables: the next generation. *Proceedings of the 14th International Conference on Information Systems* (Orlando, FL: ICIS), 213-224.
- Lohse, G.L. (1997). The role of working memory on graphical information processing. *Behaviour & Information Technology*, 16 (6), 297-308.

- López, F. J., Almaraz, J., Fernandez, P., & Shanks, D. R. (1999). Adquisición progresiva del conocimiento sobre relaciones predictivas: Curvas de aprendizaje en juicios de contingencia [Gradual acquisition of knowledge of predictive relationships: Learning curves in contingency judgments]. *Psicothema*, *11*, 337-349.
- López, F. J., Cobos, P. L., Cano, A., & Shanks, D. R. (1998). The rational analysis of human causal and probability judgment. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 314–352). Oxford: Oxford University Press.
- López, F. J., Shanks, D. R., Almaraz, J., & Fernández, P. (1998). Effects of trial order on contingency judgments: A comparison of associative and probabilistic contrast accounts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 67.
- Lovibond, P. F. (2003). Causal beliefs and conditioned responses: Retrospective reevaluation induced by experience and by instruction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 97–106.
- Lovibond, P. F., Been, S.-L., Mitchell, C. J., Bouton, M. E., & Frohardt, R. (2003). Forward and backward blocking of causal judgment is enhanced by additivity of effect magnitude. *Memory and Cognition*, *31*, 133–142.
- Lowe, R.K. (1993). Constructing a mental representation from an abstract technical diagram. *Learning and Instruction*, *3* (3), 157-179.
- Luhmann, C.C., & Ahn, W. (2005). The meaning and computation of causal power: Comment on Cheng (1997) and Novick and Cheng (2004). *Psychological Review*, *112*(3), 685-693.
- Macho, S., & Burkart, J. (2002). Recursive retrospective reevaluation of causal judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 1171-1186.
- Mackie, J.L. (1965). Causes and conditions. *American Philosophical Quarterly*, *ii*, 245-264.
- Mackie, J. L. (1974). *The cement of the universe*. Oxford: Clarendon-Press.
- Maccoby, E., & Jacklin, C. (1974). *The psychology of sex differences*. Stanford, California: Stanford University Press.



- Maichle, U. (1992). Kognitive Prozesse bei der Interpretation von Zeitdiagrammen: Eine experimentelle Studie zum Untertest "Diagramme und Tabellen". In G. Trost (Hrsg.), *Test für medizinische Studiengänge (TMS). Studien zur Evaluation. 16. Arbeitsbericht.* Bonn.
- Maichle, U. (1994). Cognitive processes in understanding line graphs. In W. Schnotz & R.W. Kulhavy (Eds.), *Comprehension of graphics* (pp. 207-226). North-Holland: Elsevier.
- Mandel, D. R., & Lehman, D. R. (1998). Integration of contingency information in judgements of cause, covariation, and probability. *Journal of Experimental Psychology: General, 127*, 269–285.
- Markus, G.F., Vijayan, S., Rao, S.B., and Vishton, P.M. (1999). Rule learning by seven month old infants. *Science, 283*, 77-80.
- Marr, D. (1982). *Vision*. San Francisco, CA: W.H. Freeman.
- Matute, H., Arcediano, F., & Miller, R. R. (1996). Test question modulates cue competition between causes and between effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 182-196.
- Matute, H., Vegas, S., & de Marez, P. (2002). Flexible use of recent information in causal and predictive judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 714–725.
- May, M., & Oestermeier, U. (2001). *Interdisciplinary perspectives on causation*. Bern Studies in the History and Philosophy of Science.
- Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*, 1-19.
- Mayer, R. E. (2001). *Multimedia learning*. New York, NY: Cambridge University Press.
- McKenzie, C. R. M. (1994). The accuracy of intuitive judgment strategies: Covariation assessment and Bayesian inference. *Cognitive Psychology, 26*, 209–239.
- Meixner, U. (2001). *Theorie der Kausalität: Ein Leitfaden zum Kausalbegriff in zwei Teilen*. Paderborn: Mentis.
- Melchers, K. G., Lachnit, H., & Shanks, D. R. (2004). Within-compound associations in retrospective revaluation and in direct learning: A challenge for comparator theory. *Quarterly Journal of Experimental Psychology, 57B*, 25-53.

- Melchers, K.G., Üngör, M., & Lachnit, H. (2005). The experimental task influences cue competition in human causal learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 31(4), 477-483.
- Melz, E. R., Cheng, P. W., Holyoak, K. J., & Waldmann, M. R. (1993). Cue competition in human categorization: Contingency or the Rescorla-Wagner rule. comments on Shanks (1991). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1398-1410.
- Mendelson, R., & Shultz, T. R. (1976). Covariation and temporal contiguity as principles of causal inference in young children. *Journal of Experimental Child Psychology*, 22(3), 408-412.
- Meyer, J. (1997). A new look at an old study on information display: Washburne (1927) reconsidered, *Human Factors*, 39, 333-340.
- Meyer, J. (2000). Performance with tables and graphs: effects of training and a Visual Search Model. *Ergonomics*, 43 (11), 1840-1865.
- Meyer, J., Shinar, D., & Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors*, 39, 268 - 286.
- Michotte, A. E. (1963). *The perception of causality*. New York, NY: Basic Books
- Mill, J.S. (1979). *A system of logic ratiocinative and inductive, collected works VII & VIII*. Toronto: Toronto University Press (Original work published 1843).
- Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge: Cambridge University Press.
- Morris, M.W., Nisbett, R.E., & Peng, K. (1995). Causal attribution across domains and cultures. In D. Sperber, D. Premack & A.J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 577-612). New York, NY: Clarendon.
- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Noel, M.-P., Désert, M., Aubrun, A., & Seron, X. (2001). Involvement of short-term memory in complex mental calculation. *Memory & Cognition*, 29 (1), 34-42.
- Norman, D. A. (1988). *The psychology of everyday things*. New York: Basic Books.
- Norman, D. A. (1991). Cognitive Artifacts. In J. M. Carroll (Ed.), *Designing interaction: Psychology at the human-computer interface*. New York: Cambridge University Press.

- Norman, D. A. (1993). *Things that make us smart*. Reading, MA: Addison-Wesley.
- Novick, L. R., & Cheng, P. W. (2004). Assessing interactive causal power. *Psychological Review*, 111, 455-485.
- Oberauer, K., Demmrich, A., Mayr, U., & Kliegl, R. (2001). Dissociating retention and access in working memory: An age-comparative study of mental arithmetic. *Memory & Cognition*, 29 (1), 18-33.
- Oestermeier, U. (1997). Begriffliche und empirische Fragen der Kausalkognition. *Kognitionswissenschaft*, 6(2), 70-85.
- Oestermeier, U. (2001). The mathematisation of the sciences and the observability of causal relations. In M. May & U. Oestermeier (Eds.), *Interdisciplinary perspectives on causation* (pp. 141-158). Bern.
- Oestermeier, U., & Hesse, F.W. (2000). Verbal and visual causal arguments. *Cognition*, 75, 65-104.
- Olson, D.R. (1996). Towards a psychology of literacy: on the relations between speech and writing. *Cognition*, 60, 83-104.
- Ong, W.J. (1982). *Orality and literacy: The technologizing of the world*. London: Methuen.
- Opwis, K., & Lüer, G. (1996). Modelle der Repräsentation von Wissen. In D. Albert & K.-H. Stapf (Hrsg.), *Enzyklopädie der Psychologie. Themenbereich C: Theorie und Forschung. Serie II: Kognition. Band 4: Gedächtnis* (S. 339-433). Göttingen: Hogrefe.
- Park, J.-K., & Cheng, P. W. (1995). *Boundary conditions on overexpectation in causal learning with discrete trials: A test of the power PC theory*. Unpublished manuscript.
- Payne, J.W., Bettman, J.R., & Johnson, E.J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.
- Payne, J.W., Bettman, J.R., & Johnson, E.J. (1990). The adaptive decision maker: Effort and accuracy in choice. In R. M. Hogarth (ed.), *Insights in decision making: A tribute to Hillel J. Einhorn*. Chicago: The University of Chicago Press.
- Pearce, J.M.(1987).A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, 94, 61–73.
- Pearce, J. M. (1994). Similarity and discrimination – A selective review and a connectionist model. *Psychological Review*, 101, 587-607.

- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87(6), 532-552.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo, CA: Morgan Kaufmann Publishers.
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge, UK: Cambridge University Press.
- Perales, J.C., Catena, A., and Maldonado, A. (2004). Inferring non-observed correlations from causal scenarios: The role of causal knowledge. *Learning and Motivation*, 35, 115-135.
- Perales, J. C., Catena, A., Shanks, D. R., & González, J. A. (2005). Dissociation between judgments and outcome-expectancy measures in covariation learning: A Signal Detection Theory approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 1105-1120.
- Perales, J. C., & Shanks, D. (2003). Normative and descriptive accounts of the influence of power and contingency on causal judgment. *Quarterly Journal of Experimental Psychology*, 56A, 977-1007.
- Perales, J. C. & Shanks, D. R. (in press). Models of covariation-based causal judgment: A review and synthesis. *Psychonomic Bulletin & Review*.
- Peterson, C.R., & Beach, L.R. (1967). Man as an intuitive statistician. *Psychological Bulletin*, 68, 29-46.
- Pineño, O., & Miller, R. R. (2007). Comparing associative, statistical, and inferential reasoning accounts of human contingency learning. *Quarterly Journal of Experimental Psychology*, 60 (3), 310-329.
- Pinker, S. & Prince, A. (1988). On language and connectionism: Analysis of parallel distributed processing model of language acquisition. *Cognition*, 29, 73-193.
- Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73-126). Hillsdale: Erlbaum.
- Price, P. C., & Yates, J. F. (1993). Judgmental overshadowing: Further evidence of cue interaction in contingency judgment. *Memory & Cognition*, 21, 561-572.

- Price, P. C., & Yates, J. F. (1995). Associative and rule-based accounts of cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1639-1655.
- Reips, U.-D. (1997). *Erwerb von Kausalwissen: Lernrichtung und Diagnostizität von Effekten*. Unveröffentlichte Dissertation, Tübingen.
- Reisberg, D. (1987). *External representations and the advantages of externalizing one's thoughts*. Paper presented at the Proceedings of the Cognitive Science Society, Seattle, WA.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II. Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.
- Rickheit, G., & Strohner, H. (1993). *Grundlagen der kognitiven Sprachverarbeitung*. Tübingen/Basel: Francke.
- Rozenblit, L., & Keil, F. C. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, 26, 521-562.
- Ross, M., & Fletcher, G.J.O. (1985). Attribution and Social Perception. In G. Lindzey, & E. Aronson (Eds.), *The handbook of social psychology* (3<sup>rd</sup> edition), (pp. 73-122). New York: Random House.
- Rumelhart, D.E. (1975). Notes on a schema for stories. In D.G. Bobrow & A. Collings (Eds.), *Representation and understanding*. New York: Academic Press.
- Rumelhart, D.E. (1980). Schemata: The building blocks of cognition. In R.J. Spiro, B.C. Bruce, and W.F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension* (pp. 33-58). Hillsdale, NJ: Erlbaum.
- Rumelhart, D.E. (1986). Parallel distributed processing. Explorations in the microstructure of cognition. In D.E. Rumelhart, J.L. McClelland & the PDP research group (Eds.), *Foundations*. Cambridge, MA: MIT Press.
- Russo, J.E. (1977). The value of unit price information. *Journal of Marketing Research*, 14, 193-201.
- Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton, NJ: Princeton University Press.

- Salomon, G. (1979/1994). *Interaction of media, cognition, and learning*. Erlbaum, Hillsdale, NJ.
- Sanfey, A., & Hastie, R. (1998). Does evidence presentation format affect judgment? An experimental evaluation of displays of data for judgment. *Psychological Science, 9* (2), 99-103.
- Scaife, M., & Rogers, Y. (1996). External cognition: how do graphical representations work? *International Journal of Human-Computer Studies, 45*, 185-214.
- Schaller, M., & O'Brian, M. (1992). "Intuitive analysis of covariance" and group stereotype formation. *Personality and Social Psychology Bulletin, 18*, 776-785.
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.
- Schlottmann, A. (1999). Seeing in happen and knowing how it works: How children understand the relation between perceptual causality and underlying mechanism. *Developmental Psychology, 35*(5), 303-317.
- Schnotz, W. (2002). Wissenserwerb mit Texten, Bildern und Diagrammen. In L. J. Issing & P. Klimsa (Hrsg.), *Informationen und Lernen mit Multimedia und Internet* (3., vollständig überarbeitete Aufl.) (S. 65-81). Weinheim: Beltz, PVU.
- Schnotz, W. & Bannert, M. (1999). Einflüsse der Visualisierungsform auf die Konstruktion mentaler Modelle beim Text- und Bildverstehen. *Zeitschrift für experimentelle Psychologie, 46*, 217-236.
- Schustack, M. W., & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General, 110*, 101-120.
- Shah, P. (2001). Graph comprehension. The role of format, content, and individual differences. In M. Anderson, B. Meyer, and P. Olivier (Eds.), *Diagrammatic Representation and Reasoning* (pp. 173-185). Berlin: Springer-Verlag.
- Shah, P., Freedman, E., & Vekiri, I. (2005). The comprehension of quantitative information in graphical displays. In P. Shah & A. Miyake, (Eds.). *The Cambridge Handbook of Visuospatial Thinking* (pp. 426-476). New York: Cambridge University Press.
- Shah, P. & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review, 14* (1), 47-69.

- Shah, P., Mayer, R.E., & Hegarty, M. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology, 91*, 690-702.
- Shaklee, H. (1983). Human covariation judgment: Accuracy and strategy. *Learning and Motivation, 14*, 433-448.
- Shaklee, H., & Elek, S. (1988). Cause and covariate: Development of two related concepts. *Cognitive Development, 3*, 1-13.
- Shaklee, H., & Hall, L. (1983). Methods of assessing strategies for judging covariation between events. *Journal of Educational Psychology, 75* (4), 583-594.
- Shaklee, H., & Mims, M. (1981). Development of rule use in judgments of covariation between adults. *Child Development, 52*, 317-325.
- Shaklee, H., & Mims, M. (1982). Sources of error in judging event covariation: Effects of memory demands. *Journal of Experimental Psychology: Human Learning and Memory, 8*, 208-224.
- Shaklee, H., & Tucker, D. (1980). A rule analysis of judgments of covariation between events. *Memory & Cognition, 8*, 459-467.
- Shaklee, H., & Wasserman, E.A. (1986). Judging interevent contingencies: Being right for the wrong reasons. *Bulletin of the Psychonomic Society, 24* (2), 91-94.
- Shanks, D. R. (1985). Forward and backward blocking in human contingency judgment. *Quarterly Journal of Experimental Psychology, 37B*, 1-21.
- Shanks, D. R. (1987). Acquisition functions in contingency judgment. *Learning and Motivation, 18*, 147-166.
- Shanks, D.R. (1991). On similarities between causal judgments in experienced and described situations. *Psychological Science, 2*, 341-350.
- Shanks, D. R. (1993). Human instrumental learning: A critical review of data and theory. *British Journal of Psychology, 84*, 319-354.
- Shanks, D. R. (1995). *The psychology of associative learning*. Cambridge, UK: Cambridge University Press.
- Shanks, D. R. (2002). Tests of the power PC theory of causal induction with negative contingencies. *Experimental Psychology, 49*, 81-88.

- Shanks, D. R. (2004). Judging covariation and causation. In D.J. Koehler & N. Harvey (Eds.), *Handbook of judgment and decision making* (pp. 220-239). Oxford: Blackwell.
- Shanks, D. R. (2007). Associationism and Cognition: Human Contingency Learning at 25. *Quarterly Journal of Experimental Psychology*, 60 (3), 291-309.
- Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 21, pp. 229-261). New York: Academic Press.
- Shanks, D.R., & Dickinson, A. (1991). Instrumental judgment and performance under variations in action–outcome contingency and contiguity. *Memory and Cognition*, 19, 353–360.
- Shanks, D. R., Holyoak, K. J., & Medin, D. L. (Eds.), *The psychology of learning and motivation, Vol. 34: Causal learning*. San Diego, CA: Academic Press.
- Shanks, D. R., & Lopez, F. J. (1996). Causal order does not affect cue selection in human associative learning. *Memory and Cognition*, 24, 511-522.
- Shanks, D. R., López, F. J., Darby, R. J., & Dickinson, A. (1996). Distinguishing associative and probabilistic contrast theories of human contingency judgment. *The psychology of learning and motivation, Vol. 34: Causal learning*. (pp. 265–312). San Diego, CA: Academic Press.
- Shanks, D. R., Pearson, S. M., & Dickinson, A. (1989). Temporal contiguity and the judgment of causality. *Quarterly Journal of Experimental Psychology*, 41B, 139-159.
- Shugan, S.M. (1980). The cost of thinking. *Journal of Consumer Research*, 7, 99-111.
- Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development*, 47, 1-51.
- Simkin, D., & Hastie, R. (1987). An information-processing analysis of graph perception. *Journal of the American Statistical Association*, 82, 454-465.
- Skirbekk, G. & Gilje, N. (1993). *Geschichte der Philosophie*. Bde 1 und 2. Frankfurt am Main: Suhrkamp Taschenbuch Verlag.
- Smedslund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology*, 4, 165-173.
- Smith, M.C., & Magee, L.E. (1980). Tracing the time course of picture-word processing. *Journal of Experimental Psychology: General*, 109, 373-392.



- Sodian, B. (2002). Entwicklung begrifflichen Wissens. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie*, fünfte vollständig überarbeitete Auflage. Psychologie Verlags Union: Weinheim.
- Spelke, E.S., Phillips, A., & Woodward, A.L. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack & A.J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 44-78). New York, NY: Clarendon.
- Spellmann, B.A. (1996). Conditionalizing causality. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol. 34: Causal learning* (pp. 167-206). San Diego: Academic Press.
- Spellman, B. A., Price, C. M., & Logan, J. M. (2001). How two causes are different from one: The use of (un)conditional information in Simpson's paradox. *Memory & Cognition*, 29, 193-208.
- Spence, I. (1990). Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 683-692.
- Spence, I., & Lewandowsky, S. (1991). Displaying proportions and percentages. *Applied Cognitive Psychology*, 5, 61-77.
- Sperber, D., Premack, D. & Premack, A. (1995). *Causal cognition: A multidisciplinary debate*. New York, NY: Clarendon.
- Spirtes, P., Glymour, C., & Scheines, R. (1993/2001). *Causation, prediction, and search*. New York: Springer-Verlag.
- Stegmüller, W. (1983). *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie. Bd. 1 Erklärung, Begründung, Kausalität. 2., verbesserte und erweiterte Aufl.. Teil D*. Berlin Heidelberg New York: Springer-Verlag.
- Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, 19, 97-140.
- Steyvers, M., Tenenbaum, J. B., Wagenmakers, E-J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, 27, 453-489.
- Stone, D.N. & Schkade, D.A. (1991). Numeric and linguistic information representation in multiattribute choice. *Organizational Behavior and Human Decision Processes*, 49, 42-59.

- Strube, G., & Schlieder, C. (1996). Wissen und Wissensrepräsentation. In G. Strube et al. (Hrsg.), *Wörterbuch der Kognitionswissenschaft* (S. 799-815). Stuttgart: Klett-Cotta.
- Suppes, P. (1970). *A probabilistic theory of causality*. Amsterdam: North Holland.
- Suppes, P. (1984). *Probabilistic metaphysics*. Oxford: Basil Blackwell.
- Süß, H.-M. (1996). *Intelligenz, Wissen und Problemlösen. Kognitive Voraussetzungen für erfolgreiches Handeln bei computersimulierten Problemen*. Göttingen: Hogrefe.
- Süß, H.-M. (1999). Intelligenz und komplexes Problemlösen: Perspektiven für eine Kooperation zwischen differentiell-psychometrischer und kognitionspsychologischer Forschung. *Psychologische Rundschau*, 50(4), 220-228.
- Suthers, D.D. (2001). Towards a systematic study of representational guidance for collaborative learning discourse. *Journal of Universal Computer Science*, 7 (3), Electronic publication: [http://www.jucs.org/jucs\\_7\\_3/towards\\_a\\_systematic\\_study](http://www.jucs.org/jucs_7_3/towards_a_systematic_study).
- Tan, J. K. H., & Benbasat, I. (1990). Processing of graphical information: A decomposition taxonomy to match data extraction tasks and graphical representation. *Information Systems Research*, 1, 416–439.
- Tangen, J. M., & Allan, L. G. (2004). Cue-interaction and judgments of causality: Contributions of causal and associative processes. *Memory & Cognition*, 32 (1), 107-124.
- Tenenbaum, J. B., & Griffiths, T. L. (2001). Structure learning in human causal induction. In T. Leen, T. Dietterich, & V. Tresp (Eds.), *Advances in neural information processing systems* (Vol. 13, pp. 59–65). Cambridge, MA: MIT Press.
- Tenenbaum, J. B., & Griffiths, T. L. (2003). Theory-based causal inference. In S. Becker, S. Thrun, & K. Obermayer (Eds.), *Advances in neural information processing systems* (Vol. 15, pp. 35-42). Cambridge, MA: MIT Press.
- Tufte, E.R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.
- Tufte, E.R. (1990). *Envisioning information*. Cheshire, CT: Graphics Press.
- Tukey, J.W. (1990). Data based graphics: Visual displays in the decades to come. *Statistical Science*, 5, 327-329.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.

- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1131.
- Tversky, B. (2001). Spatial schemas in depictions. In M. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 79-111). Cambridge: MIT Press. Oder:
- Vallée-Tourangeau, F., Hollingsworth, L., & Murphy, R. A. (1998). "Attentional bias" in correlation judgments: Smedslund (1963) revisited. *Scandinavian Journal of Psychology*, *39*, 221-233.
- Vallée-Tourangeau, F., Murphy, R. A., & Drew, S. (1997). Causal judgements that violate the predictions of the power PC theory of causal induction. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the nineteenth annual conference of the Cognitive Science Society* (pp. 775-780). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Vallée-Tourangeau, F., Murphy, R. A., Drew, S., & Baker, A. G. (1998). Judging the importance of constant and variable candidate causes: A test of the Power PC Theory. *Quarterly Journal of Experimental Psychology*, *51A*, 65-84.
- Van der Meer, E. (2003). Verstehen von Kausalzusammenhängen. In G. Rickheit (Hrsg.), *Psycholinguistik: ein internationales Handbuch, Abschnitt IV Sprachrezeption*, Kapitel 48 (S. 631-643). Berlin: de Gruyter.
- Van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Van Hamme, L. J., Kao, S. F., & Wassermann, E. A. (1993). Judging interevent relations: From cause to effect and from effect to cause. *Memory and Cognition*, *21*, 802-808.
- Van Hamme L.J., & Wasserman, E.A. (1993). Cue competition in causality judgments: The role of manner in information presentation. *Bulletin of the Psychonomic Society*, *31*(5), 457-460.
- Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation*, *25*, 127-151.
- Van Someren, M.W., Reimann, P., Boshuizen, H.P.A., & de Jong, T. (1998). *Learning with multiple representations*. Oxford: Elsevier Science.
- Vandorpe, S., De Houwer, J., & Beckers, T. (2005). Further evidence for the role of inferential reasoning in forward blocking. *Memory & Cognition*, *33*(6), 1047-1056.

- Vessey, I. (1991). Cognitive fit: a theory-based analysis of the graphs versus tables literature. *Decision Sciences*, 22, 219 - 240.
- Vessey, I. (1994). The effect of information presentation on decision making: a cost-benefit analysis. *Information & Management*, 27, 103 - 119.
- Waldmann, M. R. (1994). *Der Erwerb von Kausalwissen: Prädiktives und diagnostisches Lernen mit Kausalmodellen*. Unveröffentlichte Habilitationsschrift, Tübingen.
- Waldmann, M. R. (1996). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol. 34: Causal learning* (pp. 47-88). San Diego: Academic Press.
- Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 53-76.
- Waldmann, M. R. (2001). Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychological Bulletin & Review*, 8, 600-608.
- Waldmann, M. R., & Hagmayer, Y. (2001). Estimating causal strength: The role of structural knowledge and processing effort. *Cognition*, 82, 27-58.
- Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General*, 121, 222-236.
- Waldmann, M. R., & Holyoak, K. J. (1997). Determining whether causal order affects cue selection in human contingency learning: Comments on Shanks and Lopez (1996). *Memory & Cognition*, 25, 125-134.
- Waldmann, M. R., Holyoak, K. J., & Fratianne, A. (1995). Causal models and the acquisition of category structure. *Journal of Experimental Psychology: General*, 124, 181-206.
- Waldmann, M. R., & Martignon, L. (1998). A Bayesian network of causal learning. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 1102-1107). Mahwah, NJ: Erlbaum.
- Ward, W. C., & Jenkins, H. M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology*, 19, 231-241.

- Washburne, J.N. (1927). An experimental study of various graphic, tabular and textual methods of presenting quantitative material. *Journal of Educational Psychology*, 18, 361 - 376, 465 - 476.
- Wasserman, E. A. (1990). Detecting response-outcome relations: Toward an understanding of the causal texture of the environment. In G. H. Bower (Ed.), *The psychology of learning and motivation*, Vol. 26. New York: Academic Press.
- Wasserman, E. A., & Berglan, L. R. (1998). Backward blocking and recovery from overshadowing in human causal judgement: The role of within-compound associations. *Quarterly Journal of Experimental Psychology*, 51B, 121-138.
- Wasserman, E. A., Chatlosh, D. L., & Neunaber, D. J. (1983). Perception of causal relations in humans: Factors affecting judgments of response-outcome contingencies under free operant procedures. *Learning and motivation*, 14, 406-432.
- Wasserman, E. A., Dorner, W. W., & Kao, S. F. (1990). Contributions of specific cell information to judgments of interevent contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 509-521.
- Wasserman, E. A., Elek, S. M., Chatlosh, D. L., & Baker, A. G. (1993). Rating causal relations: Role of probability in judgments of response-outcome contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 174-188.
- Wasserman, E. A., Kao, S.-F., Van Hamme, L. J., Katagiri, M., & Young, M. E. (1996). Causation and association. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation*, Vol. 34: *Causal learning* (pp. 47-88). San Diego: Academic Press.
- Wasserman, E.A. & Neunaber, D. J. (1986). College student's responding to and rating of contingency relations: The role of temporal contiguity. *Journal of the Experimental Analysis of Behavior*, 46, 15-35.
- Wasserman, E.A., & Shaklee, H. (1984). Judging response-outcome relations: The role of response-outcome contingency, outcome probability, and method of information presentation. *Memory and Cognition*, 12, 270-286.
- Weidenmann, B. (2006). *Lernen mit Medien*. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie: Ein Lehrbuch*. 5., vollständig überarbeitete Auflage (S. 423-476). Weinheim: Beltz Psychologie Verlags Union.

- White, P.A. (1989). A theory of causal processing. *British Journal of Psychology*, 80, 431-454.
- White, P.A. (1993). *Psychological Metaphysics*. London: Routledge.
- White, P.A. (1995). Use of prior beliefs in the assignment of causal roles: Causal powers versus regularity-based accounts. *Memory and Cognition*, 23, 243-254.
- White, P.A. (1998). Causal judgement: Use of different types of contingency information as confirmatory and disconfirmatory. *European Journal of Cognitive Psychology*, 10, 131-170.
- White, P.A. (2000a). Causal judgement from contingency information: Relation between subjective reports and individual tendencies in judgement. *Memory and Cognition*, 28, 415-426.
- White, P.A. (2000b). Causal judgement from contingency information: The interpretation of factors common to all instances. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1083-1102.
- White, P.A. (2002a). Causal attribution from covariation information: The evidential evaluation model. *European Journal of Social Psychology*, 32, 667-684.
- White, P.A. (2002b). Causal judgement from contingency information: Judging interactions between two causal candidates. *Quarterly Journal of Experimental Psychology*, 55A, 819-838.
- White, P.A. (2002c). Perceiving a strong causal relation in a weak contingency: Further investigation of the evidential evaluation model of causal judgement. *Quarterly Journal of Experimental Psychology*, 55A, 97-114.
- White, P.A. (2003a). Causal judgement as the evaluation of evidence: The use of confirmatory and disconfirmatory information. *Quarterly Journal of Experimental Psychology*, 56A, 491-513.
- White, P.A. (2003b). Effects of wording and stimulus format on the use of contingency information in causal judgement. *Memory and Cognition*, 31, 231-242.
- White, P.A. (2003c). Making causal judgments from the proportion of confirming instances: The pCI rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 710-727.

- White, P.A. (2004). Causal judgement from contingency information: A systematic test of the pCI rule. *Memory and Cognition*, 32, 353–368.
- White, P.A. (2005a). Cue interaction effects in causal judgement: An interpretation in terms of the evidential evaluation model. *The Quarterly Journal of Experimental Psychology*, 58B (2), 99-140.
- White, P.A. (2005b). The Power PC Theory and causal powers: Comment on Cheng (1997) and Novick and Cheng (2004). *Psychological Review*, Vol. 112, No. 3, 675-684.
- Wickens, C.D., & Andre, A.D. (1990). Proximity compatibility and information display: effects of colour, space and objectness on information integration. *Human Factors*, 32, 61-78.
- Wickens, C.D., & Carswell, C.M. (1995). The proximity compatibility principle: its psychological foundation and its relevance to display design. *Human Factors*, 37, 473-494.
- Wilkinson, L. (1999). Graphs for research in counseling psychology. *Counseling Psychology*, 27, 384-407.
- Williams, D.A. (1996). A comparative analysis of negative contingency learning in humans and nonhumans. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation*, Vol. 34: Causal learning (pp. 89-131). San Diego, CA: Academic Press.
- Wilson, E.V., & Addo, T.B.A. (1994a). The joint effects of interactions between data display and task variables on task performance. In J.I. DeGross, S.L. Huff, & M.C. Munro (Eds.), *Proceedings of the 15th International Conference on Information Systems* (pp. 241-353). Vancouver, B.C.
- Wilson, E.V., & Addo, T.B.A. (1994b). An investigation of the relative presentation efficiency of computer-displayed graphs. *Information Management* 26, 105-115.
- Wilson, R. A., & Keil, F. C. (1998). The shadows and shallows of explanation. *Minds and Machines*, 8, 137-159.
- Winn, W.D. (1987). Charts, graphs, and diagrams in educational materials. In D. M. Willows & H. A. Houghton (Eds.), *The psychology of illustration*. Vol. 1 (pp. 152-198). New York: Springer.

- Winn, W.D. (1994). Contributions of perceptual and cognitive processes to the comprehension of graphics. In W. Schnotz & R.W. Kulhavy (Eds.), *Comprehension of graphics* (pp. 3-27). Amsterdam: North-Holland.
- Woodward, J. (2003). *Making things happen. A theory of causal explanation*. Oxford: Oxford University Press.
- Wu, M., & Cheng, P.W. (1999). Why causation need not follow from statistical association: Boundary conditions for the evaluation of generative and preventive causal powers. *Psychological Science, 10*, 92-97.
- Yates, J.F., & Curley, S.P. (1986). Contingency judgment: Primacy effects and attention decrement. *Acta Psychologica, 62*, 293-302.
- Young, M.E. (1995). On the origin of personal causal theories. *Psychonomic Bulletin and Review, 2 (1)*, 83-104.
- Zacks, J. & Tversky, B. (1999). Bars and lines: A study of graphic communication. *Memory & Cognition, 27 (6)*, 1073-1079.
- Zelazny, G. (1986). *Wie aus Zahlen Bilder werden. Wirtschaftsdaten überzeugend präsentiert*. Wiesbaden: Gabler.
- Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science, 21*, 179–217.
- Zhang, J. (2000). External representations in complex information processing tasks. In A. Kent (Ed.), *Encyclopedia of library and information science*. New York: Marcel Dekker, Inc.
- Zhang, J., Johnson, K.A., Malin, J.T., & Smith, J.W. (2002). Human-Centered Information Visualization. Paper presented at the *International Workshop on Dynamic Visualization and Learning in Tübingen, Germany, Knowledge Media Research Center, July 18-19, 2002*.
- Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science, 18*, 87–122.



## Abbildungsverzeichnis

<b>Abbildung 1:</b> Kombinationen in einer 2 (Wirkung tritt ein oder tritt nicht ein) x 2 (Ursache ist gegeben oder nicht gegeben) Kontingenztabelle. Die Kleinbuchstaben a, b, c und d stehen für Ereignishäufigkeiten.	Seite	24
<b>Abbildung 2:</b> Bildung von Kausalurteilen in erlebten Situationen (nach Shanks, 1991).	Seite	66
<b>Abbildung 3:</b> Bildung von Kausalurteilen in beschriebenen Situationen (nach Shanks, 1991).	Seite	67
<b>Abbildung 4:</b> Schwarz-Weiß-Version des Stimulus-Materials von Buehner et al. (2003, Ex 2).	Seite	79
<b>Abbildung 5:</b> Zusammenspiel von Kognition, sowie medien-, aufgaben- und personenbezogenen Faktoren bei der Beurteilung der Stärke von Kausalzusammenhängen.	Seite	132
<b>Abbildung 6:</b> Experiment 1: 2 x 2-Design.	Seite	143
<b>Abbildung 7:</b> Experiment 1: Kontrollbedingung Protokollliste.	Seite	144
<b>Abbildung 8:</b> Experiment 1: Wirksamkeitsskala.	Seite	160
<b>Abbildung 9:</b> Experiment 1: Mittlere Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) in Abhängigkeit von <i>Präsentationsform</i> und <i>Informationsart</i> .	Seite	162
<b>Abbildung 10:</b> Experiment 1: Mittlere Bearbeitungszeiten in Minuten : Sekunden (in Klammern Standardabweichungen) in Abhängigkeit von <i>Präsentationsform</i> und <i>Inhaltsart</i> .	Seite	164
<b>Abbildung 11:</b> Experiment 1: Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von den beiden Faktoren <i>Präsentationsform</i> und <i>Inhaltsart</i> .	Seite	165
<b>Abbildung 12:</b> Experiment 1: Verteilung von Probanden, die regel-basierte oder keine regel-basierten Strategien eingesetzt haben, in Abhängigkeit von <i>Präsentationsform</i> (Stapeldiagramm vs. Tabelle) und <i>Inhaltsart</i> (Wahrscheinlichkeiten vs. Häufigkeiten).	Seite	176
<b>Abbildung 13:</b> Experiment 1: Verteilung von Probanden, die entweder durchgängig oder nicht durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von <i>Inhaltsart</i> (Wahrscheinlichkeiten vs. Häufigkeiten) und <i>Präsentationsform</i> (Stapeldiagramm vs. Tabelle).	Seite	177
<b>Abbildung 14:</b> Experiment 1: Anzahl Probanden, die durchgängig oder teilweise regel-basierte Strategien eingesetzt haben und von Probanden, die einfache Heuristiken oder unklare Strategien eingesetzt haben, in Abhängigkeit von der Versuchsbedingung.	Seite	178

- Abbildung 15:** Experiment 1: Anzahl Probanden, die durchgängig oder nicht durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von der Versuchsbedingung. Seite 178
- Abbildung 16:** Experiment 1: Prozentanteile von Probanden, die entweder Häufigkeiten oder Wahrscheinlichkeiten dargeboten bekommen haben, an der Gesamtzahl der Nutzer von entweder der  $\Delta P$ -Regel oder der Power PC Theorie. Seite 180
- Abbildung 17:** Experiment 1: Mittlere Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) in Abhängigkeit von *Studienausrichtung* und *Versuchsbedingung*. Seite 186
- Abbildung 18:** Experiment 1: Häufigkeit des durchgängigen oder zumindest teilweisen Einsatzes regel-basierter Strategien in Abhängigkeit von der *Studienausrichtung* ("NW" = mathematisch-naturwissenschaftlich"; "GW" = geisteswissenschaftlich bzw. nicht mathematisch-naturwissenschaftlich). Seite 187
- Abbildung 19:** Experiment 1: Grafische Veranschaulichung der unterschiedlichen Haupteffekte auf die *Anzahl regel-konformer Lösungen* für die Gesamtstichprobe und die beiden Teilstichproben (Leibniz-Kollegiaten und Studenten). In der angehängten Tabelle sind signifikante Unterschiede **fett** gedruckt, Trends wurden unterstrichen. Seite 191
- Abbildung 20:** Experiment 2: Hohe visuelle Strukturierung des darstellenden Mediums und summarische Darbietung: Stapeldiagramm mit Häufigkeiten (kurz *SD sum*). Seite 206
- Abbildung 21:** Experiment 2: Hohe visuelle Strukturierung des darstellenden Mediums und Einzelfalldarbietung: Vierfeldertafel mit Positionskreuz (kurz *VT tbt*). Seite 206
- Abbildung 22:** Experiment 2: Geringe visuelle Strukturierung und summarische Darbietung: Fliesstext mit Häufigkeitsangaben (kurz *FT sum*). Seite 207
- Abbildung 23:** Experiment 2: Geringe visuelle Strukturierung und Einzelfalldarbietung: Fliesstext mit Aussagen zu Einzelereignissen (kurz *FT tbt*). Seite 207
- Abbildung 24:** Experiment 2: Ratingskala zur Wirksamkeitsbeurteilung. Seite 219
- Abbildung 25:** Experiment 2: Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*. Seite 222
- Abbildung 26:** Experiment 2: Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*. Seite 224
- Abbildung 27:** Experiment 2: Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen in Abhängigkeit von der Versuchsbedingung. Seite 225

- Abbildung 28:** Experiment 2: Mittlere Bearbeitungszeit in Minuten : Sekunden (in Klammern Standardabweichungen) in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*. Seite 227
- Abbildung 29:** Experiment 2: Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*. Seite 230
- Abbildung 30:** Experiment 2: Verteilung von Probanden, die regel-basierte oder keine regel-basierten Strategien eingesetzt haben, in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*. Seite 239
- Abbildung 31:** Experiment 2: Verteilung von Probanden, die entweder durchgängig regel-basierte oder nicht durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*. Seite 239
- Abbildung 32:** Experiment 2: Verteilung von Probanden, die durchgängig oder teilweise regel-basierte Strategien eingesetzt haben und von Probanden, die einfache Heuristiken oder unklare Strategien eingesetzt haben in Abhängigkeit von der Versuchsbedingung. Seite 240
- Abbildung 33:** Experiment 2: Verteilung von Probanden, die entweder durchgängig regel-basierte oder nicht durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von der Versuchsbedingung. Seite 241
- Abbildung 34:** Experiment 2: Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) getrennt nach *Studienausrichtung* in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*. Seite 247
- Abbildung 35:** Experiment 2: Anzahl von Probanden, die durchgängig oder zumindest teilweise regel-basierte Strategien in Abhängigkeit von der *Studienausrichtung* ("NW" = mathematisch-naturwissenschaftlich"; "GW" = geisteswissenschaftlich bzw. nicht mathematisch-naturwissenschaftlich) eingesetzt haben. Seite 251
- Abbildung 36:** Experiment 2: Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) getrennt nach Intelligenzunterschieden im *numerischen Denken* (*ND hoch* vs. *ND gering*) und in Abhängigkeit von den beiden Faktoren *Darbietungsart* und *visuelle Strukturierung*. Seite 255
- Abbildung 37:** Experiment 3: Leere Vierfeldertafel, wie sie die Probanden zum Anfertigen der Strichlisten unter der Bedingung *N strukturiert* vorgelegt bekamen. Seite 269
- Abbildung 38:** Experiment 3: Vorhergesagter Einfluss des Faktors *Notizen anfertigen* auf die abhängige Variable *Anzahl regel-konformer Lösungen*: Seite 272

- Abbildung 39:** Experiment 3: Mittlere Anzahl regel-konformer Lösungen (in Klammern Standardabweichungen) in Abhängigkeit von *Notizen anfertigen*. Seite 277
- Abbildung 40:** Experiment 3: Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen in Abhängigkeit von *Notizen anfertigen*. Seite 279
- Abbildung 41:** Experiment 3: *Anzahl regel-konformer Lösungen / Bearbeitungszeit* in Abhängigkeit von *Notizen anfertigen* (in Klammern Standardabweichungen). Seite 289
- Abbildung 42:** Experiment 3: *Anzahl regel-konformer Lösungen / Bearbeitungszeit* in Abhängigkeit von *Notizen anfertigen* (in Klammern Standardabweichungen). Seite 293
- Abbildung 43:** Experiment 3: Verteilung von Probanden nach Strategietyp (durchgängig oder teilweise regel-basiert versus nicht regel-basiert) in Abhängigkeit von *Notizen anfertigen*. Seite 297
- Abbildung 44:** Experiment 3: Verteilung von Probanden nach Strategietyp (teilweise oder gar nicht regel-basiert vs. durchgängig regel-basiert) in Abhängigkeit von *Notizen anfertigen*. Seite 298
- Abbildung 45:** Experiment 3: Maße der zentralen Tendenz für die abhängige Variable *Anzahl regel-konformer Lösungen* (in Klammern Standardabweichungen) getrennt nach Studienausrichtung in Abhängigkeit von *Notizen anfertigen*. Seite 303
- Abbildung 46:** Experiment 3: Anzahl Probanden, die durchgängig oder zumindest teilweise regel-basierte Strategien in Abhängigkeit von *Studienausrichtung* eingesetzt haben. Seite 307
- Abbildung 47:** Experiment 3: Anzahl Probanden mit über- oder unterdurchschnittlichen Werten im numerischen Denken in Abhängigkeit von *Studienausrichtung* und *Notizen anfertigen*. Seite 308
- Abbildung 48:** Experiment 3: Maße der zentralen Tendenz für *Anzahl regel-konformer Lösungen* bei Probanden mit hohen und geringen Werten im *numerischen Denken* (ND hoch vs. ND gering) in Abhängigkeit von *Notizen anfertigen*. Seite 311

## Tabellenverzeichnis

<b>Tabelle 1:</b>	Experiment 1: Anzahl der Probanden in Abhängigkeit von Strategietyp und Versuchsbedingung (in Klammern Prozentangaben). Zelhäufigkeiten > 5 sind <b>fett</b> markiert.	Seite	175
<b>Tabelle 2:</b>	Experiment 1: Anzahl der Probanden, die durchgängig oder teilweise regel-basierte Strategien eingesetzt haben, in Abhängigkeit von der Art der regel-basierten Strategie und der Versuchsbedingung (in Klammern Prozentangaben).	Seite	179
<b>Tabelle 3:</b>	Experiment 1: Mittlere Bearbeitungszeiten (in Minuten : Sekunden) und mittlere Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von <i>Versuchsbedingung</i> und <i>Studienausrichtung</i> ("NW" = mathematisch-naturwissenschaftlich ausgerichtete Studienfächer; "GW" = geisteswissenschaftlich ausgerichtete Studienfächer	Seite	187
<b>Tabelle 4:</b>	Experiment 1: Mittelwerte (in Klammern Standardabweichungen) für die abhängigen Variablen <i>Anzahl regel-konformer Lösungen</i> , <i>Bearbeitungszeit</i> und <i>Verarbeitungseffizienz</i> in Abhängigkeit von der jeweiligen Teilstichprobe (Leibniz-Kollegiaten vs. immatrikulierte Studenten) und der Versuchsbedingung.	Seite	189
<b>Tabelle 5:</b>	Experiment 1: Ergebnisse aus den 2 ( <i>Präsentationsform</i> ) x 2 ( <i>Informationsart</i> ) Varianzanalysen für die beiden Probandengruppen (Leibniz-Kollegiaten und Studenten) mit der abhängigen Variable <i>Anzahl regel-konformer Lösungen</i> . Signifikante P-Werte sind <b>fett</b> gedruckt, Trends wurden <u>unterstrichen</u> .	Seite	190
<b>Tabelle 6:</b>	Experiment 1: Rangkorrelationen (Kendall-Tau-b) zwischen den Werten für die abhängigen Variablen und den Werten für die Items aus der Nachbefragung. In Klammern sind die p-Werte angegeben, signifikante Korrelationen mit einem p-Wert < .05 sind <b>fett</b> gedruckt; Korrelationen mit p-Werten < .10 (Trend) wurden <u>unterstrichen</u> .	Seite	194
<b>Tabelle 7:</b>	Experiment 1: Partialkorrelationen zwischen den Werten für die abhängigen Variablen und den Werten für die Items aus der Nachbefragung, bei denen der Einfluss der vier Experimentalbedingungen und der Kontrollbedingung kontrolliert wurde. In Klammern sind die Anzahl der Freiheitsgrade und die p-Werte angegeben. Signifikante Korrelationen mit einem p-Wert < .05 sind <b>fett</b> gedruckt; Korrelationen mit p-Werten < .10 (Trend) <u>unterstrichen</u> und Korrelationen mit p-Werten < .15 (Trendnähe) <i>kursiv</i> gedruckt.	Seite	195
<b>Tabelle 8:</b>	Experiment 2: 2 x 2 – Design.	Seite	208

<b>Tabelle 9:</b>	Experiment 2: Ergebnisse aus den 2 (Darbietungsart) x 2 (visuelle Strukturierung) Kovarianzanalysen für die abhängige Variable <i>Anzahl regel-konformer Lösungen</i> . Signifikante P-Werte sind <b>fett</b> gedruckt, Trends wurden <u>unterstrichen</u> .	Seite 223
<b>Tabelle 10:</b>	Experiment 2: Ergebnisse aus Häufigkeitsvergleichen, in denen untersucht wurde, ob die Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen von je zwei Versuchsbedingungen unabhängig ist oder nicht. Signifikante P-Werte ( $\alpha < .05$ ) sind <b>fett</b> markiert.	Seite 226
<b>Tabelle 11:</b>	Experiment 2: Ergebnisse aus den 2 x 2 Kovarianzanalysen, in denen der Einfluss der Faktoren <i>Darbietungsart</i> und <i>visuelle Strukturierung</i> , sowie der Kovariate <i>numerisches Denken</i> auf die drei Zeitmaße <i>total time</i> , <i>time on task</i> und <i>rating time</i> untersucht wurde. Signifikante P-Werte sind <b>fett</b> gedruckt.	Seite 228
<b>Tabelle 12:</b>	Experiment 2: Ergebnisse aus den 2 (Darbietungsart) x 2 (visuelle Strukturierung) Kovarianzanalysen für das Effizienzmaß <i>Anzahl regel-konformer Lösungen / Bearbeitungszeit</i> . Signifikante P-Werte sind <b>fett</b> gedruckt, Trends wurden <u>unterstrichen</u> .	Seite 231
<b>Tabelle 13:</b>	Experiment 2: Anzahl der Probanden in Abhängigkeit von Strategietyp und Versuchsbedingung (in Klammern Prozentangaben). Zellhäufigkeiten > 5 sind <b>fett</b> markiert.	Seite 237
<b>Tabelle 14:</b>	Experiment 2: Anzahl der Probanden, die durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von der Art der regel-basierten Strategie und der Versuchsbedingung (in Klammern Prozentangaben).	Seite 242
<b>Tabelle 15:</b>	Experiment 2: Ergebnisse für die abhängige Variable <i>Anzahl regel-konformer Lösungen</i> aus der dreifaktoriellen Varianzanalyse mit dem zusätzlichen Faktor <i>Studienausrichtung</i> über alle Probanden hinweg und den 2 ( <i>Darbietungsart</i> ) x 2 ( <i>visuelle Strukturierung</i> ) Varianzanalysen für die beiden nach <i>Studienausrichtung</i> getrennten Probandengruppen. Signifikante P-Werte sind <b>fett</b> gedruckt, Trends wurden <u>unterstrichen</u> .	Seite 248
<b>Tabelle 16:</b>	Experiment 2: Bearbeitungszeit in Minuten : Sekunden und Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von <i>Studienausrichtung</i> ("NW" = mathematisch-naturwissenschaftliche Ausrichtung; "GW" = geisteswissenschaftliche Ausrichtung) und den vier Experimentalbedingungen <i>SD sum</i> , <i>FT sum</i> , <i>VT tbt</i> und <i>FT tbt</i> .	Seite 249
<b>Tabelle 17:</b>	Experiment 2: Ergebnisse aus den 2 ( <i>Darbietungsart</i> ) x 2 ( <i>visuelle Strukturierung</i> ) Varianzanalysen für die beiden Teilstichproben ( <i>ND hoch</i> und <i>ND gering</i> ) mit der abhängigen Variable <i>Anzahl regel-konformer Lösungen</i> . Signifikante P-Werte sind <b>fett</b> gedruckt, Trends wurden <u>unterstrichen</u> .	Seite 256
<b>Tabelle 18:</b>	Experiment 3: 1 x 4 – Design.	Seite 269

- Tabelle 19:** Experiment 3: Ergebnisse aus der univariaten Kovarianzanalyse mit dem Faktor *Notizen anfertigen* für die abhängige Variable *Anzahl regel-konformer Lösungen*. Signifikante P-Werte sind **fett** gedruckt. Seite 278
- Tabelle 20:** Experiment 3: Paarweise Vergleiche zwischen den Bonferroni-adjustierten Gruppenmittelwerten für die abhängige Variable *Anzahl regel-konformer Lösungen* in Abhängigkeit von *Notizen anfertigen*. Signifikante P-Werte ( $\alpha < .05$ ) sind **fett** markiert. Seite 278
- Tabelle 21:** Experiment 3: Ergebnisse aus paarweisen Häufigkeitsvergleichen, in denen untersucht wurde, ob die Verteilung der Anzahl von Probanden mit mehr als drei oder weniger als vier regel-konformen Lösungen von je zwei Versuchsbedingungen unabhängig ist oder nicht. Signifikante P-Werte sind **fett** markiert, Trends unterstrichen. Seite 280
- Tabelle 22:** Experiment 3: Ergebnisse aus den univariaten Kovarianzanalysen, in denen der Einfluss von *Notizen anfertigen* und *numerisches Denken* auf die drei Zeitmaße *total time*, *time on task* und *rating time* untersucht wurde. Signifikante P-Werte sind **fett** gedruckt. Seite 289
- Tabelle 23:** Experiment 3: Paarweise Vergleiche zwischen den Bonferroni-adjustierten Gruppenmittelwerten für die Variablen *total time*, *time on task* und *rating time*. Signifikante P-Werte sind **fett** gedruckt. Seite 290
- Tabelle 24:** Experiment 3: Ergebnisse aus univariaten Kovarianzanalysen, in denen der Einfluss von *Notizen anfertigen* und *numerischem Denken* auf die Verarbeitungseffizienz untersucht wurde. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen. Seite 294
- Tabelle 25:** Experiment 3: Paarweise Vergleiche zwischen den Bonferroni-adjustierten Gruppenmittelwerten für die abhängige Variable *Verarbeitungseffizienz* in Abhängigkeit vom Faktor *Notizen anfertigen*. Signifikante P-Werte ( $\alpha < .05$ ) sind **fett** markiert. Seite 294
- Tabelle 26:** Experiment 3: Anzahl der Probanden in Abhängigkeit von Strategietyp und Versuchsbedingung (in Klammern Prozentangaben). Zellhäufigkeiten > 5 sind **fett** markiert. Seite 296
- Tabelle 27:** Experiment 3: Ergebnisse von Häufigkeitsanalysen zur Verteilung der Strategiewahl (durchgängig oder teilweise regelbasiert versus nicht regelbasiert) in Abhängigkeit von den Bedingungsvergleichen *ohne N sum* vs. *ohne N tbt*; *N strukturiert* vs. *ohne N tbt* und *N frei* vs. *ohne N tbt*. Signifikante P-Werte ( $\alpha < .05$ ) sind **fett** markiert. Seite 298
- Tabelle 28:** Experiment 3: Ergebnisse von Häufigkeitsanalysen zur Verteilung der Strategiewahl (durchgängig versus nicht durchgängig regelbasiert) in Abhängigkeit von den drei Bedingungsvergleichen *ohne N sum* vs. *ohne N tbt*; *N strukturiert* vs. *ohne N tbt* und *N frei* vs. *ohne N tbt*. Signifikante P-Werte ( $\alpha < .05$ ) sind **fett** markiert. Seite 299

- Tabelle 29:** Experiment 3: Anzahl Probanden, die durchgängig regel-basierte Strategien eingesetzt haben, in Abhängigkeit von der Art der regel-basierten Strategie und der Versuchsbedingung (in Klammern Prozentangaben). Seite 299
- Tabelle 30:** Experiment 3: Ergebnisse für die abhängige Variable *Anzahl regel-konformer Lösungen* aus der zweifaktoriellen Varianzanalyse mit den beiden Faktoren *Studienausrichtung* und *Notizen anfertigen*, sowie den beiden einfaktoriellen Varianzanalysen mit dem Faktor *Notizen anfertigen* für die beiden nach *Studienausrichtung* getrennten Probandengruppen. Signifikante P-Werte sind **fett** gedruckt. Seite 304
- Tabelle 31:** Experiment 3: Einzelkontraste (Least significant difference, Tukey HSD; Scheffé-Test) in der Teilstichprobe Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung für die abhängige Variable *Anzahl regel-konformer Lösungen* in Abhängigkeit von *Notizen anfertigen*. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen. Seite 305
- Tabelle 32:** Experiment 3: Einzelkontraste (Student-Newman-Keuls-Test) in der Teilstichprobe Probanden mit mathematisch-naturwissenschaftlicher Studienausrichtung für die abhängige Variable *Anzahl regel-konformer Lösungen* in Abhängigkeit von *Notizen anfertigen*. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen. Seite 305
- Tabelle 33:** Experiment 3: Bearbeitungszeit und Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von *Studienausrichtung* und *Notizen anfertigen*. Seite 306
- Tabelle 34:** Experiment 3: Ergebnisse für die abhängige Variable *Anzahl regel-konformer Lösungen* aus der 2 (*numerisches Denken*) x 4 (*Notizen anfertigen*) Varianzanalyse, sowie den beiden einfaktoriellen Varianzanalysen für die Probandengruppen *ND hoch* und *ND gering* mit dem Faktor *Notizen anfertigen*. Signifikante P-Werte sind **fett** gedruckt, Trends wurden unterstrichen. Seite 312
- Tabelle 35:** Experiment 3: Bearbeitungszeit und Verarbeitungseffizienz (in Klammern Standardabweichungen) in Abhängigkeit von *numerisches Denken* und *Notizen anfertigen*. Seite 313
- Tabelle A1:** Experiment 1: Zusammensetzung der Teilnehmer hinsichtlich ihrer realen oder beabsichtigten Studienfächer und sonstiger demografischer Merkmale. Seite A 3
- Tabelle A2:** Experiment 2: Zusammensetzung der Teilnehmer hinsichtlich ihrer realen oder beabsichtigten Studienfächer und sonstiger demografischer Merkmale. Seite A 14
- Tabelle A3:** Experiment 2: Anzahl der Probanden, die in Abhängigkeit von Strategietyp und Versuchsbedingung die Strategie im Laufe der Aufgabenbearbeitung gewechselt haben. Zelhäufigkeiten > 5 sind **fett** markiert. Seite A 22



- Tabelle A4:** Experiment 2: Signifikante Faktoren und Interaktionen für die abhängigen Variablen *Bearbeitungszeit* und *Verarbeitungseffizienz* aus der dreifaktoriellen Varianzanalyse mit dem zusätzlichen Faktor *Studienausrichtung* über alle Probanden hinweg und den 2 (*Darbietungsart*) x 2 (*Visuelle Strukturierung*) Varianzanalysen für die beiden Probandengruppen mit und ohne mathematisch-naturwissenschaftliche Studienausrichtung. Seite A 23
- Tabelle A5:** Experiment 3: Zusammensetzung der Teilnehmer hinsichtlich ihrer realen oder beabsichtigten Studienfächer und sonstiger demografischer Merkmale. Seite A 24
- Tabelle A6:** Experiment 3: Anzahl der Probanden, die in Abhängigkeit von Strategietyp und Versuchsbedingung die Strategie im Laufe der Aufgabenbearbeitung gewechselt haben. Seite A 30
- Tabelle A7:** Experiment 3: Signifikante Faktoren und Interaktionen für die abhängigen Variablen *Bearbeitungszeit* und *Verarbeitungseffizienz* aus der zweifaktoriellen Varianzanalyse mit den Faktoren *Studienausrichtung* und *Notizen anfertigen* über alle Probanden hinweg und den univariaten Varianzanalysen mit dem vierstufigen Faktor *Notizen anfertigen* für die beiden Probandengruppen mit und ohne mathematisch-naturwissenschaftliche Studienausrichtung. Seite A 31
- Tabelle A8:** Experiment 3: Signifikante Faktoren und Interaktionen für die abhängigen Variablen *Bearbeitungszeit* und *Verarbeitungseffizienz* aus der zweifaktoriellen Varianzanalyse mit den Faktoren *numerisches Denken (hoch vs. gering)* und *Notizen anfertigen* über alle Probanden hinweg und den univariaten Varianzanalysen mit dem vierstufigen Faktor *Notizen anfertigen* für die beiden Probandengruppen mit hohen oder geringen Werten im numerischen Denken. Seite A 32

# **Anhang**

## Anhangsverzeichnis

### A1 Anhang zu Experiment 1:

A1.1	Experiment 1: Zusammensetzung der Teilnehmer .....	A 3
A1.2	Experiment 1: Instruktionen .....	A 4
A1.3	Experiment 1: Fragen zur Vorgehensweise.....	A 9
A1.4	Experiment 1: Nachbefragung .....	A 10
A1.5	Experiment 1: Eingesetzte Präsentationsformen und Inhaltsarten .....	A 12

### A2 Anhang zu Experiment 2:

A2.1	Experiment 2: Zusammensetzung der Teilnehmer .....	A 14
A2.2	Experiment 2: Instruktionen .....	A 15
A2.3	Experimente 2 und 3: Fragen zur Vorgehensweise .....	A 21
A2.4	Experiment 2: Weitere Ergebnisse .....	A 22

### A3 Anhang zu Experiment 3:

A3.1	Experiment 3: Zusammensetzung der Teilnehmer .....	A 24
A3.2	Experiment 3: Instruktionen .....	A 25
A3.3	Experiment 3: Weitere Ergebnisse .....	A 30

### A4 Items zur Erfassung der Kovariate numerisches Denken

Die Items wurden dem Berliner Intelligenzstruktur-Test (BIS-Test, Jäger, Süß & Beauducel, 1997) entnommen und in der Nachuntersuchung zu den Experimenten 2 und 3 eingesetzt.

A4.0	Kurzbeschreibung der neun ausgewählten Items aus dem BIS-Test .....	A 32
A4.1	BIS-Item SI - Sieben-Teilbar (BN) .....	A 33
A4.2	BIS-Item RD - Rechnerisches Denken (KN) .....	A 35
A4.3	BIS-Item ZZ - Zweistellige Zahlen (MN) .....	A 37
A4.4	BIS-Item TL - Tabellen-Lesen (KN) .....	A 39
A4.5	BIS-Item ZN - Zahlenreihen (KN) .....	A 41
A4.6	BIS-Item XG - X-Größer (BN) .....	A 43
A4.7	BIS-Item SC - Schätzen (KN) .....	A 45
A4.8	BIS-Item RZ - Rechen-Zeichen (BN) .....	A 47
A4.9	BIS-Item BR - Buchstabenreihen (KN) .....	A 49

# Anhang A1

## Anhang A1.1: Experiment 1: Zusammensetzung der Teilnehmer

**Tabelle A1:** Experiment 1: Zusammensetzung der Teilnehmer hinsichtlich ihrer realen oder beabsichtigten Studienfächer und sonstiger demografischer Merkmale.

Vp	Studienrichtung	Studienfach/ Studienwunsch	Alter	Geschlecht	Präsentationsform	Inhaltsart	Vp	Studienrichtung	Studienfach/ Studienwunsch	Alter	Geschlecht	Präsentationsform	Inhaltsart
1	Geist	Pädagogik	28	männlich	Tabelle	Häufigkeiten	41	Natur	Medizin	19	weiblich	Tabelle	Häufigkeiten
2	Natur	Betriebswirtschaftslehre	28	Weiblich	Stapeldiagramm	Häufigkeiten	42	Geist	Medienwissenschaften/Jüdische Studien	20	weiblich	Stapeldiagramm	Häufigkeiten
3	Geist	Nordistik, Romanistik	22	Weiblich	Tabelle	Wahrscheinlichkeiten	43	Natur	Architektur	20	männlich	Tabelle	Wahrscheinlichkeiten
4	Geist	Jura	24	Weiblich	Stapeldiagramm	Wahrscheinlichkeiten	44	Geist	Germanistik	27	weiblich	Stapeldiagramm	Wahrscheinlichkeiten
5	Geist	Jura/Rhetorik	23	männlich	Protokollliste	Einzelereignisse	45	Natur	Geologie	20	weiblich	Protokollliste	Einzelereignisse
6	Geist	Philosophie	20	männlich	Tabelle	Häufigkeiten	46	Geist	Sportwissenschaft	21	männlich	Tabelle	Häufigkeiten
7	Natur	Medizin	28	männlich	Stapeldiagramm	Häufigkeiten	47	Natur	Volkswirtschaftslehre/Politik	21	männlich	Stapeldiagramm	Häufigkeiten
8	Geist	Jura	24	weiblich	Tabelle	Wahrscheinlichkeiten	48	Natur	Physik	20	weiblich	Tabelle	Wahrscheinlichkeiten
9	Natur	Medizin	26	weiblich	Stapeldiagramm	Wahrscheinlichkeiten	49	Geist	Philosophie/ Alte Geschichte	33	männlich	Stapeldiagramm	Wahrscheinlichkeiten
10	Geist	Allgemeine Sprachwissenschaft	36	männlich	Protokollliste	Einzelereignisse	50	Geist	Jura	25	weiblich	Protokollliste	Einzelereignisse
11	Natur	Umweltnaturwissenschaften	20	weiblich	Tabelle	Häufigkeiten	51	Natur	Medizin	22	weiblich	Tabelle	Häufigkeiten
12	Geist	Politik/Soziologie	20	weiblich	Stapeldiagramm	Häufigkeiten	52	Geist	Germanistik	22	weiblich	Stapeldiagramm	Häufigkeiten
13	Natur	Psychologie	20	weiblich	Tabelle	Wahrscheinlichkeiten	53	Geist	Empirische Kulturwissenschaften	21	weiblich	Tabelle	Wahrscheinlichkeiten
14	Natur	Biochemie	19	weiblich	Stapeldiagramm	Wahrscheinlichkeiten	54	Geist	Germanistik & Politikwissenschaft	19	weiblich	Stapeldiagramm	Wahrscheinlichkeiten
15	Geist	Politik/Internationale Beziehungen	20	weiblich	Protokollliste	Einzelereignisse	55	Geist	Sozialpädagogik	52	weiblich	Protokollliste	Einzelereignisse
16	Natur	Wirtschaftsmathematik	19	weiblich	Tabelle	Häufigkeiten	56	Natur	Biochemie	34	weiblich	Tabelle	Häufigkeiten
17	Natur	Volkswirtschaftslehre regional	20	weiblich	Stapeldiagramm	Häufigkeiten	57	Natur	Psychologie	24	weiblich	Stapeldiagramm	Häufigkeiten
18	Geist	Kommunikationswissenschaften	19	weiblich	Tabelle	Wahrscheinlichkeiten	58	Natur	Bioinformatik	21	weiblich	Tabelle	Wahrscheinlichkeiten
19	Natur	Medizin	21	männlich	Stapeldiagramm	Wahrscheinlichkeiten	59	Natur	Geografie	33	weiblich	Stapeldiagramm	Wahrscheinlichkeiten
20	Natur	Psychologie	19	weiblich	Protokollliste	Einzelereignisse	60	Natur	Medizin	51	männlich	Protokollliste	Einzelereignisse
21	Geist	Germanistik, Geschichte	20	weiblich	Tabelle	Häufigkeiten	61	Geist	Rhetorik/Politik	22	männlich	Tabelle	Häufigkeiten
22	Geist	Theaterwissenschaften/ Regie	21	weiblich	Stapeldiagramm	Häufigkeiten	62	Natur	Betriebswirtschaftslehre	23	weiblich	Stapeldiagramm	Häufigkeiten
23	Geist	Musik-, Theaterwissenschaften oder Germanistik	21	weiblich	Tabelle	Wahrscheinlichkeiten	63	Natur	Psychologie	21	weiblich	Tabelle	Wahrscheinlichkeiten
24	Natur	Medizin	20	weiblich	Stapeldiagramm	Wahrscheinlichkeiten	64	Natur	Psychologie	21	weiblich	Stapeldiagramm	Wahrscheinlichkeiten
25	Natur	Medizin	20	weiblich	Protokollliste	Einzelereignisse	65	Geist	Politik, Neuere englische Literatur	27	weiblich	Protokollliste	Einzelereignisse
26	Geist	Philosophie, Theologie, Philologie	21	männlich	Tabelle	Häufigkeiten	66	Natur	Medizin	25	weiblich	Tabelle	Häufigkeiten
27	Geist	Philosophie	20	männlich	Stapeldiagramm	Häufigkeiten	67	Natur	Physik (promoviert)	35	männlich	Stapeldiagramm	Häufigkeiten
28	Geist	Politik	19	weiblich	Tabelle	Wahrscheinlichkeiten	68	Geist	Englisch, Sport	24	männlich	Tabelle	Wahrscheinlichkeiten
29	Geist	Jura	19	weiblich	Stapeldiagramm	Wahrscheinlichkeiten	69	Geist	Theologie/ Italienisch	48	männlich	Stapeldiagramm	Wahrscheinlichkeiten
30	Natur	Molekulare Biotechnologie/ Medizin	20	weiblich	Protokollliste	Einzelereignisse	70	Natur	Mathematik	26	männlich	Protokollliste	Einzelereignisse
31	Natur	Medizin	20	weiblich	Tabelle	Häufigkeiten	71	Geist	Deutsch, Latein	21	weiblich	Tabelle	Häufigkeiten
32	Natur	Medizin oder Jura	20	weiblich	Stapeldiagramm	Häufigkeiten	72	Geist	Germanistik, Geschichte	27	weiblich	Stapeldiagramm	Häufigkeiten
33	Natur	Medizin	20	weiblich	Tabelle	Wahrscheinlichkeiten	73	Natur	Bioinformatik	26	weiblich	Tabelle	Wahrscheinlichkeiten
34	Natur	Zahnmedizin	20	weiblich	Stapeldiagramm	Wahrscheinlichkeiten	74	Geist	Slawistik Englisch, Geschichte	27	weiblich	Stapeldiagramm	Wahrscheinlichkeiten
35	Natur	Medizin oder Lehramt	20	weiblich	Protokollliste	Einzelereignisse	75	Geist	Spanisch, Englisch	20	männlich	Protokollliste	Einzelereignisse
36	Natur	Mathematik Französisch	20	weiblich	Tabelle	Häufigkeiten	76	Geist	Politik, Soziologie	21	weiblich	Tabelle	Häufigkeiten
37	Geist	Theologie/Islamwissenschaften, Politik	20	weiblich	Stapeldiagramm	Häufigkeiten	77	Geist	Linguistik, Neuere Geschichte	25	männlich	Stapeldiagramm	Häufigkeiten
38	Natur	Psychologie/ Lehramt	20	männlich	Tabelle	Wahrscheinlichkeiten	78	Geist	Englisch, Französisch	20	weiblich	Tabelle	Wahrscheinlichkeiten
39	Natur	Psychologie oder Medizin	20	weiblich	Stapeldiagramm	Wahrscheinlichkeiten	79	Geist	Literaturwissenschaften, Soziologie	22	weiblich	Stapeldiagramm	Wahrscheinlichkeiten
40	Geist	Medien- und Kommunikationswissenschaft	19	weiblich	Protokollliste	Einzelereignisse	80	Natur	Biochemie	26	männlich	Protokollliste	Einzelereignisse

## Anhang A1.2: Experiment 1: Instruktionen

### Instruktion – Bedingung SH

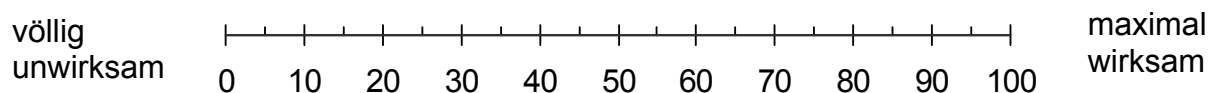
Wie Sie vielleicht wissen, kann man durch das Versprühen von Silberjodid die Regentätigkeit von Wolken beeinflussen. Das in der Luft vorhandene Wasser lagert sich an die Silberjodidkristalle an und regnet ab, sobald eine kritische Tröpfchengröße erreicht wird. Diese Technik wird in Trockengebieten genutzt, um Wolken, die eigentlich nicht abregnen würden, doch noch zum Regnen zu bringen.

Ein Chemieunternehmen arbeitet an der Entwicklung neuer Substanzen, die Wolken mittels Wolkenimpfung zum Abregnen bringen sollen. Dazu wurden mit einem Flugzeug in einem Trockengebiet sechs Testflüge unternommen, bei denen jeweils die Wirksamkeit von einer Substanz untersucht wurde. Pro Substanz wurden bei stabiler Wetterlage 20 geimpfte und 20 ungeimpfte Wolken gleichen Wolkentyps beobachtet. Keine dieser Wolken regnete zu Beginn der Beobachtungsphase. Für jede Substanz finden sie die Testergebnisse auf einem separaten Blatt in einem Säulendiagramm zusammen gefasst. Die zusammen gefassten Ergebnisse enthalten vier verschiedene Arten von Informationen:

1. Anzahl der geimpften Wolken, die im Anschluss an eine Impfung zu regnen begonnen haben.
2. Anzahl der geimpften Wolken, die im Anschluss an eine Impfung nicht zu regnen begonnen haben
3. Anzahl der nicht geimpften Wolken, die spontan zu regnen begonnen haben
4. Anzahl der nicht geimpften Wolken, die nicht zu regnen begonnen haben.

Ihre Aufgabe ist es nun, an Hand der im Diagramm zusammen gefassten Ergebnisse die Wirksamkeit einer jeden Substanz zu beurteilen. Bitte prüfen sie dazu die Säuleneinträge des Diagramms so genau wie möglich. Zur Einschätzung der Wirksamkeit steht Ihnen für jede Substanz eine Skala zum Ankreuzen zur Verfügung. Die Skala reicht von 0 bis 100. 0 bedeutet dabei, dass diese Substanz völlig unwirksam ist und keinerlei Einfluss auf das Regenverhalten der Wolken hat, 100 bedeutet, dass diese Substanz maximal wirksam ist und man damit jede Wolke zum Abregnen bringen kann.

#### Abbildung 1: Wirksamkeitsskala



Wenn Sie die Wirksamkeit der Substanz nicht nur geschätzt, sondern einen genauen Wert berechnet haben, so tragen Sie diesen Wert in das dafür zur Verfügung stehende Kästchen ein:

Sie stehen bei der Bearbeitung der Aufgaben nicht unter Zeitdruck. Gehen Sie die Aufgabenblätter der Reihe nach durch und blättern sie nicht zurück. Wenn Sie wollen, können Sie sich auf den Blättern Notizen machen. Am Ende des Versuchs werden wir Sie zu Ihrer Vorgehensweise befragen. Das Instruktionsblatt können Sie während des ganzen Versuchs einsehen. Wenn Sie Fragen haben, wenden Sie sich an den Versuchsleiter.

## Instruktion - Bedingung SW

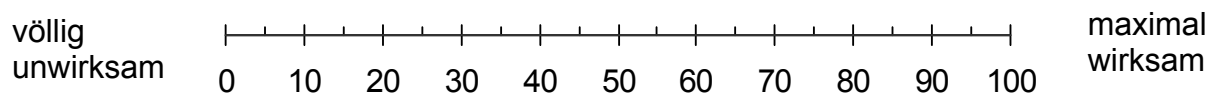
Wie Sie vielleicht wissen, kann man durch das Versprühen von Silberjodid die Regentätigkeit von Wolken beeinflussen. Das in der Luft vorhandene Wasser lagert sich an die Silberjodidkristalle an und regnet ab, sobald eine kritische Tröpfchengröße erreicht wird. Diese Technik wird in Trockengebieten genutzt, um Wolken, die eigentlich nicht abregnen würden, doch noch zum Regnen zu bringen.

Ein Chemieunternehmen arbeitet an der Entwicklung neuer Substanzen, die Wolken mittels Wolkenimpfung zum Abregnen bringen sollen. Dazu wurden mit einem Flugzeug in einem Trockengebiet sechs Testflüge unternommen, bei denen jeweils die Wirksamkeit von einer Substanz untersucht wurde. Pro Substanz wurden bei stabiler Wetterlage 20 geimpfte und 20 ungeimpfte Wolken gleichen Wolkentyps beobachtet. Keine dieser Wolken regnete zu Beginn der Beobachtungsphase. Für jede Substanz finden sie die Testergebnisse auf einem separaten Blatt in einem Säulendiagramm zusammen gefasst. Die zusammen gefassten Ergebnisse enthalten vier verschiedene Arten von Informationen:

1. Prozentzahl der geimpften Wolken, die im Anschluss an eine Impfung zu regnen begonnen haben.
2. Prozentzahl der geimpften Wolken, die im Anschluss an eine Impfung nicht zu regnen begonnen haben
3. Prozentzahl der nicht geimpften Wolken, die spontan zu regnen begonnen haben
4. Prozentzahl der nicht geimpften Wolken, die nicht zu regnen begonnen haben.

Ihre Aufgabe ist es nun, an Hand der im Diagramm zusammen gefassten Ergebnisse die Wirksamkeit einer jeden Substanz zu beurteilen. Bitte prüfen sie dazu die Säuleneinträge des Diagramms so genau wie möglich. Zur Einschätzung der Wirksamkeit steht Ihnen für jede Substanz eine Skala zum Ankreuzen zur Verfügung. Die Skala reicht von 0 bis 100. 0 bedeutet dabei, dass diese Substanz völlig unwirksam ist und keinerlei Einfluss auf das Regenverhalten der Wolken hat, 100 bedeutet, dass diese Substanz maximal wirksam ist und man damit jede Wolke zum Abregnen bringen kann.

### Abbildung 1: Wirksamkeitsskala



Wenn Sie die Wirksamkeit der Substanz nicht nur geschätzt, sondern einen genauen Wert berechnet haben, so tragen Sie diesen Wert in das dafür zur Verfügung stehende Kästchen ein:

Sie stehen bei der Bearbeitung der Aufgaben nicht unter Zeitdruck. Gehen Sie die Aufgabenblätter der Reihe nach durch und blättern sie nicht zurück. Wenn Sie wollen, können Sie sich auf den Blättern Notizen machen. Am Ende des Versuchs werden wir Sie zu Ihrer Vorgehensweise befragen. Das Instruktionsblatt können Sie während des ganzen Versuchs einsehen. Wenn Sie Fragen haben, wenden Sie sich an den Versuchsleiter.

## Instruktion – Bedingung TH

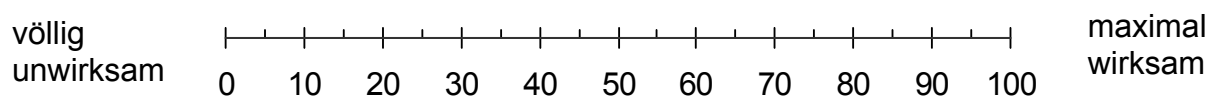
Wie Sie vielleicht wissen, kann man durch das Versprühen von Silberjodid die Regentätigkeit von Wolken beeinflussen. Das in der Luft vorhandene Wasser lagert sich an die Silberjodidkristalle an und regnet ab, sobald eine kritische Tröpfchengröße erreicht wird. Diese Technik wird in Trockengebieten genutzt, um Wolken, die eigentlich nicht abregnen würden, doch noch zum Regnen zu bringen.

Ein Chemieunternehmen arbeitet an der Entwicklung neuer Substanzen, die Wolken mittels Wolkenimpfung zum Abregnen bringen sollen. Dazu wurden mit einem Flugzeug in einem Trockengebiet sechs Testflüge unternommen, bei denen jeweils die Wirksamkeit von einer Substanz untersucht wurde. Pro Substanz wurden bei stabiler Wetterlage 20 geimpfte und 20 ungeimpfte Wolken gleichen Wolkentyps beobachtet. Keine dieser Wolken regnete zu Beginn der Beobachtungsphase. Für jede Substanz finden sie die Testergebnisse auf einem separaten Blatt in einer Tabelle zusammen gefasst. Die zusammen gefassten Ergebnisse enthalten vier verschiedene Arten von Informationen:

1. Anzahl der geimpften Wolken, die im Anschluss an eine Impfung zu regnen begonnen haben.
2. Anzahl der geimpften Wolken, die im Anschluss an eine Impfung nicht zu regnen begonnen haben
3. Anzahl der nicht geimpften Wolken, die spontan zu regnen begonnen haben
4. Anzahl der nicht geimpften Wolken, die nicht zu regnen begonnen haben.

Ihre Aufgabe ist es nun, an Hand der in der Tabelle zusammen gefassten Ergebnisse die Wirksamkeit einer jeden Substanz zu beurteilen. Bitte prüfen sie dazu die Tabelleneinträge so genau wie möglich. Zur Einschätzung der Wirksamkeit steht Ihnen für jede Substanz eine Skala zum Ankreuzen zur Verfügung. Die Skala reicht von 0 bis 100. 0 bedeutet dabei, dass diese Substanz völlig unwirksam ist und keinerlei Einfluss auf das Regenverhalten der Wolken hat, 100 bedeutet, dass diese Substanz maximal wirksam ist und man damit jede Wolke zum Abregnen bringen kann.

### Abbildung 1: Wirksamkeitsskala



Wenn Sie die Wirksamkeit der Substanz nicht nur geschätzt, sondern einen genauen Wert berechnet haben, so tragen Sie diesen Wert in das dafür zur Verfügung stehende Kästchen ein:

Sie stehen bei der Bearbeitung der Aufgaben nicht unter Zeitdruck. Gehen Sie die Aufgabenblätter der Reihe nach durch und blättern sie nicht zurück. Wenn Sie wollen, können Sie sich auf den Blättern Notizen machen. Am Ende des Versuchs werden wir Sie zu Ihrer Vorgehensweise befragen. Das Instruktionsblatt können Sie während des ganzen Versuchs einsehen. Wenn Sie Fragen haben, wenden Sie sich an den Versuchsleiter.

## Instruktion – TW

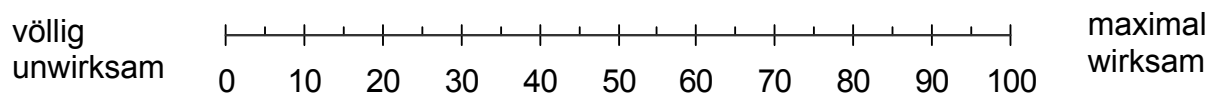
Wie Sie vielleicht wissen, kann man durch das Versprühen von Silberjodid die Regentätigkeit von Wolken beeinflussen. Das in der Luft vorhandene Wasser lagert sich an die Silberjodidkristalle an und regnet ab, sobald eine kritische Tröpfchengröße erreicht wird. Diese Technik wird in Trockengebieten genutzt, um Wolken, die eigentlich nicht abregnen würden, doch noch zum Regnen zu bringen.

Ein Chemieunternehmen arbeitet an der Entwicklung neuer Substanzen, die Wolken mittels Wolkenimpfung zum Abregnen bringen sollen. Dazu wurden mit einem Flugzeug in einem Trockengebiet sechs Testflüge unternommen, bei denen jeweils die Wirksamkeit von einer Substanz untersucht wurde. Pro Substanz wurden bei stabiler Wetterlage 20 geimpfte und 20 ungeimpfte Wolken gleichen Wolkentyps beobachtet. Keine dieser Wolken regnete zu Beginn der Beobachtungsphase. Für jede Substanz finden sie die Testergebnisse auf einem separaten Blatt in einer Tabelle zusammen gefasst. Die zusammen gefassten Ergebnisse enthalten vier verschiedene Arten von Informationen:

1. Prozentzahl der geimpften Wolken, die im Anschluss an eine Impfung zu regnen begonnen haben.
2. Prozentzahl der geimpften Wolken, die im Anschluss an eine Impfung nicht zu regnen begonnen haben
3. Prozentzahl der nicht geimpften Wolken, die spontan zu regnen begonnen haben
4. Prozentzahl der nicht geimpften Wolken, die nicht zu regnen begonnen haben.

Ihre Aufgabe ist es nun, an Hand der in der Tabelle zusammen gefassten Ergebnisse die Wirksamkeit einer jeden Substanz zu beurteilen. Bitte prüfen sie dazu die Tabelleneinträge so genau wie möglich. Zur Einschätzung der Wirksamkeit steht Ihnen für jede Substanz eine Skala zum Ankreuzen zur Verfügung. Die Skala reicht von 0 bis 100. 0 bedeutet dabei, dass diese Substanz völlig unwirksam ist und keinerlei Einfluss auf das Regenverhalten der Wolken hat, 100 bedeutet, dass diese Substanz maximal wirksam ist und man damit jede Wolke zum Abregnen bringen kann.

### Abbildung 1: Wirksamkeitsskala



Wenn Sie die Wirksamkeit der Substanz nicht nur geschätzt, sondern einen genauen Wert berechnet haben, so tragen Sie diesen Wert in das dafür zur Verfügung stehende Kästchen ein:

Sie stehen bei der Bearbeitung der Aufgaben nicht unter Zeitdruck. Gehen Sie die Aufgabenblätter der Reihe nach durch und blättern sie nicht zurück. Wenn Sie wollen, können Sie sich auf den Blättern Notizen machen. Am Ende des Versuchs werden wir Sie zu Ihrer Vorgehensweise befragen. Das Instruktionsblatt können Sie während des ganzen Versuchs einsehen. Wenn Sie Fragen haben, wenden Sie sich an den Versuchsleiter.



## Instruktion – Bedingung PL (Kontrollbedingung)

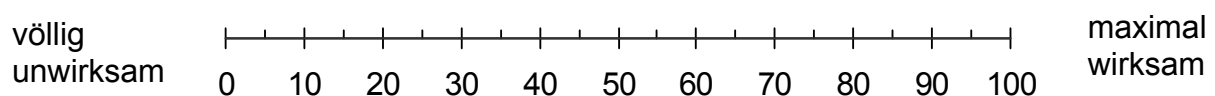
Wie Sie vielleicht wissen, kann man durch das Versprühen von Silberjodid die Regentätigkeit von Wolken beeinflussen. Das in der Luft vorhandene Wasser lagert sich an die Silberjodidkristalle an und regnet ab, sobald eine kritische Tröpfchengröße erreicht wird. Diese Technik wird in Trockengebieten genutzt, um Wolken, die eigentlich nicht abregnen würden, doch noch zum Regnen zu bringen.

Ein Chemieunternehmen arbeitet an der Entwicklung neuer Substanzen, die Wolken mittels Wolkenimpfung zum Abregnen bringen sollen. Dazu wurden mit einem Flugzeug in einem Trockengebiet sechs Testflüge unternommen, bei denen jeweils die Wirksamkeit von einer Substanz untersucht wurde. Pro Substanz wurden bei stabiler Wetterlage 20 geimpfte und 20 ungeimpfte Wolken gleichen Wolkentyps beobachtet. Keine dieser Wolken regnete zu Beginn der Beobachtungsphase. Für jede Substanz finden sie die Testergebnisse auf einem separaten Blatt in einer Liste zusammen gefasst. Die zusammen gefassten Ergebnisse enthalten vier verschiedene Arten von Informationen:

1. Eine Wolke wurde mit der Substanz geimpft und fing im Anschluss an die Impfung zu regnen an.
2. Eine Wolke wurde mit der Substanz geimpft und fing im Anschluss an die Impfung nicht zu regnen an.
3. Eine Wolke wurde nicht mit der Substanz geimpft und fing spontan an zu regnen.
4. Eine Wolke wurde nicht mit der Substanz geimpft und fing nicht zu regnen an.

Ihre Aufgabe ist es nun, an Hand der in der Liste zusammen gefassten Ergebnisse die Wirksamkeit einer jeden Substanz zu beurteilen. Bitte prüfen sie dazu die Listeneinträge so genau wie möglich. Zur Einschätzung der Wirksamkeit steht Ihnen für jede Substanz eine Skala zum Ankreuzen zur Verfügung. Die Skala reicht von 0 bis 100. 0 bedeutet dabei, dass diese Substanz völlig unwirksam ist und keinerlei Einfluss auf das Regenverhalten der Wolken hat, 100 bedeutet, dass diese Substanz maximal wirksam ist und man damit jede Wolke zum Abregnen bringen kann.

### Abbildung 1: Wirksamkeitsskala



Wenn Sie die Wirksamkeit der Substanz nicht nur geschätzt, sondern einen genauen Wert berechnet haben, so tragen Sie diesen Wert in das dafür zur Verfügung stehende Kästchen ein:

Sie stehen bei der Bearbeitung der Aufgaben nicht unter Zeitdruck. Gehen Sie die Aufgabenblätter der Reihe nach durch und blättern sie nicht zurück. Wenn Sie wollen, können Sie sich auf den Blättern Notizen machen. Am Ende des Versuchs werden wir Sie zu Ihrer Vorgehensweise befragen. Das Instruktionsblatt können Sie während des ganzen Versuchs einsehen. Wenn Sie Fragen haben, wenden Sie sich an den Versuchsleiter.

**Anhang A1.3:** Experiment 1: Fragen zur Vorgehensweise**Fragen zur Vorgehensweise**

Wie würden Sie Ihre Vorgehensweise bei der Bestimmung der Wirksamkeit einer Substanz in eigenen Worten beschreiben ?

Haben Sie versucht die Wirksamkeit einer Substanz zu berechnen? Wenn ja, bitte beschreiben Sie Ihren Rechenweg (ggf. mit Formel), oder falls Sie mehrere Rechenwege ausprobiert haben, Ihre Rechenwege (ggf. mit Formeln).

## Anhang A1.4: Experiment 1: Nachbefragung

### Nachbefragung

Sie haben an einem Experiment teil genommen, in dem der Einfluss verschiedener Darstellungsformen (z.B. Tabelle, Diagramm und Ereignisaufzählung) und Informationsarten (z.B. Darbietung von Informationen als Häufigkeiten oder Wahrscheinlichkeiten) auf den Erwerb von Wissen über kausale Zusammenhänge untersucht wurde. Auf den beiden Beiblättern sehen Sie die fünf in diesem Experiment eingesetzten Darstellungsformen. Bitte sehen Sie sich die Inhalte auf diesen Blättern kurz an. Es ist für uns von besonderem Interesse, **welche Erfahrungen Sie mit diesen oder ähnlich gestalteten Informationsdarbietungsformen** in der Schule, im Studium (Bücher, Fachzeitschriften, Computerprogramme, Internet), im Beruf oder in der Freizeit (Bücher, Zeitschriften, Internet, Fernsehen) gemacht haben. Bei den nachfolgenden Fragen handelt es sich um Aufgaben zum Ankreuzen.

Wie häufig haben oder hatten Sie in Schule, Studium, Beruf und Freizeit mit Listen, in denen Ereignisse protokolliert werden, zu tun ?

	Ein- bis mehrmals pro Woche	weniger als einmal die Woche	weniger als einmal im Monat	weniger als einmal im Jahr	gar nie
Schule					
Studium					
Beruf					
Freizeit					

Wie häufig haben oder hatten Sie in Schule, Studium, Beruf und Freizeit mit Tabellen zu tun ?

	Ein- bis mehrmals pro Woche	weniger als einmal die Woche	weniger als einmal im Monat	weniger als einmal im Jahr	gar nie
Schule					
Studium					
Beruf					
Freizeit					

Wie häufig haben oder hatten Sie in Schule, Studium, Beruf und Freizeit mit Diagrammen oder Grafiken zu tun ?

	Ein- bis mehrmals pro Woche	weniger als einmal die Woche	weniger als einmal im Monat	weniger als einmal im Jahr	gar nie
Schule					
Studium					
Beruf					
Freizeit					

Wie oft erstellen Sie selber Ereignislisten, Tabellen oder Diagramme/Grafiken ?

	ein- bis mehrmals pro Woche	weniger als einmal die Woche	weniger als einmal im Monat	weniger als einmal im Jahr	gar nie
Liste					
Tabelle					
Diagramm					

Wie häufig rechnen Sie mit den Grundrechenarten (Addition, Subtraktion, Multiplikation, Division) ?

ein- bis mehrmals pro Woche	weniger als einmal die Woche	weniger als einmal im Monat	weniger als einmal im Jahr	gar nie

Wie sicher fühlen Sie sich in den Grundrechenarten ?

- sehr sicher
- eher sicher
- weder sicher noch unsicher
- eher unsicher
- sehr unsicher

Wie häufig wenden Sie Prozentrechnen an ?

ein- bis mehrmals pro Woche	weniger als einmal die Woche	weniger als einmal im Monat	weniger als einmal im Jahr	gar nie

Wie sicher fühlen Sie sich im Prozentrechnen ?

- sehr sicher
- eher sicher
- weder sicher noch unsicher
- eher unsicher
- sehr unsicher

Wie häufig arbeiten Sie mit einem Tabellenkalkulationsprogramm (z.B. Excel, QuattroPro etc.)

ein- bis mehrmals pro Woche	weniger als einmal die Woche	weniger als einmal im Monat	weniger als einmal im Jahr	gar nie

Vielen Dank für die Teilnahme an unserem Experiment !

## Anhang A1.5: Experiment 1: Eingesetzte Präsentationsformen und Inhaltsarten

### 1. Listen, in denen Ereignisabfolgen protokolliert werden

Zustand geimpfter und ungeimpfter regnender und nicht regnender Wolken am Ende des Beobachtungszeitraums.

Wolke	geimpft?	Zustand der Wolke am Ende des Beobachtungszeitraums
1	Ja	Kein Regen
2	Nein	Kein Regen
3	Ja	Regen
4	Nein	Kein Regen
5	Ja	Kein Regen
6	Nein	Regen
7	Ja	Regen
8	Nein	Kein Regen
9	Ja	Regen
10	Nein	Kein Regen
11	Ja	Kein Regen
12	Nein	Kein Regen
13	Ja	Regen
14	Nein	Kein Regen
15	Ja	Kein Regen
16	Nein	Regen
17	Ja	Kein Regen
18	Nein	Kein Regen
19	Ja	Regen
20	Nein	Kein Regen
21	Ja	Kein Regen
22	Nein	Regen
23	Ja	Regen
24	Nein	Kein Regen
25	Ja	Regen
26	Nein	Kein Regen
27	Ja	Kein Regen
28	Nein	Regen
29	Ja	Regen
30	Nein	Kein Regen
31	Ja	Kein Regen
32	Nein	Kein Regen
33	Ja	Kein Regen
34	Nein	Regen
35	Ja	Kein Regen
36	Nein	Kein Regen
37	Ja	Regen
38	Nein	Kein Regen
39	Ja	Regen
40	Nein	Kein Regen

## 2. Tabelle mit Häufigkeitsangaben

Anzahl regnender und nicht regnender Wolken bei geimpften und nicht geimpften Wolken am Ende des Beobachtungszeitraums

	geimpfte Wolken	nicht geimpfte Wolken
regnende Wolken	10	5
nicht regnende Wolken	10	15

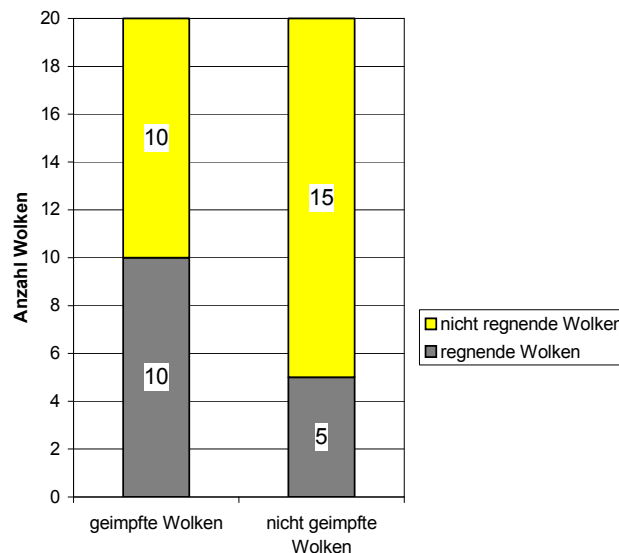
## 3. Tabelle mit Wahrscheinlichkeitsangaben

Wahrscheinlichkeit regnender und nicht regnender Wolken in Prozent bei geimpften und nicht geimpften Wolken am Ende des Beobachtungszeitraums

	geimpfte Wolken	nicht geimpfte Wolken
regnende Wolken	50 %	25 %
nicht regnende Wolken	50 %	75 %

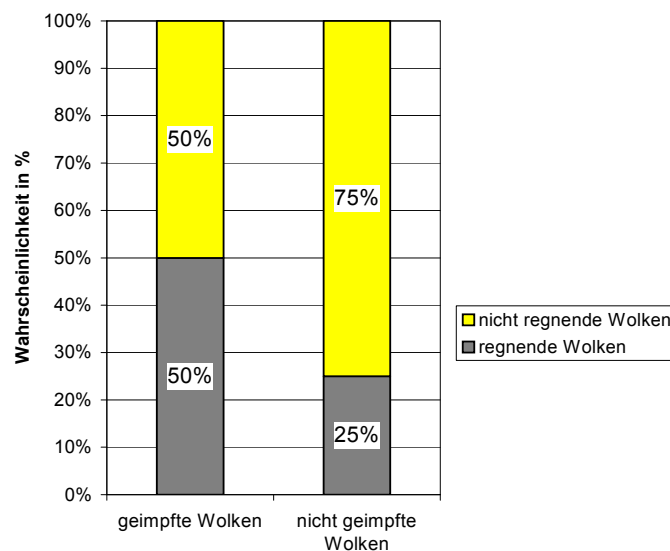
## 4. Diagramm mit Häufigkeitsangaben

Anzahl regnender und nicht regnender Wolken bei geimpften und nicht geimpften Wolken am Ende des Beobachtungszeitraums



## 5. Diagramm mit Wahrscheinlichkeitsangaben

Wahrscheinlichkeit regnender und nicht regnender Wolken in Prozent bei geimpften und nicht geimpften Wolken am Ende des Beobachtungszeitraums



## Anhang A2

### Anhang A2.1: Experiment 2: Zusammensetzung der Teilnehmer

**Tabelle A2:** Experiment 2: Zusammensetzung der Teilnehmer hinsichtlich ihrer realen oder beabsichtigten Studienfächer und sonstiger demografischer Merkmale.

Vp	Studienrichtung	Studienfach/ Studienwunsch	Alter	Geschlecht	Präsentationsform	Darbietungsart	Vp	Studienrichtung	Studienfach/ Studienwunsch	Alter	Geschlecht	Präsentationsform	Darbietungsart
1	Geist	Germanistik	22	weiblich	Text	summarisch	61	Natur	Maschinenbau	25	männlich	Text	summarisch
2	Geist	Jura	22	weiblich	Text	einzelfallweise	62	Natur	Geologie-Doktorand	35	männlich	Text	einzelfallweise
3	Geist	Geschichte, Französisch	20	weiblich	Grafik	summarisch	63	Natur	BWL	21	weiblich	Grafik	summarisch
4	Geist	Geschichte, Französisch	20	männlich	Grafik	einzelfallweise	64	Natur	Biologie	21	männlich	Grafik	einzelfallweise
7	Geist	Ev. Theologie	42	männlich	Text	summarisch	67	Natur	VWL	21	weiblich	Text	summarisch
8	Geist	Ev. Theologie	27	männlich	Text	einzelfallweise	68	Natur	Biologie	19	weiblich	Text	einzelfallweise
9	Geist	Sozialpädagogik	27	weiblich	Grafik	summarisch	69	Natur	VWL, Politik, Spanisch	22	weiblich	Grafik	summarisch
10	Geist	Pädagogik	25	weiblich	Grafik	einzelfallweise	70	Natur	VWL	19	weiblich	Grafik	einzelfallweise
13	Geist	Pädagogik	21	weiblich	Text	summarisch	73	Natur	BWL	27	weiblich	Text	summarisch
14	Geist	Jura	22	männlich	Text	einzelfallweise	74	Natur	Medizin	26	weiblich	Text	einzelfallweise
15	Geist	Französisch, Linguistik	24	weiblich	Grafik	summarisch	75	Natur	Bioinformatik	22	weiblich	Grafik	summarisch
16	Geist	Jura	26	weiblich	Grafik	einzelfallweise	76	Natur	Psychologie	20	weiblich	Grafik	einzelfallweise
19	Geist	Pädagogik	22	weiblich	Text	summarisch	79	Natur	Biologie, Deutsch	23	weiblich	Text	summarisch
20	Geist	Englisch, Französisch	24	weiblich	Text	einzelfallweise	80	Natur	Chemie-Ingenieurin	44	weiblich	Text	einzelfallweise
21	Geist	Theologie, Soziologie	23	männlich	Grafik	summarisch	81	Natur	Medizin	22	männlich	Grafik	summarisch
22	Geist	Germanistik, Geograf	25	weiblich	Grafik	einzelfallweise	82	Natur	Biologie	23	weiblich	Grafik	einzelfallweise
25	Geist	Romanistik	26	weiblich	Text	summarisch	85	Natur	Mathe, Sport	21	weiblich	Text	summarisch
26	Geist	Englisch, Geschichte	25	weiblich	Text	einzelfallweise	86	Natur	Biologie, Politik	26	weiblich	Text	einzelfallweise
27	Geist	Kunstgeschichte, EKW	21	weiblich	Grafik	summarisch	87	Natur	BWL	25	weiblich	Grafik	summarisch
28	Geist	Kunstgeschichte	24	weiblich	Grafik	einzelfallweise	88	Natur	Biologie	21	männlich	Grafik	einzelfallweise
31	Geist	Pädagogik, EKW	20	weiblich	Text	summarisch	91	Natur	Medizin	21	weiblich	Text	summarisch
32	Geist	Rhetorik	21	weiblich	Text	einzelfallweise	92	Natur	Medizin	24	weiblich	Text	einzelfallweise
33	Geist	Englisch, Spanisch	29	weiblich	Grafik	summarisch	93	Natur	Biologie	24	männlich	Grafik	summarisch
34	Geist	Romanistik, EKW	32	weiblich	Grafik	einzelfallweise	94	Natur	Physik	24	männlich	Grafik	einzelfallweise
37	Geist	Jura	24	weiblich	Text	summarisch	97	Natur	Pharmazie	57	männlich	Text	summarisch
38	Geist	Islamkunde, Religion	23	männlich	Text	einzelfallweise	98	Natur	Medizin	22	weiblich	Text	einzelfallweise
39	Geist	Sport	19	weiblich	Grafik	summarisch	99	Natur	Psychologie	22	weiblich	Grafik	summarisch
40	Geist	Germanistik, Japanologie	26	männlich	Grafik	einzelfallweise	100	Natur	Medizin	27	weiblich	Grafik	einzelfallweise
43	Geist	Pädagogik	28	weiblich	Text	summarisch	103	Natur	Mathe, Deutsch	20	weiblich	Text	summarisch
44	Geist	Deutsch, Kath. Theologie	24	weiblich	Text	einzelfallweise	104	Natur	Biochemie	22	männlich	Text	einzelfallweise
45	Natur	Patholinguistik	21	männlich	Grafik	summarisch	105	Natur	Geografie	21	männlich	Grafik	summarisch
46	Geist	Kunstgeschichte	22	weiblich	Grafik	einzelfallweise	106	Natur	Mathe, Philosophie,	20	männlich	Grafik	einzelfallweise
49	Geist	Politikwissenschaft	20	weiblich	Text	summarisch	109	Natur	Physik	20	männlich	Text	summarisch
50	Geist	Philosophie, Englisch	19	weiblich	Text	einzelfallweise	110	Natur	Biologie, Theologie	20	weiblich	Text	einzelfallweise
51	Geist	Theologie	25	männlich	Grafik	summarisch	111	Geist	Theater- und Medienwissenschaften	20	weiblich	Grafik	summarisch
52	Geist	Jura	19	weiblich	Grafik	einzelfallweise	112	Natur	Physik	25	männlich	Grafik	einzelfallweise
55	Geist	Religions-, Kommunikationswissenschaften	21	weiblich	Text	summarisch	115	Natur	Biologie	34	weiblich	Text	summarisch
56	Geist	Englisch, Musik	20	weiblich	Text	einzelfallweise	116	Natur	Mathematik	22	männlich	Text	einzelfallweise
57	Geist	Rhetorik, Soziologie	20	weiblich	Grafik	summarisch	117	Natur	BWL	28	männlich	Grafik	summarisch
58	Geist	Soziologie	26	weiblich	Grafik	einzelfallweise	118	Natur	Biologie	23	weiblich	Grafik	einzelfallweise

## Anhang A2.2: Experiment 2: Instruktionen

### Instruktion – Bedingung *SD sum* (Ex 2) und *ohne N sum* (Ex 3)

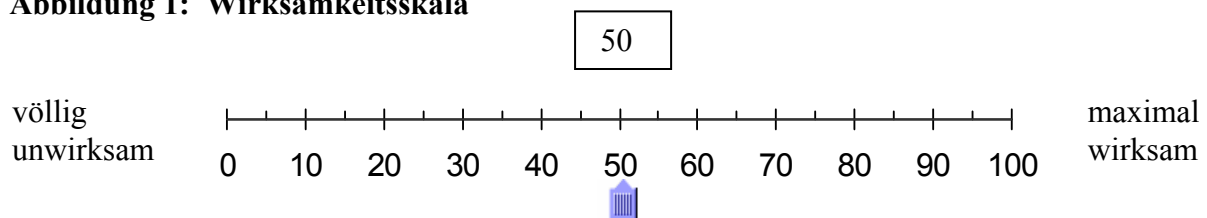
Wie Sie vielleicht wissen, kann man durch das Versprühen von Silberjodid die Regentätigkeit von Wolken beeinflussen. Das in der Luft vorhandene Wasser lagert sich an die Silberjodidkristalle an und regnet ab, sobald eine kritische Tröpfchengröße erreicht wird. Diese Technik wird in Trockengebieten genutzt, um Wolken, die eigentlich nicht abregnen würden, doch noch zum Regnen zu bringen.

Ein Chemieunternehmen arbeitet an der Entwicklung neuer Substanzen, die Wolken mittels Wolkenimpfung zum Abregnen bringen sollen. Dazu wurden mit einem Flugzeug in einem Trockengebiet sechs Testflüge unternommen, bei denen jeweils die Wirksamkeit von einer Substanz untersucht wurde. Pro Substanz wurden bei stabiler Wetterlage 20 geimpfte und 20 ungeimpfte Wolken beobachtet. Die 40 Wolken weisen untereinander keine wesentlichen Unterschiede auf und sind miteinander vergleichbar. Die meteorologischen Rahmenbedingungen sind für alle Wolken und bei jedem Einsatz einer Substanz gleich. Zu Beginn einer jeden Beobachtungsphase regnete keine der Wolken. Für jede Substanz finden Sie die Testergebnisse auf dem Computer-Bildschirm in einem Säulendiagramm zusammengefasst. Das Diagramm und sein Inhalt werden Ihnen auf Seite 2 erläutert. Die darin zusammengefassten Ergebnisse enthalten vier verschiedene Arten von Informationen:

1. Anzahl der geimpften Wolken, die im Anschluss an eine Impfung zu regnen begonnen haben.
2. Anzahl der geimpften Wolken, die im Anschluss an eine Impfung nicht zu regnen begonnen haben
3. Anzahl der nicht geimpften Wolken, die spontan zu regnen begonnen haben
4. Anzahl der nicht geimpften Wolken, die nicht zu regnen begonnen haben.

Ihre Aufgabe ist es nun, an Hand der im Diagramm zusammengefassten Ergebnisse die Wirksamkeit der jeweiligen Substanz zu beurteilen. Bitte schauen Sie sich dabei die Säuleneinträge des Diagramms so aufmerksam wie möglich an. Das Diagramm bleibt maximal 7 Minuten eingeblendet. Wenn Sie sich schon vorher über die Wirksamkeit der Substanz im Klaren sind, können Sie die Weitertaste drücken. Sie sehen dann zur Einschätzung der Wirksamkeit einer Substanz eine Skala eingeblendet, auf der Sie mit Hilfe der Maus einen Regler betätigen und Zahlenwerte eingeben können. Die Skala reicht von 0 bis 100. "0" bedeutet dabei, dass die Substanz völlig unwirksam ist und keinerlei Einfluss auf das Regenverhalten der Wolken hat, "100" bedeutet, dass diese Substanz maximal wirksam ist und man damit jede Wolke zum Abregnen bringen kann.

**Abbildung 1: Wirksamkeitsskala**



Erst wenn Sie mit Hilfe der Maus einen Zahlenwert eingetragen haben, können Sie die Weitertaste anklicken um die nächste Aufgabe zu bearbeiten. Der Versuch beginnt mit einer Probesubstanz, deren Wirksamkeit Sie ebenfalls beurteilen sollen. Sie können die Probesubstanz dazu nutzen, um sich über Ihre Vorgehensweise bei der Beurteilung der folgenden 6 Substanzen klar zu werden. Am Ende des Versuchs werden wir Sie zu Ihrer Vorgehensweise befragen. Das Instruktionsblatt können Sie während des ganzen Versuchs einsehen. Während des Versuchs sollen Sie sich keine Notizen machen. Bei Fragen wenden Sie sich an den Versuchsleiter.



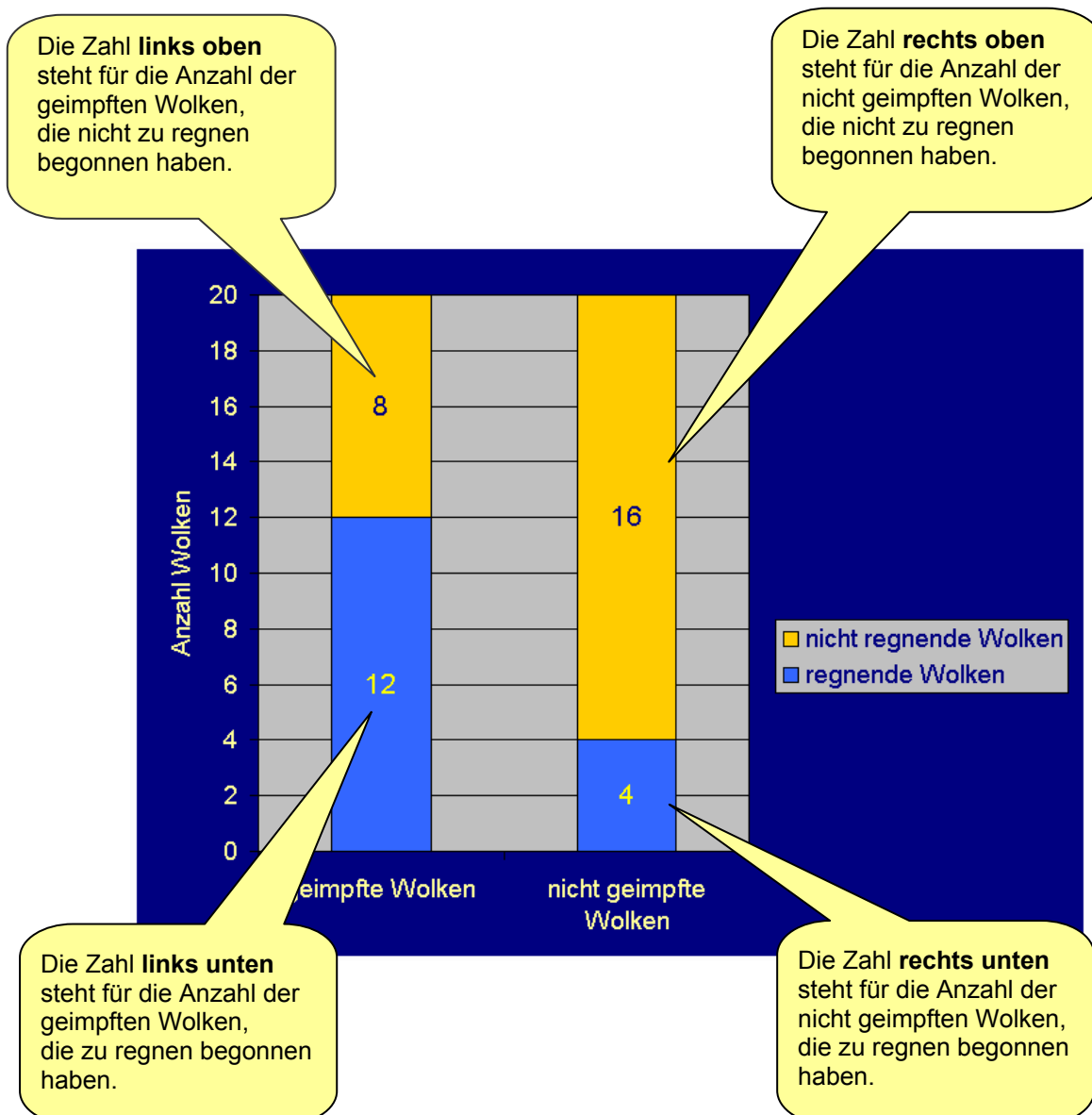
## Erläuterung zum Säulendiagramm

Wie Sie aus der Instruktion wissen, sollen Sie die Wirksamkeiten von Substanzen an Hand von Testergebnissen beurteilen, die in einem Säulendiagramm zusammengefasst sind. Dieses Säulendiagramm und seinen Inhalt wollen wir Ihnen nun kurz erläutern. Damit wollen wir sicher stellen, dass Sie den Diagramminhalt verstehen.

Wie Sie wissen, enthält unser Säulendiagramm vier verschiedene Arten von Informationen (Anzahl geimpfter Wolken, die zu regnen begonnen haben; Anzahl geimpfter Wolken, die nicht zu regnen begonnen haben; Anzahl nicht geimpfter Wolken, die zu regnen begonnen haben; Anzahl nicht geimpfter Wolken, die nicht zu regnen begonnen haben). Diese Informationen sind im Säulendiagramm räumlich unterschiedlich angeordnet.

Im Experiment sieht unser Säulendiagramm nun folgendermaßen aus: Bitte schauen Sie sich Abbildung 2 in Ruhe an und lesen Sie die Inhalte. Wenn Sie Fragen haben oder Ihnen irgend etwas unklar sein sollte, fragen Sie bitte den Versuchsleiter.

**Abbildung 2:** Säulendiagramm



## Instruktion – Bedingung *VT tbt* (Ex 2) und *ohne N tbt* (Ex 3)

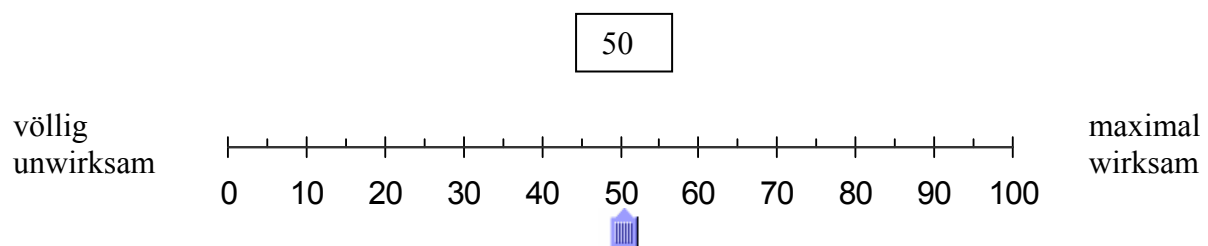
Wie Sie vielleicht wissen, kann man durch das Versprühen von Silberjodid die Regentätigkeit von Wolken beeinflussen. Das in der Luft vorhandene Wasser lagert sich an die Silberjodidkristalle an und regnet ab, sobald eine kritische Tröpfchengröße erreicht wird. Diese Technik wird in Trockengebieten genutzt, um Wolken, die eigentlich nicht abregnen würden, doch noch zum Regnen zu bringen.

Ein Chemieunternehmen arbeitet an der Entwicklung neuer Substanzen, die Wolken mittels Wolkenimpfung zum Abregnen bringen sollen. Dazu wurden mit einem Flugzeug in einem Trockengebiet sechs Testflüge unternommen, bei denen jeweils die Wirksamkeit von einer Substanz untersucht wurde. Pro Substanz wurden bei stabiler Wetterlage 20 geimpfte und 20 ungeimpfte Wolken beobachtet. Die 40 Wolken weisen untereinander keine wesentlichen Unterschiede auf und sind miteinander vergleichbar. Die meteorologischen Rahmenbedingungen sind für alle Wolken und bei jedem Einsatz einer Substanz gleich. Zu Beginn einer jeden Beobachtungsphase regnete keine der Wolken. Die Ergebnisse für jede der 40 beobachteten Wolken werden Ihnen nacheinander auf dem Computer-Bildschirm in einer sogenannten Vierfeldertafel dargeboten. Die Vierfeldertafel und ihr Inhalt werden Ihnen auf Seite 2 erläutert. Durch die Plazierung eines Kreuzes wird für jeden der 40 Fälle angezeigt, welches Ergebnis eingetroffen ist. Entsprechend den 4 Feldern der Vierfeldertafel sind vier verschiedene Ergebnisse möglich:

1. Die Wolke wurde geimpft und fing zu regnen an.
2. Die Wolke wurde geimpft und fing nicht zu regnen an.
3. Die Wolke wurde nicht geimpft und fing zu regnen an.
4. Die Wolke wurde nicht geimpft und fing nicht zu regnen an.

Ihre Aufgabe ist es nun, alle nacheinander dargebotenen Beobachtungsergebnisse so aufmerksam wie möglich zu registrieren und abschließend die Wirksamkeit einer jeden Substanz zu beurteilen. Nach Durchsicht der 40 Beobachtungsergebnisse sehen Sie dann zur Einschätzung der Wirksamkeit einer Substanz eine Skala eingeblendet, auf der Sie mit Hilfe der Maus einen Regler betätigen und Zahlenwerte eingeben können. Die Skala reicht von 0 bis 100. "0" bedeutet dabei, dass die Substanz völlig unwirksam ist und keinerlei Einfluss auf das Regenverhalten der Wolken hat, "100" bedeutet, dass diese Substanz maximal wirksam ist und man damit jede Wolke zum Abregnen bringen kann.

**Abbildung 1: Wirksamkeitsskala**



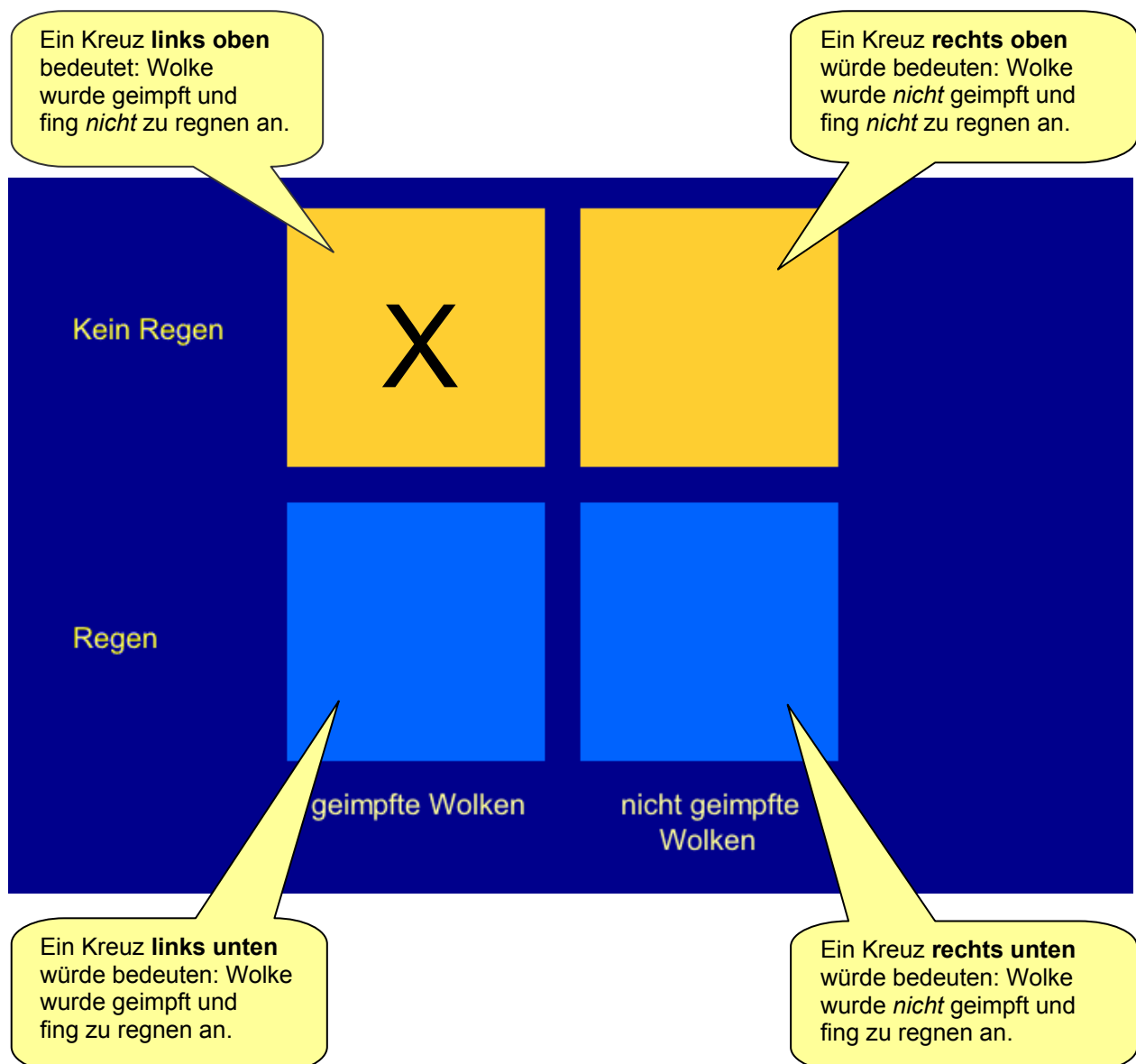
Erst wenn Sie mit Hilfe der Maus einen Zahlenwert eingetragen haben, können Sie die Weertaste anklicken um die nächste Aufgabe zu bearbeiten. Der Versuch beginnt mit einer Probesubstanz, deren Wirksamkeit Sie ebenfalls beurteilen sollen. Sie können die Probesubstanz dazu nutzen, um sich über Ihre Vorgehensweise bei der Beurteilung der folgenden 6 Substanzen klar zu werden. Am Ende des Versuchs werden wir Sie zu Ihrer Vorgehensweise befragen. Das Instruktionsblatt können Sie während des ganzen Versuchs einsehen. Während des Versuchs sollen Sie sich keine Notizen machen. Bei Fragen wenden Sie sich an den Versuchsleiter.

## Erläuterung zur Vierfeldertafel

Wie Sie aus der Instruktion wissen, sollen Sie die Wirksamkeiten von Substanzen an Hand einer so genannten Vierfeldertafel beurteilen. Diese Vierfeldertafel und ihren Inhalt wollen wir Ihnen nun kurz erläutern. Damit wollen wir sicherstellen, dass Sie den Inhalt der Tafel verstehen.

Jedes der vier möglichen Beobachtungsergebnisse (Wolke geimpft - Regen; Wolke ungeimpft - Regen; Wolke geimpft – Kein Regen; Wolke nicht geimpft - kein Regen) wird durch die Platzierung eines schwarzen Kreuzes auf einem der vier Felder angezeigt. In Abbildung 2 sehen Sie die Vierfeldertafel, wie sie in unserem Experiment verwendet wird. Beispielhaft befindet sich das Kreuz links oben, es könnte sich auch links unten, rechts oben oder rechts unten stehen. Bitte schauen Sie sich Abbildung 2 in Ruhe an und lesen Sie die Inhalte. Wenn Sie Fragen haben oder Ihnen irgendetwas unklar sein sollte, fragen Sie bitte den Versuchsleiter.

**Abbildung 2:** Vierfeldertafel



## Instruktion – Bedingung *FT sum*

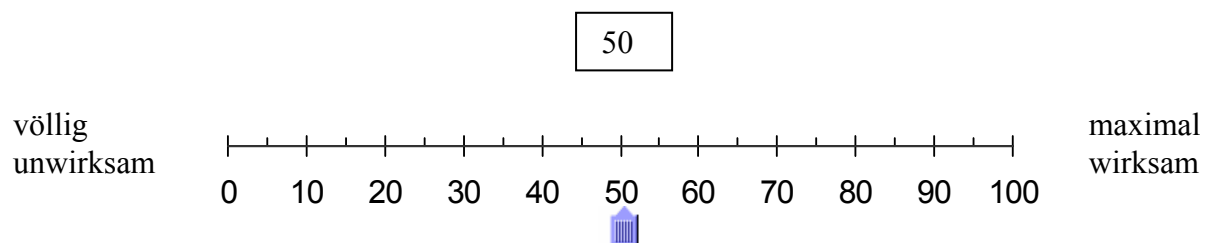
Wie Sie vielleicht wissen, kann man durch das Versprühen von Silberjodid die Regentätigkeit von Wolken beeinflussen. Das in der Luft vorhandene Wasser lagert sich an die Silberjodidkristalle an und regnet ab, sobald eine kritische Tröpfchengröße erreicht wird. Diese Technik wird in Trockengebieten genutzt, um Wolken, die eigentlich nicht abregnen würden, doch noch zum Regnen zu bringen.

Ein Chemieunternehmen arbeitet an der Entwicklung neuer Substanzen, die Wolken mittels Wolkenimpfung zum Abregnen bringen sollen. Dazu wurden mit einem Flugzeug in einem Trockengebiet sechs Testflüge unternommen, bei denen jeweils die Wirksamkeit von einer Substanz untersucht wurde. Pro Substanz wurden bei stabiler Wetterlage 20 geimpfte und 20 ungeimpfte Wolken beobachtet. Die 40 Wolken weisen untereinander keine wesentlichen Unterschiede auf und sind miteinander vergleichbar. Die meteorologischen Rahmenbedingungen sind für alle Wolken und bei jedem Einsatz einer Substanz gleich. Zu Beginn einer jeden Beobachtungsphase regnete keine der Wolken. Für jede Substanz finden Sie die Testergebnisse auf dem Computer-Bildschirm in einem kurzen Text zusammen gefasst. Die zusammen gefassten Ergebnisse enthalten vier verschiedene Arten von Informationen:

1. Anzahl der geimpften Wolken, die im Anschluss an eine Impfung zu regnen begonnen haben.
2. Anzahl der geimpften Wolken, die im Anschluss an eine Impfung nicht zu regnen begonnen haben
3. Anzahl der nicht geimpften Wolken, die spontan zu regnen begonnen haben
4. Anzahl der nicht geimpften Wolken, die nicht zu regnen begonnen haben.

Ihre Aufgabe ist es nun, an Hand der im Text zusammen gefassten Ergebnisse die Wirksamkeit der jeweiligen Substanz zu beurteilen. Bitte konzentrieren Sie sich bei der Bearbeitung dieser Aufgabe so gut wie möglich. Der Text bleibt maximal 2 Minuten und 40 Sekunden eingeblendet. Wenn Sie sich schon vorher über die Wirksamkeit der Substanz im Klaren sind, können Sie die Weitertaste drücken. Sie sehen dann zur Einschätzung der Wirksamkeit einer Substanz eine Skala eingeblendet, auf der Sie mit Hilfe der Maus Zahlenwerte eingeben können. Die Skala reicht von 0 bis 100. "0" bedeutet dabei, dass die Substanz völlig unwirksam ist und keinerlei Einfluss auf das Regenverhalten der Wolken hat, "100" bedeutet, dass diese Substanz maximal wirksam ist und man damit jede Wolke zum Abregnen bringen kann.

**Abbildung 1: Wirksamkeitsskala**



Erst wenn Sie mit Hilfe der Maus einen Zahlenwert eingetragen haben, können Sie die Weitertaste anklicken um die nächste Aufgabe zu bearbeiten. Der Versuch beginnt mit einer Probesubstanz, deren Wirksamkeit Sie ebenfalls beurteilen sollen. Sie können die Probesubstanz dazu nutzen, um sich über Ihre Vorgehensweise bei der Beurteilung der folgenden 6 Substanzen klar zu werden. Am Ende des Versuchs werden wir Sie zu Ihrer Vorgehensweise befragen. Das Instruktionsblatt können Sie während des ganzen Versuchs einsehen. Während des Versuchs sollen Sie sich keine Notizen machen. Bei Fragen wenden Sie sich an den Versuchsleiter.

## Instruktion – Bedingung *FT tbt*

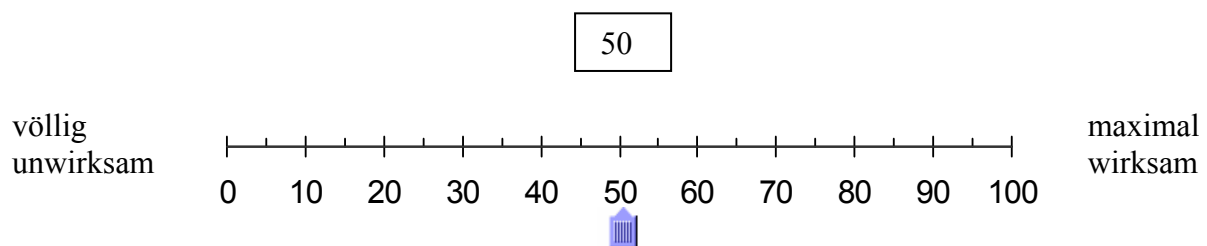
Wie Sie vielleicht wissen, kann man durch das Versprühen von Silberjodid die Regentätigkeit von Wolken beeinflussen. Das in der Luft vorhandene Wasser lagert sich an die Silberjodidkristalle an und regnet ab, sobald eine kritische Tröpfchengröße erreicht wird. Diese Technik wird in Trockengebieten genutzt, um Wolken, die eigentlich nicht abregnen würden, doch noch zum Regnen zu bringen.

Ein Chemieunternehmen arbeitet an der Entwicklung neuer Substanzen, die Wolken mittels Wolkenimpfung zum Abregnen bringen sollen. Dazu wurden mit einem Flugzeug in einem Trockengebiet sechs Testflüge unternommen, bei denen jeweils die Wirksamkeit von einer Substanz untersucht wurde. Pro Substanz wurden bei stabiler Wetterlage 20 geimpfte und 20 ungeimpfte Wolken beobachtet. Die 40 Wolken weisen untereinander keine wesentlichen Unterschiede auf und sind miteinander vergleichbar. Die meteorologischen Rahmenbedingungen sind für alle Wolken und bei jedem Einsatz einer Substanz gleich. Zu Beginn einer jeden Beobachtungsphase regnete keine der Wolken. Die Ergebnisse für jede der 40 beobachteten Wolken werden Ihnen nacheinander auf dem Computer-Bildschirm in einer Textzeile dargeboten. Es sind vier verschiedene Ergebnisse möglich:

1. Die Wolke wurde geimpft und fing zu regnen an.
2. Die Wolke wurde geimpft und fing nicht zu regnen an.
3. Die Wolke wurde nicht geimpft und fing zu regnen an.
4. Die Wolke wurde nicht geimpft und fing nicht zu regnen an.

Ihre Aufgabe ist es nun, alle nacheinander dargebotenen Beobachtungsergebnisse so aufmerksam wie möglich zu registrieren und abschließend die Wirksamkeit einer jeden Substanz zu beurteilen. Nach Durchsicht der 40 Beobachtungsergebnisse sehen Sie dann zur Einschätzung der Wirksamkeit einer Substanz eine Skala eingeblendet, auf der Sie mit Hilfe der Maus einen Regler betätigen und Zahlenwerte eingeben können. Die Skala reicht von 0 bis 100. "0" bedeutet dabei, dass die Substanz völlig unwirksam ist und keinerlei Einfluss auf das Regenverhalten der Wolken hat, "100" bedeutet, dass diese Substanz maximal wirksam ist und man damit jede Wolke zum Abregnen bringen kann.

**Abbildung 1: Wirksamkeitsskala**



Erst wenn Sie mit Hilfe der Maus einen Zahlenwert eingetragen haben, können Sie die Weitertaste anklicken um die nächste Aufgabe zu bearbeiten. Der Versuch beginnt mit einer Probesubstanz, deren Wirksamkeit Sie ebenfalls beurteilen sollen. Sie können die Probesubstanz dazu nutzen, um sich über Ihre Vorgehensweise bei der Beurteilung der folgenden 6 Substanzen klar zu werden. Am Ende des Versuchs werden wir Sie zu Ihrer Vorgehensweise befragen. Das Instruktionsblatt können Sie während des ganzen Versuchs einsehen. Während des Versuchs sollen Sie sich keine Notizen machen. Bei Fragen wenden Sie sich an den Versuchsleiter.

**Anhang A2.3:** Experimente 2 und 3: Fragen zur Vorgehensweise

## **Fragen zur Vorgehensweise**

Wie würden Sie Ihre Vorgehensweise bei der Bestimmung der Wirksamkeit einer Substanz in eigenen Worten beschreiben ?

Haben Sie versucht die Wirksamkeit einer Substanz zu berechnen? Wenn ja, bitte beschreiben Sie Ihren Rechenweg (ggf. mit Formel), oder falls Sie mehrere Rechenwege ausprobiert haben, Ihre Rechenwege (ggf. mit

## Anhang A2.4: Experiment 2: Weitere Ergebnisse

**Tabelle A3:** Experiment 2: Anzahl der Probanden, die in Abhängigkeit von Strategietyp und Versuchsbedingung die Strategie im Laufe der Aufgabenbearbeitung gewechselt haben. Zelhäufigkeiten > 5 sind **fett** markiert.

	Wechsler unklar zu ...									Wechsler ab zu ...		Wechsler $\Delta P$ zu ...	Gesamt
	ab	<u>ab</u> 2	ab, 25	ab, $\Delta P$	ab, PPC	25	$\Delta P$	$\Delta P$ , PPC	PPC	$\Delta P$	25	PPC	
<b>Versuchsbedingung</b>													
SD sum	2	0	0	2	1	0	1	1	1	1	0	1	10 18.9 %
FT sum	3	0	<b>6</b>	0	1	1	0	1	2	1	0	1	16 30.2 %
VT tbt	3	0	2	2	0	2	2	0	1	0	1	0	13 24.5 %
FT tbt	4	1	1	2	1	0	4	0	1	0	0	0	14 26.4 %
<b>Gesamt</b>	12 22.6 %	1 1.9 %	9 16,9 %	6 11.3 %	3 5.7 %	3 5.7 %	7 13.2 %	2 3.8 %	5 9.4 %	2 3.8 %	1 1.9 %	2 3.8 %	53 100 % <sup>1</sup>

<sup>1</sup> In Tabelle A3 werden 53 Probanden als Wechsler klassifiziert, in Tabelle 13 nur 45. Der Unterschied zwischen den beiden Tabellen kommt folgendermaßen zu Stande: Von den 53 Wechslern haben 2 zwischen regel-basierten Strategien gewechselt, 1 von *ab* zu 25, 3 von *unklar* zu 25 und 1 zu *ab/2*. Diese 7 Probanden wurden in Tabelle 13 anders zugeordnet, nämlich 2 x *regel-basiert*, 1 x *ab* und 4 x *unklar*. Ein weiterer Proband (53 – 7 = 46; 46 – 45 Wechsler = 1 Wechsler) wurde in Tabelle 13 ebenfalls als *unklar* klassifiziert, weil er nicht erklären konnte, wie er zu einer regel-basierten Strategie gekommen ist und somit eher rein zufällig zu einer regel-basierten Strategie wechselte. In Tabelle A3 wurde dieser Proband als Wechsler klassifiziert.

**Tabelle A4:** Experiment 2: Signifikante Faktoren und Interaktionen für die abhängigen Variablen *Bearbeitungszeit* und *Verarbeitungseffizienz* aus der dreifaktoriellen Varianzanalyse mit dem zusätzlichen Faktor *Studienausrichtung* über alle Probanden hinweg und den 2 (*Darbietungsart*) x 2 (*Visuelle Strukturierung*) Varianzanalysen für die beiden Probandengruppen mit und ohne mathematisch-naturwissenschaftliche Studienausrichtung.

Probanden- gruppe	abhängige Variable	signifikante Faktoren und Interaktionen	Trends	nicht signifikante Faktoren und Interaktionen
Gesamt	<b>Bearbeitungs- zeit</b>	Visuelle Strukturierung (p=.041); Darbietungsart (p<.001); Darbietungsart x Visuelle Strukturierung (p=.005)		Studienausrichtung (p=.174); Studienausrichtung x Visuelle Strukturierung (p=.653); Studienausrichtung x Darbietungsart (p=.414); Studienausrichtung x Darbietungsart x Visuelle Strukturierung (p=.376)
	<b>Effizienz</b>	Studienausrichtung (p=.013); Darbietungsart (p<.001); Studienausrichtung x Darbietungsart (p=.027)	Visuelle Strukturierung (p=.073); Studienausrichtung x Visuelle Strukturierung (p=.087); Studienausrichtung x Darbietungsart x Visuelle Strukturierung (p=.076)	Darbietungsart x Visuelle Strukturierung (p=.112)
Nur NW	<b>Bearbeitungs- zeit</b>	Visuelle Strukturierung (p<.001); Darbietungsart (p<.001)		Darbietungsart x Visuelle Strukturierung (p=.106)
	<b>Effizienz</b>	Visuelle Strukturierung (p=.020); Darbietungsart (p<.001); Darbietungsart x Visuelle Strukturierung (p=.024)		
Nur GW	<b>Bearbeitungs- zeit</b>	Darbietungsart (p<.001); Darbietungsart x Visuelle Strukturierung (p=.022)		Visuelle Strukturierung (p=.312)
	<b>Effizienz</b>	Darbietungsart (p=.027)		Visuelle Strukturierung (p=.951); Darbietungsart x Visuelle Strukturierung (p=.891)



## Anhang A3

### Anhang A3.1: Experiment 3: Zusammensetzung der Teilnehmer

**Tabelle A5:** Experiment 3: Zusammensetzung der Teilnehmer hinsichtlich ihrer realen oder beabsichtigten Studienfächer und sonstiger demografischer Merkmale.

Vp	Studienrichtung	Studienfach/ Studienwunsch	Alter	Geschlecht	Bedingung	Vp	Studienrichtung	Studienfach/ Studienwunsch	Alter	Geschlecht	Bedingung
3	Geist	Geschichte, Französisch	20	weiblich	ohne Notizen summarisch	63	Natur	BWL	21	weiblich	ohne Notizen summarisch
4	Geist	Geschichte, Französisch	20	männlich	ohne Notizen einzelfallweise	64	Natur	Biologie	21	männlich	ohne Notizen einzelfallweise
5	Geist	Pädagogik	31	weiblich	Notizen frei einzelfallweise	65	Natur	BWL	24	weiblich	Notizen frei einzelfallweise
6	Geist	Pädagogik	27	weiblich	Notizen strukturiert einzelfallweise	66	Natur	Geowissenschaften	29	männlich	Notizen strukturiert einzelfallweise
9	Geist	Sozialpädagogik	27	weiblich	ohne Notizen summarisch	69	Natur	VWL, Politik, Spanisch	22	weiblich	ohne Notizen summarisch
10	Geist	Pädagogik	25	weiblich	ohne Notizen einzelfallweise	70	Natur	VWL	19	weiblich	ohne Notizen einzelfallweise
11	Geist	Ev. Theologie	24	weiblich	Notizen frei einzelfallweise	71	Natur	BWL	19	weiblich	Notizen frei einzelfallweise
12	Geist	Germanistik, Englisch	26	weiblich	Notizen strukturiert einzelfallweise	72	Natur	VWL	23	weiblich	Notizen strukturiert einzelfallweise
15	Geist	Französisch, Linguistik	24	weiblich	ohne Notizen summarisch	75	Natur	Bioinformatik	22	weiblich	ohne Notizen summarisch
16	Geist	Jura	26	weiblich	ohne Notizen einzelfallweise	76	Natur	Psychologie	20	weiblich	ohne Notizen einzelfallweise
17	Geist	Musikwissenschaften	29	männlich	Notizen frei einzelfallweise	77	Natur	BWL	28	weiblich	Notizen frei einzelfallweise
18	Geist	Ev. Theologie	26	weiblich	Notizen strukturiert einzelfallweise	78	Natur	Psychologie	24	weiblich	Notizen strukturiert einzelfallweise
21	Geist	Theologie, Soziologie	23	männlich	ohne Notizen summarisch	81	Natur	Medizin	22	männlich	ohne Notizen summarisch
22	Geist	Germanistik, Geograf	25	weiblich	ohne Notizen einzelfallweise	82	Natur	Biologie	23	weiblich	ohne Notizen einzelfallweise
23	Geist	EKW	31	weiblich	Notizen frei einzelfallweise	83	Natur	Biochemie	35	weiblich	Notizen frei einzelfallweise
24	Geist	Germanistik	22	weiblich	Notizen strukturiert einzelfallweise	84	Natur	Psychologie	21	weiblich	Notizen strukturiert einzelfallweise
27	Geist	Kunstgeschichte, EKW	21	weiblich	ohne Notizen summarisch	87	Natur	BWL	25	weiblich	ohne Notizen summarisch
28	Geist	Kunstgeschichte	24	weiblich	ohne Notizen einzelfallweise	88	Natur	Biologie	21	männlich	ohne Notizen einzelfallweise
29	Geist	Jura	24	männlich	Notizen frei einzelfallweise	89	Natur	Informatik	19	männlich	Notizen frei einzelfallweise
30	Geist	Philosophie	19	männlich	Notizen strukturiert einzelfallweise	90	Natur	Medizin	24	männlich	Notizen strukturiert einzelfallweise
33	Geist	Englisch, Spanisch	29	weiblich	ohne Notizen summarisch	93	Natur	Biologie	24	männlich	ohne Notizen summarisch
34	Geist	Romanistik, EKW	32	weiblich	ohne Notizen einzelfallweise	94	Natur	Physik	24	männlich	ohne Notizen einzelfallweise
35	Geist	Jura	24	weiblich	Notizen frei einzelfallweise	95	Natur	Medizin	20	weiblich	Notizen frei einzelfallweise
36	Geist	Philosophie	25	weiblich	Notizen strukturiert einzelfallweise	96	Natur	Mathematik	22	männlich	Notizen strukturiert einzelfallweise
39	Geist	Sport	19	weiblich	ohne Notizen summarisch	99	Natur	Psychologie	22	weiblich	ohne Notizen summarisch
40	Geist	Germanistik, Japanologie	26	männlich	ohne Notizen einzelfallweise	100	Natur	Medizin	27	weiblich	ohne Notizen einzelfallweise
41	Geist	Medienwissenschaft	30	weiblich	Notizen frei einzelfallweise	101	Natur	Physik	20	weiblich	Notizen frei einzelfallweise
42	Geist	Jura	26	männlich	Notizen strukturiert einzelfallweise	102	Natur	Bioinformatik	23	männlich	Notizen strukturiert einzelfallweise
45	Natur	Patholinguistik	21	männlich	ohne Notizen summarisch	105	Natur	Geografie	21	männlich	ohne Notizen summarisch
46	Geist	Kunstgeschichte	22	weiblich	ohne Notizen einzelfallweise	106	Natur	Mathe, Philosophie,	20	männlich	ohne Notizen einzelfallweise
47	Geist	Psychologie	41	weiblich	Notizen frei einzelfallweise	107	Natur	Psychologie	19	weiblich	Notizen frei einzelfallweise
48	Geist	Sport, Englisch	21	weiblich	Notizen strukturiert einzelfallweise	108	Natur	Pharmazie-Doktorand	27	männlich	Notizen strukturiert einzelfallweise
51	Geist	Theologie	25	männlich	ohne Notizen summarisch	111	Geist	Theater- und Medienwissenschaften	20	weiblich	ohne Notizen summarisch
52	Geist	Jura	19	weiblich	ohne Notizen einzelfallweise	112	Natur	Physik	25	männlich	ohne Notizen einzelfallweise
53	Geist	EKW	30	männlich	Notizen frei einzelfallweise	113	Natur	Bioinformatik	26	weiblich	Notizen frei einzelfallweise
54	Geist	EKW, Philosophie	21	weiblich	Notizen strukturiert einzelfallweise	114	Natur	Psychologie	21	männlich	Notizen strukturiert einzelfallweise
57	Geist	Rhetorik, Soziologie	20	weiblich	ohne Notizen summarisch	117	Natur	BWL	28	männlich	ohne Notizen summarisch
58	Geist	Soziologie	26	weiblich	ohne Notizen einzelfallweise	118	Natur	Biologie	23	Weiblich	ohne Notizen einzelfallweise
59	Geist	Komparatistik	20	weiblich	Notizen frei einzelfallweise	119	Natur	Biologie, Germanistik	27	Weiblich	Notizen frei einzelfallweise
60	Geist	Jura	24	weiblich	Notizen strukturiert einzelfallweise	120	Natur	Geologie	29	männlich	Notizen strukturiert einzelfallweise

## Anhang A3.2: Experiment 3: Instruktionen

### Instruktion – Bedingung *N* strukturiert

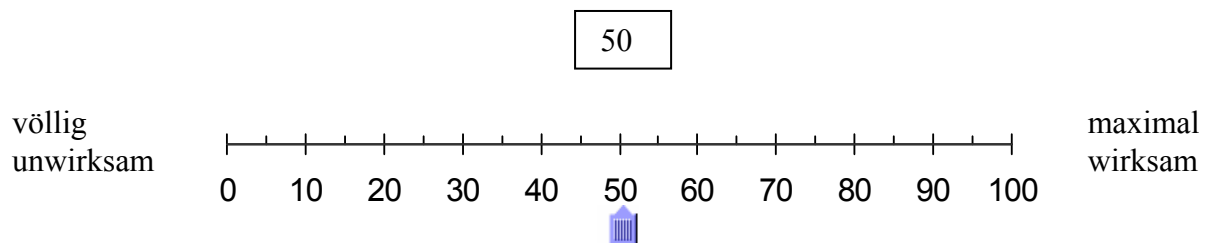
Wie Sie vielleicht wissen, kann man durch das Versprühen von Silberjodid die Regentätigkeit von Wolken beeinflussen. Das in der Luft vorhandene Wasser lagert sich an die Silberjodidkristalle an und regnet ab, sobald eine kritische Tröpfchengröße erreicht wird. Diese Technik wird in Trockengebieten genutzt, um Wolken, die eigentlich nicht abregnen würden, doch noch zum Regnen zu bringen.

Ein Chemieunternehmen arbeitet an der Entwicklung neuer Substanzen, die Wolken mittels Wolkenimpfung zum Abregnen bringen sollen. Dazu wurden mit einem Flugzeug in einem Trockengebiet sechs Testflüge unternommen, bei denen jeweils die Wirksamkeit von einer Substanz untersucht wurde. Pro Substanz wurden bei stabiler Wetterlage 20 geimpfte und 20 ungeimpfte Wolken beobachtet. Die 40 Wolken weisen untereinander keine wesentlichen Unterschiede auf und sind miteinander vergleichbar. Die meteorologischen Rahmenbedingungen sind für alle Wolken und bei jedem Einsatz einer Substanz gleich. Zu Beginn einer jeden Beobachtungsphase regnete keine der Wolken. Die Ergebnisse für jede der 40 beobachteten Wolken werden Ihnen nacheinander auf dem Computer-Bildschirm in einer sogenannten Vierfeldertafel dargeboten. Die Vierfeldertafel und ihr Inhalt werden Ihnen auf Seite 2 erläutert. Durch die Plazierung eines Kreuzes wird für jeden der 40 Fälle angezeigt, welches Ergebnis eingetroffen ist. Entsprechend den 4 Feldern der Vierfeldertafel sind vier verschiedene Ergebnisse möglich:

1. Die Wolke wurde geimpft und fing zu regnen an.
2. Die Wolke wurde geimpft und fing nicht zu regnen an.
3. Die Wolke wurde nicht geimpft und fing zu regnen an.
4. Die Wolke wurde nicht geimpft und fing nicht zu regnen an.

Ihre Aufgabe ist es nun, alle nacheinander dargebotenen Beobachtungsergebnisse so aufmerksam wie möglich zu registrieren und abschließend die Wirksamkeit einer jeden Substanz zu beurteilen. Nach Durchsicht der 40 Beobachtungsergebnisse sehen Sie dann zur Einschätzung der Wirksamkeit einer Substanz eine Skala eingeblendet, auf der Sie mit Hilfe der Maus einen Regler betätigen und Zahlenwerte eingeben können. Die Skala reicht von 0 bis 100. "0" bedeutet dabei, dass die Substanz völlig unwirksam ist und keinerlei Einfluss auf das Regenverhalten der Wolken hat, "100" bedeutet, dass diese Substanz maximal wirksam ist und man damit jede Wolke zum Abregnen bringen kann.

**Abbildung 1: Wirksamkeitsskala**



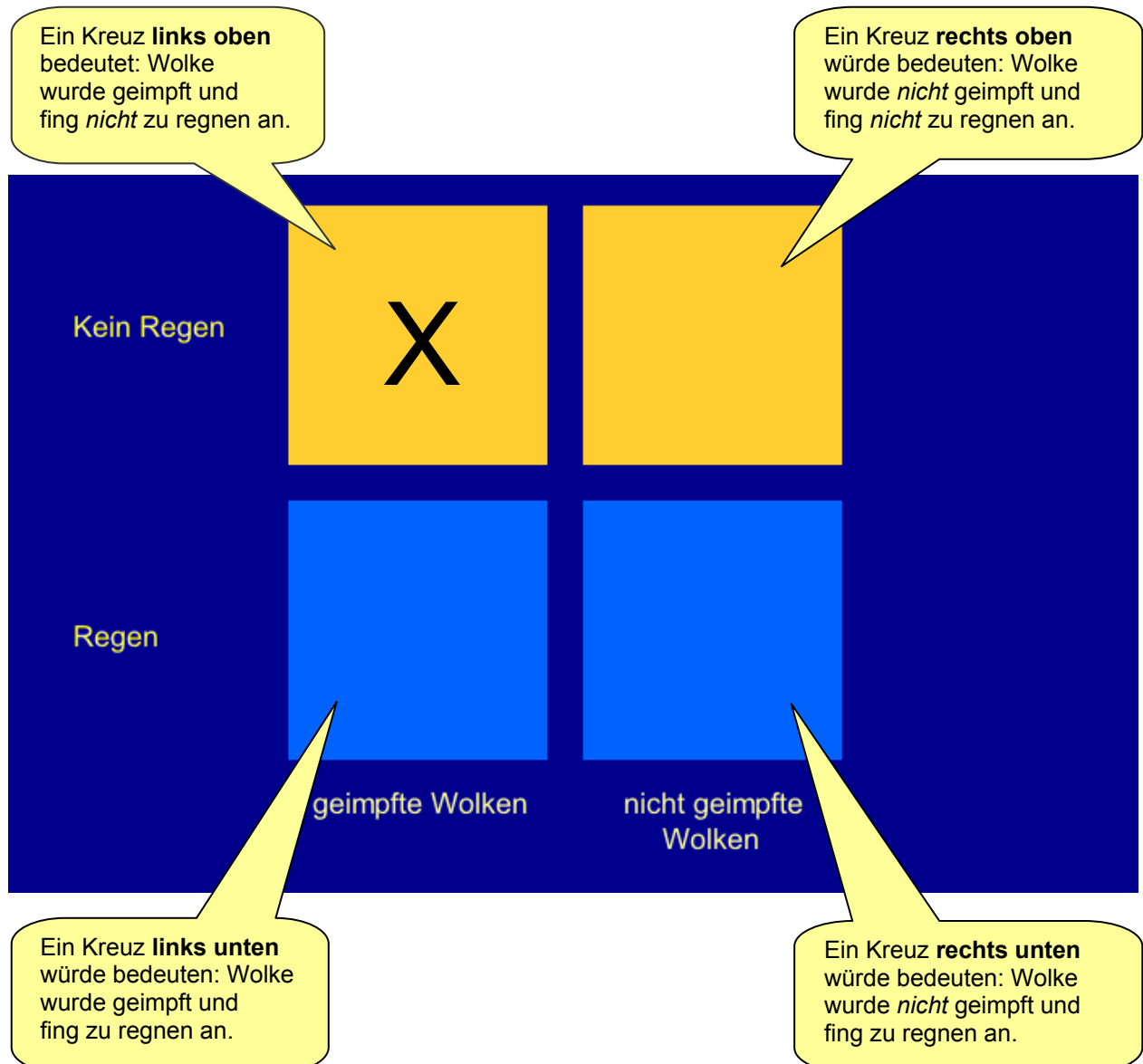
Erst wenn Sie mit Hilfe der Maus einen Zahlenwert eingetragen haben, können Sie die Weertaste anklicken um die nächste Aufgabe zu bearbeiten. Der Versuch beginnt mit einer Probesubstanz, deren Wirksamkeit Sie ebenfalls beurteilen sollen. Sie können die Probesubstanz dazu nutzen, um sich über Ihre Vorgehensweise bei der Beurteilung der folgenden 6 Substanzen klar zu werden. Am Ende des Versuchs werden wir Sie zu Ihrer Vorgehensweise befragen. Das Instruktionsblatt können Sie während des ganzen Versuchs einsehen. Während des Versuchs sollen Sie sich auf den beiliegenden Blättern Notizen machen. Eine Anleitung zum Notizen machen erhalten Sie auf Seite 3. Bei Fragen wenden Sie sich an den Versuchsleiter.

## Erläuterung zur Vierfeldertafel

Wie Sie aus der Instruktion wissen, sollen Sie die Wirksamkeiten von Substanzen an Hand einer sogenannten Vierfeldertafel beurteilen. Diese Vierfeldertafel und ihren Inhalt wollen wir Ihnen nun kurz erläutern. Damit wollen wir sicher stellen, daß Sie den Inhalt der Tafel verstehen.

Jedes der vier möglichen Beobachtungsergebnisse (Wolke geimpft - Regen; Wolke ungeimpft - Regen; Wolke geimpft – Kein Regen; Wolke nicht geimpft - kein Regen) wird durch die Plazierung eines schwarzen Kreuzes auf einem der vier Felder angezeigt. In Abbildung 2 sehen Sie die Vierfeldertafel, wie sie in unserem Experiment verwendet wird. Beispielhaft befindet sich das Kreuz links oben, es könnte sich auch links unten, rechts oben oder rechts unten stehen. Bitte schauen Sie sich Abbildung 2 in Ruhe an und lesen Sie die Inhalte. Wenn Sie Fragen haben oder Ihnen irgend etwas unklar sein sollte, fragen Sie bitte den Versuchsleiter.

**Abbildung 2:** Vierfeldertafel

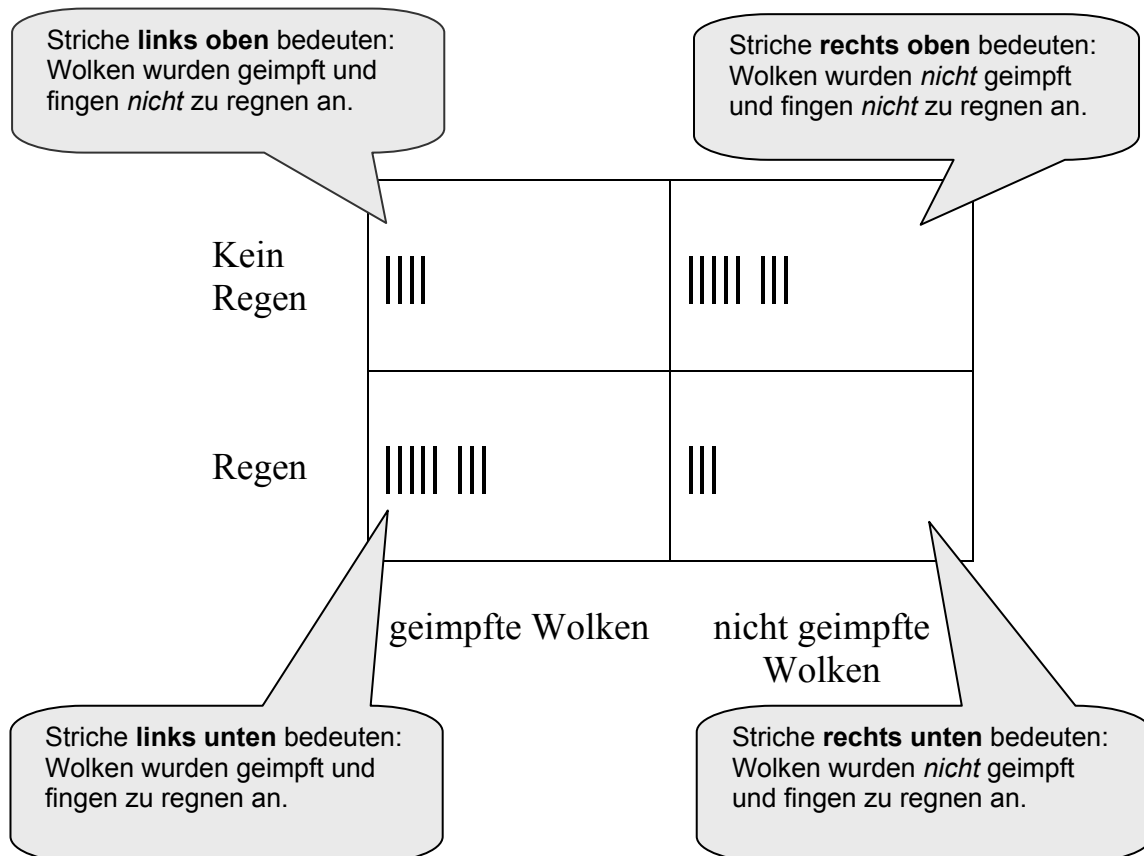


## Anleitung zum Notizen machen

Wie Sie aus der Instruktion wissen, sollen Sie sich während des Testens der Wirksamkeit einzelner Substanzen Notizen machen. Für jede Substanz steht Ihnen dazu eine Vierfeldertafel zur Verfügung. Diese Vierfeldertafel ist genauso aufgebaut wie die Vierfeldertafel in Abbildung 2 auf Seite 2. Ihre Aufgabe ist es, die Vierfeldertafel mit Strichen auszufüllen. Ein Strich steht für ein Beobachtungsergebnis. Die ausgefüllte Vierfeldertafel soll Sie bei der Beurteilung der Wirksamkeit einer Substanz unterstützen.

In Abbildung 3 sehen Sie, wie eine ausgefüllte Vierfeldertafel nach z.B. 23 von insgesamt 40 Beobachtungsergebnissen aussehen könnte. Schauen Sie sich Abbildung 3 solange an, bis Sie sie verstanden haben. Wenn Sie Fragen haben oder Ihnen irgend etwas unklar sein sollte, fragen Sie bitte den Versuchsleiter.

**Abbildung 3:** Vierfeldertafel nach 23 beobachteten Wolken



## Instruktion – Bedingung *N frei*

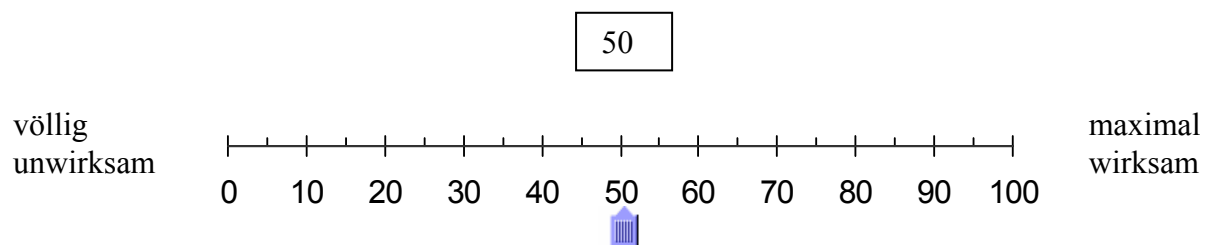
Wie Sie vielleicht wissen, kann man durch das Versprühen von Silberjodid die Regentätigkeit von Wolken beeinflussen. Das in der Luft vorhandene Wasser lagert sich an die Silberjodidkristalle an und regnet ab, sobald eine kritische Tröpfchengröße erreicht wird. Diese Technik wird in Trockengebieten genutzt, um Wolken, die eigentlich nicht abregnen würden, doch noch zum Regnen zu bringen.

Ein Chemieunternehmen arbeitet an der Entwicklung neuer Substanzen, die Wolken mittels Wolkenimpfung zum Abregnen bringen sollen. Dazu wurden mit einem Flugzeug in einem Trockengebiet sechs Testflüge unternommen, bei denen jeweils die Wirksamkeit von einer Substanz untersucht wurde. Pro Substanz wurden bei stabiler Wetterlage 20 geimpfte und 20 ungeimpfte Wolken beobachtet. Die 40 Wolken weisen untereinander keine wesentlichen Unterschiede auf und sind miteinander vergleichbar. Die meteorologischen Rahmenbedingungen sind für alle Wolken und bei jedem Einsatz einer Substanz gleich. Zu Beginn einer jeden Beobachtungsphase regnete keine der Wolken. Die Ergebnisse für jede der 40 beobachteten Wolken werden Ihnen nacheinander auf dem Computer-Bildschirm in einer sogenannten Vierfeldertafel dargeboten. Die Vierfeldertafel und ihr Inhalt werden Ihnen auf Seite 2 erläutert. Durch die Plazierung eines Kreuzes wird für jeden der 40 Fälle angezeigt, welches Ergebnis eingetroffen ist. Entsprechend den 4 Feldern der Vierfeldertafel sind vier verschiedene Ergebnisse möglich:

1. Die Wolke wurde geimpft und fing zu regnen an.
2. Die Wolke wurde geimpft und fing nicht zu regnen an.
3. Die Wolke wurde nicht geimpft und fing zu regnen an.
4. Die Wolke wurde nicht geimpft und fing nicht zu regnen an.

Ihre Aufgabe ist es nun, alle nacheinander dargebotenen Beobachtungsergebnisse so aufmerksam wie möglich zu registrieren und abschließend die Wirksamkeit einer jeden Substanz zu beurteilen. Nach Durchsicht der 40 Beobachtungsergebnisse sehen Sie dann zur Einschätzung der Wirksamkeit einer Substanz eine Skala eingeblendet, auf der Sie mit Hilfe der Maus einen Regler betätigen und Zahlenwerte eingeben können. Die Skala reicht von 0 bis 100. "0" bedeutet dabei, dass die Substanz völlig unwirksam ist und keinerlei Einfluss auf das Regenverhalten der Wolken hat, "100" bedeutet, dass diese Substanz maximal wirksam ist und man damit jede Wolke zum Abregnen bringen kann.

**Abbildung 1: Wirksamkeitsskala**



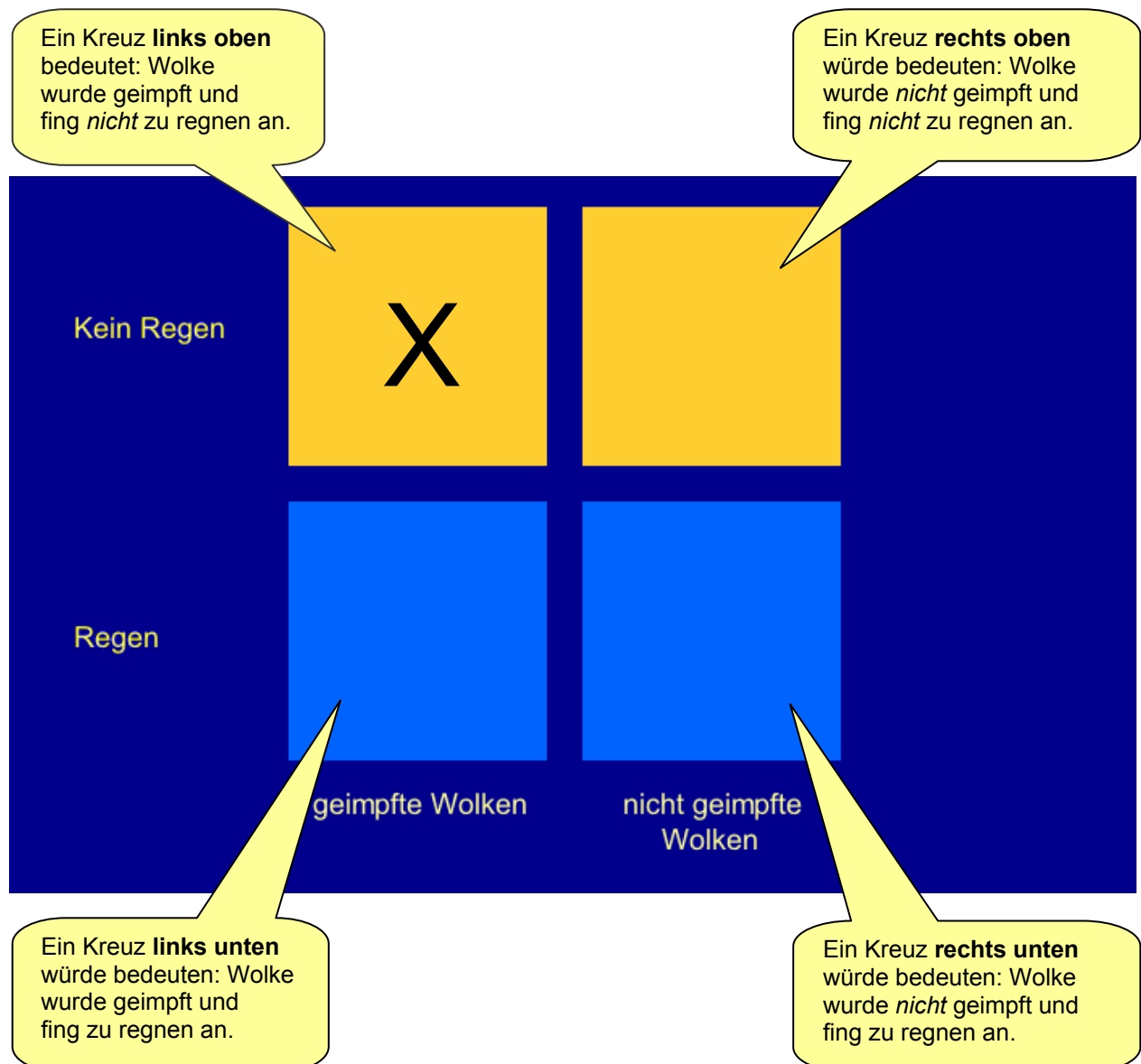
Erst wenn Sie mit Hilfe der Maus einen Zahlenwert eingetragen haben, können Sie die Weitertaste anklicken um die nächste Aufgabe zu bearbeiten. Der Versuch beginnt mit einer Probesubstanz, deren Wirksamkeit Sie ebenfalls beurteilen sollen. Sie können die Probesubstanz dazu nutzen, um sich über Ihre Vorgehensweise bei der Beurteilung der folgenden 6 Substanzen klar zu werden. Am Ende des Versuchs werden wir Sie zu Ihrer Vorgehensweise befragen. Das Instruktionsblatt können Sie während des ganzen Versuchs einsehen. Während des Versuchs können Sie sich auf beiliegenden Blättern Notizen machen. Bei Fragen wenden Sie sich an den Versuchsleiter.

## Erläuterung zur Vierfeldertafel

Wie Sie aus der Instruktion wissen, sollen Sie die Wirksamkeiten von Substanzen an Hand einer so genannten Vierfeldertafel beurteilen. Diese Vierfeldertafel und ihren Inhalt wollen wir Ihnen nun kurz erläutern. Damit wollen wir sicherstellen, dass Sie den Inhalt der Tafel verstehen.

Jedes der vier möglichen Beobachtungsergebnisse (Wolke geimpft - Regen; Wolke ungeimpft - Regen; Wolke geimpft - Kein Regen; Wolke nicht geimpft - kein Regen) wird durch die Platzierung eines schwarzen Kreuzes auf einem der vier Felder angezeigt. In Abbildung 2 sehen Sie die Vierfeldertafel, wie sie in unserem Experiment verwendet wird. Beispielhaft befindet sich das Kreuz links oben, es könnte sich auch links unten, rechts oben oder rechts unten stehen. Bitte schauen Sie sich Abbildung 2 in Ruhe an und lesen Sie die Inhalte. Wenn Sie Fragen haben oder Ihnen irgendetwas unklar sein sollte, fragen Sie bitte den Versuchsleiter.

**Abbildung 2:** Vierfeldertafel



### Anhang A3.3: Experiment 3: Weitere Ergebnisse

**Tabelle A6:** Experiment 3: Anzahl der Probanden, die in Abhängigkeit von Strategietyp und Versuchsbedingung die Strategie im Laufe der Aufgabenbearbeitung gewechselt haben.

Versuchsbedingung	Wechsler unklar zu ...							Wechsler ab zu ...			Wechsler AP zu ...		Sonstige	Gesamt			
	ab 25	ab, ΔP	ab, PPC	ab, PPC	ab, ΔP, PPC	ab/2 PPC Zufall	25	ΔP	ΔP, (ER+ΔP)/2	ΔP, PPC	25	ΔP	PPC		PPC	ER * (1-BR)	(c+d) / c*a, 25
ohne N sum	2	0	2	1	0	0	0	1	0	1	1	0	1	0	0	10	
																18.9 %	
N strukturiert	3	3	0	0	1	0	1	1	0	0	1	1	1	0	3	1	1
																17	
N frei	2	1	2	2	0	1	0	1	1	1	1	0	0	1	0	0	0
																13	
ohne N tbt	3	2	2	0	0	0	2	2	0	0	1	1	0	0	0	0	0
																13	
<b>Gesamt</b>	10	6	6	3	1	1	3	5	1	2	4	2	2	1	4	1	1
																53 <sup>2</sup>	
																100 %	

<sup>2</sup> In Tabelle A6 werden 53 Probanden als Wechsler klassifiziert, in Tabelle 26 nur 41. Die Differenz von 12 Probanden kommt hauptsächlich durch 5 Probanden zustande, die zwischen zwei regel-basierten Strategien gewechselt haben und in Tabelle 26 als *regel-basiert* eingestuft worden waren und durch 6 Probanden, die lediglich bei gleichen Erfolgs- und Basisraten (Kategorie 25 - Wechsler) von einer regel-basierten Strategie Gebrauch gemacht haben und in Tabelle 26 nicht als *Wechsler regel-basiert* klassifiziert worden waren. Ein weiterer Proband ist bei zwei Aufgaben auf ein Ergebnis gekommen, dass mit den Vorhersagen der PPC-Theorie übereinstimmte, doch konnte er weder erklären, wie dieses Ergebnis zustande gekommen war, noch gab er an, dass er etwas berechnet hätte. Dieser Proband schien also eher zufällig auf eine regel-konforme Lösung gestoßen zu sein und wurde deshalb in Tabelle 26 als *unklar* eingestuft.

**Tabelle A7:** Experiment 3: Signifikante Faktoren und Interaktionen für die abhängigen Variablen *Bearbeitungszeit* und *Verarbeitungseffizienz* aus der zweifaktoriellen Varianzanalyse mit den Faktoren *Studienausrichtung* und *Notizen anfertigen* über alle Probanden hinweg und den univariaten Varianzanalysen mit dem vierstufigen Faktor *Notizen anfertigen* für die beiden Probandengruppen mit und ohne mathematisch-naturwissenschaftliche Studienausrichtung.

Probanden- gruppe	Abhängige Variable	signifikante Faktoren und Interaktionen	nicht signifikante Faktoren und Interaktionen
<b>Gesamt</b>	Bearbeitungszeit	Notizen anfertigen ( $p < .001$ )	Studienausrichtung ( $p = .120$ ) Studienausrichtung x Notizen anfertigen ( $p = .251$ )
	Effizienz	Studienausrichtung ( $p = .004$ ) Notizen anfertigen ( $p < .001$ ) Studienausrichtung x Notizen anfertigen ( $p < .001$ )	
<b>Nur NW</b>	Bearbeitungszeit	Notizen anfertigen ( $p < .001$ )	
	Effizienz	Notizen anfertigen ( $p < .001$ )	
<b>Nur GW</b>	Bearbeitungszeit	Notizen anfertigen ( $p < .001$ )	
	Effizienz	Notizen anfertigen ( $p = .035$ )	

**Tabelle A8:** Experiment 3: Signifikante Faktoren und Interaktionen für die abhängigen Variablen *Bearbeitungszeit* und *Verarbeitungseffizienz* aus der zweifaktoriellen Varianzanalyse mit den Faktoren *numerisches Denken (hoch vs. gering)* und *Notizen anfertigen* über alle Probanden hinweg und den univariaten Varianzanalysen mit dem vierstufigen Faktor *Notizen anfertigen* für die beiden Probandengruppen mit hohen oder geringen Werten im numerischen Denken.

Probanden- gruppe	Abhängige Variable	Signifikante Faktoren und Interaktionen	nicht signifikante Faktoren und Interaktionen
<b>Gesamt</b>	Bearbeitungszeit	Numerisches Denken ( $p < .001$ ) Notizen anfertigen ( $p < .001$ )	numerisches Denken x Notizen anfertigen ( $p = .815$ )
	Effizienz	Numerisches Denken ( $p = .022$ ) Notizen anfertigen ( $p < .001$ ) Numerisches Denken x Notizen anfertigen ( $p < .034$ )	
<b>Nur ND hoch</b>	Bearbeitungszeit	Notizen anfertigen ( $p < .001$ )	
	Effizienz	Notizen anfertigen ( $p < .001$ )	
<b>Nur ND gering</b>	Bearbeitungszeit	Notizen anfertigen ( $p < .001$ )	
	Effizienz	Notizen anfertigen ( $p = .004$ )	



## **Anhang A4:** Items aus dem Berliner Intelligenzstruktur-Test (BIS-Test, Jäger, Süß & Beauducel, 1997) zur Erfassung der Kovariate *numerisches Denken*.

### **Anhang A4.0:** Kurzbeschreibung der ausgewählten Items aus dem BIS-Test.

Neun aus dem BIS-Test ausgewählte Items, die den Probanden im Anschluss an die Experimente 2 und 3 zur Bearbeitung vorgelegt wurden. Die Reihenfolge entspricht der Abfolge im Test:

**SI - Sieben-Teilbar (BN)**

In Spalten mit zweistelligen Zahlen sind alle durch sieben teilbaren Zahlen zu markieren.

**RD - Rechnerisches Denken (KN)**

Textaufgaben der üblichen Art sind zu lösen.

**ZZ - Zweistellige Zahlen (MN)**

Eine Reihe zweistelliger Zahlen ist einzuprägen und unmittelbar danach in beliebiger Reihenfolge zu reproduzieren.

**TL - Tabellen-Lesen (KN)**

Gefordert wird die Auswertung von Informationen, die in Statistiken (Tabellen und Verlaufskurven) enthalten sind; 'multiple-choice' - Lösungen.

**ZN - Zahlenreihen (KN)**

Nach bestimmten Regeln aufgebaute Zahlenreihen sind um ein weiteres Glied zu ergänzen.

**XG - X-Größer (BN)**

In Zahlenreihen sind alle Zahlen zu markieren, die um X-größer sind als die unmittelbar vorausgegangene Zahl.

**SC - Schätzen (KN)**

Komplexe Rechenaufgaben, deren Ergebnisse geschätzt, bzw. durch einfache rechnerische Überlegungen gefunden werden können.

**RZ - Rechen-Zeichen (BN)**

In vorgegebenen Gleichungen stehen anstelle von Plus- oder Minuszeichen leere Kästchen. Die richtigen Rechenzeichen sind einzutragen.

**BR - Buchstabenreihen (KN)**

Nach bestimmten Regeln aufgebaute Buchstabenreihen sind um zwei Buchstaben fortzusetzen.

**Anhang A4.1:** BIS-Item SI - Sieben-Teilbar (BN)

Streichen Sie alle Zahlen durch, die **durch 7 teilbar** sind!

Arbeiten Sie in Pfeilrichtung und so schnell Sie können!

Beispiel

45	<del>28</del>	53	51	63	12	29
79	46	17	70	23	24	11

Streichen Sie alle Zahlen durch, die **durch 7 teilbar** sind!

Arbeiten Sie in Pfeilrichtung und so schnell Sie können!

11	77	51	30	16	49	72
18	62	81	35	15	44	57
21	80	47	73	79	28	68
56	37	20	60	14	34	46
80	69	42	31	40	27	63
23	56	21	12	75	70	76
28	25	36	81	45	11	35
52	77	61	17	33	59	19
59	65	42	11	40	56	13
28	50	73	66	34	58	63
38	44	56	42	24	54	41
78	49	32	74	43	19	14
23	35	45	61	17	21	39

**Anhang A4.2:** BIS-Item RD - Rechnerisches Denken (KN)

Lösen Sie folgende Aufgaben

**Beispiel**

Ein Lehrling geht mit einem Betrag von 7,30 Euro in ein Lokal.  
Dort kostet ein Glas Bier 2,50 Euro.

Wie viel Gläser Bier kann er bezahlen und wie viel Geld behält er übrig ?

**Lösung**

Er kann \_\_\_\_\_ Gläser Bier bezahlen und  
behält \_\_\_\_\_ Euro übrig.

Lösen Sie folgende Aufgaben.

Für Nebenrechnungen bitte **dieses Blatt** benutzen!

**Aufgabe 1:**

Ein Arbeiter verdient 15,20 Euro pro Stunde. Wie viel verdient er, wenn er 5 Stunden arbeitet?

Antwort: Er verdient \_\_\_\_\_ Euro.

**Aufgabe 2:**

Ein Tunnel soll 1020 m lang werden. Er wird auf beiden Seiten gleichzeitig gebaut. Auf der einen Seite werden täglich 24 m, auf der anderen Seite täglich 27 m fertig gestellt. In wie viel Tagen ist der Tunnel fertig?

Antwort: Der Tunnel ist in \_\_\_\_\_ Tagen fertig.

**Aufgabe 3:**

Ein Kaufmann kauft für 1500 Euro Zucker. Er verkauft ihn für 1900 Euro. An jedem Sack verdient er 50 Euro. Wie viel Säcke hatte er?

Antwort: Er hatte \_\_\_\_\_ Säcke.

**Aufgabe 4:**

In einem zehnstöckigen Hochhaus gibt es zwei Fahrstühle. Fahrstuhl A braucht für die Fahrt von Stockwerk zu Stockwerk 7 Sekunden und hält auf jeder Etage 4 Sekunden. Fahrstuhl B braucht für die Fahrt von Stockwerk zu Stockwerk nur 6 Sekunden, hält aber auf jeder Etage 5 Sekunden.

Beide Fahrstühle starten gleichzeitig im Erdgeschoß.

Welcher Fahrstuhl kommt zuerst im 10.Stockwerk an?

Antwort: Fahrstuhl \_\_\_\_\_ .

Um wie viele Sekunden kommt der erste Fahrstuhl eher im 10. Stockwerk an?

Antwort: Um \_\_\_\_\_ Sekunden.

**Aufgabe 5:**

Wie viel muß man zur Zahl 12 hinzu zählen, damit die Summe im gleichen Verhältnis zu 15 steht wie 30 zu 25?

Antwort: Man muß \_\_\_\_\_ hinzu zählen, damit das gleiche Verhältnis hergestellt ist.

**Anhang A4.3: BIS-Item ZZ - Zweistellige Zahlen (MN)**

Prägen Sie sich diese **zweistelligen Zahlen** gut ein !

Sie haben 1 Minute Zeit!

<b>61</b>	<b>57</b>	<b>98</b>	<b>42</b>	<b>29</b>	<b>53</b>	<b>47</b>	<b>86</b>
<b>73</b>	<b>95</b>	<b>26</b>	<b>34</b>	<b>91</b>	<b>65</b>	<b>38</b>	<b>82</b>



**Anhang A4.4:** BIS-Item TL - Tabellen-Lesen (KN)

Auf der folgenden Seite finden Sie eine Tabelle. Die Fragen darunter können Sie aufgrund der Tabelle beantworten.

Für jede Frage sind 5 Antworten zur Auswahl gegeben. Den Buchstaben vor der richtigen Antwort streichen Sie bitte durch. Nur eine Antwort ist richtig.



Die unten stehenden Fragen sind an Hand der Tabelle zu beantworten.

### Höchst- und Tiefsttemperaturen und Jahresniederschläge

Jahr	Ort X			Ort Y			Ort Z		
	Temperatur		Jahresniederschläge	Temperatur		Jahresniederschläge	Temperatur		Jahresniederschläge
	Hoch	Tief		Hoch	Tief		Hoch	Tief	
1955	+ 39	- 14	91 cm	+ 37	- 10	50 cm	+ 32	- 14	88 cm
1956	+ 36	- 12	60 cm	+ 38	- 11	48 cm	+ 38	- 11	55 cm
1957	+ 37	- 11	65 cm	+ 40	- 12	63 cm	+ 37	- 10	75 cm
1958	+ 33	- 19	90 cm	+ 36	- 14	83 cm	+ 36	- 13	80 cm
1959	+ 38	- 16	53 cm	+ 34	- 13	58 cm	+ 38	- 12	76 cm

- Wie groß ist der Unterschied zwischen der höchsten und der niedrigsten beobachteten Temperatur?  
 a) 24                                      b) 52                                      c) 40                                      d) 56  
 e) keine der genannten Temperaturen
- In welchem Jahr hat es in allen drei Orten zusammen die höchste Niederschlagsmenge gegeben?  
 a) 1955                                      b) 1959                                      c) 1957                                      d) 1956  
 e) 1958
- Wie hoch war die durchschnittliche jährliche Höchsttemperatur in Ort Z?  
 a) 37,0                                      b) 36,5                                      c) 36,2                                      d) 34,9  
 e) keine der genannten Temperaturen
- In welchem Jahr wurden in Ort X die größten Temperaturunterschiede beobachtet?  
 a) 1958                                      b) 1957                                      c) 1955                                      d) 1956  
 e) 1959
- In welchem Jahr betrug die Jahresniederschlagsmenge des Ortes Y weniger als 60 % der Jahresniederschlagsmenge des Ortes X?  
 a) 1956                                      b) 1955                                      c) 1957                                      d) 1958  
 e) 1959
- Wie groß war der Unterschied in cm zwischen der durchschnittlichen Jahresniederschlagsmenge des Ortes Y und des Ortes Z?  
 a) 12,8                                      b) 14,4                                      c) 13,6                                      d) 14,8  
 e) keine der genannten Mengen

**Anhang A4.5:** BIS-Item ZN - Zahlenreihen (KN)

Jede der folgenden Zahlenreihen ist nach einer bestimmten Regel aufgebaut.  
Sie sollen immer das **nächste Glied** der **Reihe** hinschreiben.

Welche Zahl muß an Stelle des Fragezeichens stehen?

Beispiel 1

2 5 8 11 14 17 ? \_\_\_\_\_

Beispiel 2

5 10 9 18 17 34 ? \_\_\_\_\_

**Beachten Sie bitte:**

Die Regel kann darin bestehen, daß zugezählt, abgezogen, malgenommen oder geteilt werden muß.

Welche Zahl muß an Stelle des Fragezeichens stehen?

**Schreiben** Sie die gesuchte Zahl auf die Linie am Ende der Zahlenreihe!

25	23	20	18	15	?	_____	
21	18	9	27	24	12	36	? _____
5	10	20	10	20	40	30	? _____
33	11	6	24	8	3	12	? _____
12	3	7	28	32	8	12	? _____
39	13	6	30	10	3	15	? _____
16	7	28	20	80	73	292	? _____
6	3	9	5	20	15	75	? _____
3	6	10	30	35	140	146	? _____

**Anhang A4.6: BIS-Item XG - X-Größer (BN)**

Streichen Sie alle Zahlen durch, die **um 7 größer** sind als die unmittelbar vorangegangene Zahl.

Arbeiten Sie **so rasch Sie können!**

**Beispiel**

Um 7 größer

15	20	7	10	3	<del>10</del>	14	7	12	<del>19</del>	5	<del>12</del>	26
—————→												
18	25	27	13	20	22	29	20	28	4	11	9	16
15	22	9	1	8	15	6	13	21	3	9	16	23

Streichen Sie jetzt so **schnell** wie möglich alle Zahlen durch, die um **3 (drei!)** größer sind als die unmittelbar vorangegangene Zahl.

um 3 größer

18	20	24	27	2	4	7	11	13	18	21	12	15
5	9	12	11	17	20	14	17	30	26	29	9	22
11	14	19	5	8	10	23	26	7	8	11	25	28
15	19	22	9	13	15	18	16	19	11	15	17	20
11	18	24	27	29	3	1	5	8	11	21	14	17
13	16	19	23	25	4	7	17	22	25	10	13	18
22	1	4	8	11	22	25	28	14	18	3	6	7
11	15	4	7	11	14	18	20	23	17	18	21	24
26	29	8	11	17	24	7	10	11	9	13	16	5
6	19	22	26	2	5	7	11	6	9	10	13	15

**Anhang A4.7: BIS-Item SC - Schätzen (KN)**

Die folgenden Rechenaufgaben sind leichter, als es auf den ersten Blick scheint. Ein genaues Ausrechnen ist nicht notwendig. Sie können die richtige Lösung durch **einfache rechnerische Überlegungen** finden. Nur eine der Wahllösungen a bis e ist richtig.

Streichen Sie den Buchstaben vor der richtigen Lösung durch.

## Beispiel 1

$$118\,492 - 3\,684 - 2\,106 - 4\,768 = ?$$

- a) 107 621                      b) 107 527                      ~~c) 107 934~~  
d) 98 743                        e) 99 825

## Beispiel 2

$$219 \times 5 + 712 \times 5 + 83 \times 5 = ?$$

- a) 5 072                        b) 5 067                        c) 5 074  
d) 5 071                        e) 5 070

Bei den folgenden Aufgaben werden Sie nicht immer an der letzten Stelle ablesen können, welche Lösung richtig ist. Es gibt aber immer rechnerische Anhaltspunkte, die ein herausfinden der richtigen Lösung ohne genaues Ausrechnen ermöglichen.

Gehen Sie daher möglichst von verschiedenen Seiten an die Aufgaben heran.

Streichen Sie den Buchstaben vor der richtigen Lösung durch.

1.  $711 \times 411 + 5137 = ?$

- a) 292 221                      b) 297 358                      c) 536 404  
d) 762 615                      e) 193 266

2.  $4567 \times 3187 = ?$

- a) 12 146 137                      b) 14 555 029                      c) 13 267 843  
d) 12 678 156                      e) 15 443 754

3.  $705 + 815 + 945 + 735 = ?$

- a) 3 210                              b) 2 880                              c) 3 240  
d) 3 200                              e) 3 860

4.  $56324186 - 52418218 = ?$

- a) 3 643 074                      b) 4 094 032                      c) 3 905 968  
d) 4 742 404                      e) 3 126 068

5.  $311 \times 811 + 45302 = ?$

- a) 297 523                              b) 250 601                              c) 248 631  
d) 442 904                              e) 361 222

6.  $41726 \frac{12}{27} + 89671 \frac{7}{27} + \frac{8}{27} = ?$

- a)  $120597 \frac{2}{27}$                               b) 131 398                              c) 135 108  
d)  $130397 \frac{7}{27}$                               e)  $129597 \frac{13}{27}$

7.  $59 \times 59 = ?$


- a) 3 681                                      b) 3 481                                      c) 3 701  
d) 3 440                                      e) 3 861

**Anhang A4.8: BIS-Item RZ - Rechen-Zeichen (BN)**

In den folgenden Rechnungen stehen an Stelle der Rechenzeichen (+ oder -) leere Kästchen. Setzen Sie bitte die richtigen Rechenzeichen (+ oder -) so ein, daß die Rechnung auf geht.

Arbeiten Sie in Pfeilrichtung und **so schnell Sie können!**

Beispiele


$$\begin{array}{rcccccc} 2 & \square & 2 & \square & 4 & = & 8 \\ 1 & \square & 2 & \square & 3 & = & 6 \\ 4 & \square & 3 & \square & 1 & = & 6 \\ 6 & \square & 4 & \square & 2 & = & 0 \end{array}$$



Setzen Sie + oder - so in die leeren Kästchen ein, daß die Rechnung auf geht.

Arbeiten Sie in Pfeilrichtung und **so schnell Sie können!**

↓

$$2 \square 1 \square 1 = 4$$
$$1 \square 2 \square 1 = 2$$

$$3 \square 2 \square 1 = 0$$

$$1 \square 1 \square 3 = 3$$

$$4 \square 2 \square 2 = 8$$

$$1 \square 2 \square 1 = 4$$

$$0 \square 1 \square 2 = 3$$

$$3 \square 3 \square 2 = 4$$

$$0 \square 2 \square 2 = 4$$

$$2 \square 1 \square 1 = 0$$

$$1 \square 2 \square 2 = 5$$

$$3 \square 3 \square 1 = 1$$

$$2 \square 2 \square 1 = 3$$

$$0 \square 4 \square 4 = 8$$

$$4 \square 1 \square 1 = 2$$

$$1 \square 3 \square 2 = 2$$

$$5 \square 2 \square 3 = 0$$

$$3 \square 2 \square 1 = 6$$

$$4 \square 1 \square 3 = 0$$

$$0 \square 3 \square 1 = 4$$

**Anhang A4.9: BIS-Item BR - Buchstabenreihen (KN)**

Die folgende Buchstabenreihe ist nach einer ganz bestimmten Regel aufgebaut. Wenn Sie die Regel erkannt haben, können Sie die **beiden folgenden Buchstaben** der Reihe ergänzen.

Tragen Sie bitte die **beiden folgenden Buchstaben** in die **freien Kästchen** ein:

Zur Orientierung das Alphabet:

**A B C D E F G H I J K L M N O P Q R S T U V W X Y Z**

Beispiel 1

**A Z B Y C X D W**

--	--

Beispiel 2

**A D C F E H G J**

--	--

Schreiben Sie die **beiden** folgenden Buchstaben in die **freien Kästchen**!

Zur Orientierung das Alphabet:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

A X A Y B X B Y C X C Y D

--	--

E E F G G H I I J K

--	--

E F E F C D G H G H E F I J I

--	--

B C D D C B C D E E D C D E F

--	--

C B D D C E E D F

--	--

B A B D C D F E F H G H

--	--

A D F I K N P S

--	--

C E W V G I U T K M S

--	--