

Selektive Aktivierung als Erklärung des Rückschaufehlers

Dissertation

zur Erlangung des Grades eines
Doktors der Naturwissenschaften

(Dr. rer. Nat.)

an der Universität Trier

vorgelegt von

Markus Eisenhauer

Gutachter:

Prof. Dr. K.F. Wender

PD Dr. Rüdiger F. Pohl

FB-I Psychologie
Universität Trier
Universitätsring 15
54286 Trier

FB 06 - Psychologie und
Sportwissenschaften
Justus-Liebig-Universität, Gießen
Otto-Behaghel-Str. 10 F
35394 Gießen

Juni 2000

FÜR BRITTA

Die genaue Erinnerung ist alles mögliche, ein ursachenunbelastetes Auflisten eines ganzen Gehirnlandes, das Nebenseitige als Möglichkeit, wie es anders hätte sein können, wenn es nicht gekommen wäre, wie es hereingetreten werden mußte.
(Werner Schwab, 1993; aus Volksvernichtung in Fäkaliendramen. Graz-Wien: Verlag Droschl.)

Inhaltsverzeichnis

BEFUNDE UND THEORIE

Kapitel 1 Einleitung

| | | |
|-----|----------------------------------|---|
| 1.1 | Beschreibung des Phänomens | 3 |
|-----|----------------------------------|---|

Kapitel 2 Phänomen und Befunde

| | | |
|-------|---|----|
| 2.1 | Phänomen | 4 |
| 2.1.1 | Definition | 4 |
| 2.1.2 | Untersuchungsdesign | 5 |
| 2.2 | Befunde | 10 |
| 2.2.1 | Material | 10 |
| | Fallgeschichten | 10 |
| | Prognostische Urteile | 11 |
| | Almanachfragen | 11 |
| | Politische Wahlen | 13 |
| | Evaluation von Entscheidungen | 14 |
| | Soziale Urteilsphänomene | 14 |
| | Medizinische Urteile | 15 |
| | Juristische Urteile | 16 |
| 2.2.2 | Zur Robustheit des Rückschaufehlers | 17 |
| | Instruktion und Aufklärung | 17 |
| | Wahrheit | 18 |
| | Schwierigkeit | 19 |
| | Überraschung | 19 |
| | Plausibilität | 20 |

Kapitel 3 Theorien zum Rückschaufehler

| | | |
|-------|---|----|
| 3.1 | Adaptives Lernen und Motivation | 21 |
| 3.1.1 | Rückschaufehler und adaptives Lernen | 21 |
| 3.1.2 | Motivationale Faktoren beim Rückschaufehler | 22 |
| | Selbstdarstellung | 22 |
| | Motivation versus Kognition | 23 |
| 3.2 | Kognitive Erklärungen | 24 |
| 3.2.1 | Der Rückschaufehler als Gedächtnisphänomen | 25 |
| | Erinnerungsvermögen | 25 |
| | Unmittelbare Assimilation | 25 |
| | Theorien zum Falschinformationseffekt | 27 |
| | Modell relativer Spurstärke | 31 |
| 3.2.2 | Der Rückschaufehler als Urteilsphänomen | 32 |
| | Anchoring and Adjustment | 33 |
| | Antworttendenzmodell | 33 |
| | Rekonstruktionstheorie | 36 |
| | Selektive Zugänglichkeit | 37 |
| | Selektive Aktivierung | 38 |
| 3.3 | Abschließende Anmerkung | 39 |

Kapitel 4 SARA

| | | |
|-----|---|----|
| 4.1 | Einführung | 40 |
| 4.2 | Organisation des Wissens | 40 |
| | Ablaufschema | 40 |
| | Organisation | 41 |
| | Architektur | 43 |
| | Informationseinheiten | 44 |
| 4.3 | Prozesse | 45 |
| | Sampling | 45 |
| | Abruf von Images | 45 |
| | Lernen | 46 |
| | Vergessen | 47 |
| 4.4 | Spezifische Verarbeitungsprozesse | 49 |
| | Generierung einer Schätzung | 49 |
| | Enkodierung eines Ankers | 52 |
| | Rekonstruktion einer Schätzung | 54 |
| 4.5 | Fazit | 56 |

EMPIRISCHE ARBEITEN

Kapitel 5 Experimente mit Wiedererkennenstest

| | | |
|-----|--|-----|
| 5.1 | Experiment 1: Wiedererkennen der Argumente | 60 |
| | Methode | 60 |
| | Ergebnisse | 68 |
| | Diskussion | 72 |
| 5.2 | Experiment 2: Wiedererkennen mit komprimierter Schätzphase | 76 |
| | Einleitung | 76 |
| | Methode | 77 |
| | Ergebnisse | 81 |
| | Diskussion | 86 |
| 5.3 | Experiment 3: Wiedererkennen mit relevanten Sachverhalten | 90 |
| | Einleitung | 90 |
| | Methode | 92 |
| | Ergebnisse | 95 |
| | Diskussion | 103 |

Kapitel 6 Experimente mit Erinnerungsmaßen

| | | |
|-----|--|-----|
| 6.1 | Experiment 4: Freie Wiedergabe der Argumente im Gedächtnisdesign | 109 |
| | Einleitung | 109 |
| | Methode | 111 |
| | Ergebnisse | 114 |
| | Diskussion | 125 |
| 6.2 | Experiment 5: Freie Wiedergabe der Argumente im Urteilsdesign | 131 |
| | Einleitung | 131 |
| | Methode | 131 |
| | Ergebnisse | 137 |
| | Diskussion | 140 |

FAZIT

Kapitel 7 Fazit

| | | |
|-------|---|-----|
| 7.1 | Abschließende Diskussion..... | 144 |
| 7.1.1 | Material | 145 |
| | Konfidenzskala | 145 |
| | Kausalstruktur | 145 |
| | Dichotome Entscheidung | 146 |
| | Wert der Images | 149 |
| | Wichtigkeit der Argumente | 150 |
| 7.1.2 | Allgemeine Aspekte | 151 |
| | Selektive Aktivierung versus selektiver Abruf | 151 |
| | Regressionseffekte und Rückschaufehler | 151 |
| | Reaktionszeiten | 153 |
| | Widerspruch | 153 |
| 7.2 | Abschluß und Ausblick | 155 |
| 7.2.1 | Fazit | 155 |
| 7.2.2 | Ausblick | 156 |
| | Literatur | 159 |

ANHANG

| | | |
|-----|---|-----|
| A | Sachverhalte | 172 |
| A.1 | Sachverhalte zu Experiment 1-2 | 172 |
| A.2 | Sachverhalte zu Experiment 3-5 | 190 |
| B | Instruktionen | 214 |
| B.1 | Instruktionen der Experimente am Computer | 214 |
| B.2 | Instruktionen zum Fragebogen | 217 |
| C | Erklärung | 221 |

*Wir erinnern uns, wie es uns gefällt. Würde man es sich nicht erlauben, einige wenige Lügen hinzunehmen, ich weiß nicht, wie man jemals die Vergangenheit ertrüge. Gott sei Dank, daß es, abgesehen vom Augenblick des Geschehens, nie so etwas gibt wie eine nackte Tatsache. Zehn Minuten später hat man schon begonnen, sie mit irgendeiner Art von Kruste zu überziehen. (Elizabeth Bowen, Juni 7 – 13 1989; aus *Der Wendepunkt. Ein Lebensbericht*. Reinbek: Rowohlt Verlag.)*

Danksagung

Eine solche Arbeit kann natürlich nicht ganz ohne die Hilfe anderer gelingen. Ich möchte mich daher an dieser Stelle bei Allen bedanken für die Geduld und die anregenden Diskussionen, die dieser Arbeit zugute gekommen ist.

Zunächst danke ich ganz besonders meinen Gutachtern und Betreuern PD. Dr. Rüdiger Pohl und Prof. Dr. K.-F. Wender. Ohne Ihre Unterstützung wäre diese Arbeit nicht zustande gekommen. Oliver Hardt danke ich für Diskussion und Korrektur. Besonders möchte ich auch Frau Maria Huber danken, für das sorgfältige und schnelle Korrekturlesen dieser Arbeit. Gerd Fessler, Jürgen Gehrke und Bettina Menzel danke ich ebenfalls für Korrekturlesen.

Dank gebührt natürlich auch allen, die bei der Durchführung der Experimente geholfen haben: Thomas Wiesenfeldt, Henning Schulz, Stefan Lambert.

Zusammenfassung

Beim Rückschaufehler handelt es sich um eine systematisch verzerrte Erinnerung an früher getroffene Urteile. Der Rückschaufehler ist eine robuste kognitive Täuschung, die bislang aber nur unbefriedigend erklärt wurde. In dieser Arbeit wird ein detailliertes kognitives Prozeßmodell (SARA) vorgestellt. Der grundlegende Prozeß ist demnach ein zyklischer, probabilistischer Abruf von Informationen aus dem Langzeitgedächtnis ins Arbeitsgedächtnis. Dies führt zu einer höheren Abrufwahrscheinlichkeit dieser Informationen und wird daher im Modell als „*selektive Aktivierung*“ bezeichnet. Der Rückschaufehler ist somit auf veränderte Abrufwahrscheinlichkeiten von durch die Enkodierung der Lösung *selektiv aktivierten* Informationen zurückzuführen.

In fünf Experimenten wurde das Konzept „*selektive Aktivierung*“ überprüft. Probanden sollten die Entwicklung von Sachverhalten beurteilen und sich später nach Bekanntgabe der Lösung wieder daran erinnern. Zusätzlich wurden Argumente dargeboten, die für eine Zu- bzw. Abnahme des Sachverhalts sprachen. Durch die Lösung *selektiv aktivierte* Argumente sollten zu einem verbesserten Abruf führen.

In drei Wiedererkennenstests konnte eine selektive Aktivierung der Argumente durch die Lösung nicht nachgewiesen werden. In diesen Experimenten trat allerdings auch der Rückschaufehler nicht auf. In zwei Experimenten mit freier Wiedergabe ließen sich die erwarteten Effekte dagegen zeigen. Dies kann als erster Nachweis selektiver Aktivierung interpretiert werden.

Befunde und Theorien

1 Einleitung

Aus der gegenwärtigen Perspektive betrachtet erscheinen uns vergangene Ereignisse oft plausibel. Das sich Stefanie von Gerd trennen würde, daß Markus gerade dieses eine Angebot auswählen würde, daß die Bundestagswahl so ausgehen würde, daß das so nie klappen konnte, das war uns doch schon immer klar. Die Beurteilung von Ereignissen aus der Vergangenheit ist eine diffizile Angelegenheit. Nicht nur die Schwierigkeit, sich alle Details eines vergangenen Ereignisses zu vergegenwärtigen, sondern vor allem die Unsicherheit über den Ausgang einer Situation, bei mehreren möglichen Alternativen, ist nach Bekanntwerden des Ausgangs nicht mehr nachvollziehbar. Im Nachhinein scheint die Situation eindeutig und es drängt sich das Gefühl auf, es schon „immer gewußt zu haben“ oder gewußt zu haben, daß das so gar nicht hätte funktionieren können. Allerdings stellt sich dann natürlich die naheliegende Frage, warum man dann beispielsweise die Situation nicht gleich ganz anders in Angriff genommen hat. Es ist klar, daß hier die Entscheidungssituation vollkommen falsch eingeschätzt wird. Die ursprüngliche Situation mit ihren mehreren möglichen und (unterschiedlich) wahrscheinlichen Ausgängen wird anders eingeschätzt. Entweder werden alternative Ausgänge gar nicht mehr bedacht oder die Wahrscheinlichkeit für die Ausgänge verschiebt sich stark zugunsten des eingetretenen Ereignisses.

Das Phänomen ist klar. Systematisch wird der vergangene Wissensstand in der Erinnerung *zu gut*¹ eingeschätzt. Post factum stellt sich manche Situation scheinbar eindeutig dar, die prä factum gar nicht so eindeutig schien. Dieses Phänomen, nach Bekanntwerden des Ausgangs systematisch die naive Situation (ohne Lösungswissen) zugunsten des tatsächlichen eingetretenen Ausgangs (fehl-) einzuschätzen, soll einem Vorschlag Rüdiger Pohls (1992) folgend als *Rückschaufehler* bezeichnet werden.

Wie der Rückschaufehler zustande kommt und wie er eventuell (falls überhaupt) vermieden werden kann, das ist Gegenstand der vorliegenden Arbeit.

¹ Mit „zu gut“ soll keine Wertung verbunden sein. Impliziert ist bereits die Perspektive des bekanntgewordenen Ausgangs. Gemeint ist hier, daß die Verfälschung im Mittel immer in Richtung des Ausgangs geht.

2 Phänomen und Befunde

2.1 Phänomen

2.1.1 Definition

In der Einleitung zu dieser Arbeit hatte ich den Rückschaufehler eingeführt als eine systematische (Fehl-) Einschätzung des naiven Wissensstands zugunsten des zwischenzeitlich bekanntgewordenen Ausgangs. Also die Tendenz von Individuen in der Rückschau, mit Lösungswissen, die Wahrscheinlichkeit des berichteten Ausgangs höher einzuschätzen als sie tatsächlich in der Vorausschau, ohne Lösungswissen, ausfiel. Der Rückschaufehler ist eine Projektion neuen Wissens in die Vergangenheit, begleitet von der Leugnung, daß jenes Wissen Einfluß auf das Urteil ausgeübt haben könnte. Komplex wird die Angelegenheit noch zusätzlich dadurch, daß die gleichen kognitiven Mechanismen, die durch einen Feedback in einer Urteilsaufgabe wirken, für Rückschautendenzen als auch für adaptives Lernen verantwortlich gemacht werden können. Läßt man beispielsweise die Instruktion im Rückschauparadigma, so zu antworten, als wäre der Ausgang noch unbekannt, weg, hätte man eine einfache Lernbedingung. Man erwartete geradezu, daß die Probanden ihr Wissen adaptierten.

Kohnert (1996) und Hawkins und Hastie (1990) weisen darauf hin, daß oft neben der eigentlichen Verfälschung noch eine weitere Täuschung metakognitiver Natur stattfindet: *Zum Rückschaufehler gehört nicht nur, daß Personen ihr neu erworbenes Wissen über ein Ereignis nicht ignorieren, wenn sie alte Vorhersagen rekonstruieren sollen, sondern auch, daß ihnen dieser Einfluß nicht bewußt ist.* (Kohnert, 1996, S. 24). Dieser Aspekt der Unbewußtheit über Einflussgrößen eröffnet einen metakognitiven Erklärungsrahmen, der einem automatischen adaptiven Prozeß vorgeschaltet sein könnte. So könnte beispielsweise in bestimmten Situationen, die durch die Art der Rückmeldung oder durch bewußte Instruktion die Einflussgrößen erkennbar machen, eine automatische Einflußnahme des Ausgangs (der Lösung) vermieden werden. So verschwand beispielsweise in einem Experiment von Hasher, Attig und Alba (1981; Exp. 2) und ebenfalls in einem Experiment von Erdfelder und Buchner (1998; Exp. 3) nach der Warnung, die soeben präsentierte Information sei falsch gewesen, der Rückschaufehler.

Die Unbewußtheit ist aber keine notwendige Voraussetzung für das Auftreten des Effekts. So konnten Pohl und Hell (1996) zeigen, daß der Effekt sogar dann auftritt, wenn den Probanden, neben einer ausführlichen Aufklärung über den Effekt und der Aufforderung, diesen zu vermeiden, noch jedem Probanden individuell der persönliche Bias genau erklärt wurde. Gleichzeitig wurden sie aufgefordert, in einem folgenden Experiment stärker dagegen anzugehen. Auch eine zweimalige Wiederholung konnte den Rückschaufehler nicht reduzieren.

Allgemein kann man davon ausgehen, daß der Rückschaufehler bei Fragen (oder in Situationen) auftritt, bei deren Beantwortung eine gewisse Unsicherheit herrscht. Vor der Aufforderung, einen vergangenen (oder hypothetischen) Wissensstand wiederzugeben, müssen Informationen vorliegen, die genau diese Unsicherheit reduzieren. Diese Informationen werden im Normalfall die korrekten Antworten auf die Frage (oder der tatsächliche Ausgang eines Ereignisses) sein. Im Prinzip ist es jedoch nicht von entscheidender Bedeutung, ob die Information korrekt oder wahr ist. So kann beispielsweise auch das Urteil eines anderen Probanden zu der beobachtbaren mittleren Verschiebung in Richtung der Rückmeldung führen.

Der Rückschaufehler wurde in zahlreichen Studien demonstriert. Eine gute Literaturübersicht findet sich bei Hawkins und Hastie (1990) oder bei Pohl (1993). Eine Metaanalyse von Christensen-Szalanski und Willham (1991) über 128 Experimente zum *hindsight bias* fand in lediglich sechs Experimenten keinen Rückschaufehler und belegt eindrucksvoll die Robustheit des Effekts. Die Stabilität des Rückschaufehlers legt den Gedanken nahe, daß es sich hier um ein grundlegendes Phänomen handelt. Vor allem stabile, automatische Prozesse der menschlichen Informationsverarbeitung kommen zur Erklärung des Phänomens in Betracht.

Bevor ich jedoch zu den Theorien und Erklärungen des Rückschaufehlers kommen kann, möchte ich zunächst auf das Untersuchungsparadigma und das verwendete Material eingehen.

2.1.2 Untersuchungsdesign

Es sollen hier drei verschiedene Untersuchungsdesigns vorgestellt werden: Gedächtnisdesign (*memory design*), Urteilsdesign (*hypothetical design*) und Ankerdesign (*anchoring design*). Im Wesentlichen lassen sich die Designtypen an dem Zeitpunkt,

zu dem die kritische Information (Anker, Lösung, Ausgang der Situation) präsentiert wird, unterscheiden: Wird diese Information vor Abgabe einer eigenen Schätzung präsentiert, ist die Aufgabe, eine unbeeinflusste Schätzung zu generieren, ein Urteilsprozeß (Urteils-, Ankerdesign); wird die Information erst nach Abgabe einer unabhängigen Schätzung präsentiert, handelt es sich eher um eine Gedächtnisaufgabe (Gedächtnisdesign). Diese Unterscheidung zwischen Urteils- und Gedächtnisprozessen ist auch Gegenstand einer bislang noch nicht entschiedenen Kontroverse zur Erklärung des Rückschafehlers (siehe Kap. 3). In Abb. 1.1 ist der schematische Ablauf eines Experimentes im Gedächtnisdesign anhand einer Frage dargestellt. Das Gedächtnisdesign ist gewissermaßen das vollständigste Design. Es enthält alle Elemente, die auch in den anderen Designtypen auftauchen. Daher bietet es sich an, dieses Design zur Erläuterung zu verwenden. Es besteht in seiner Grundversion aus fünf Phasen (siehe Abb. 1.1):

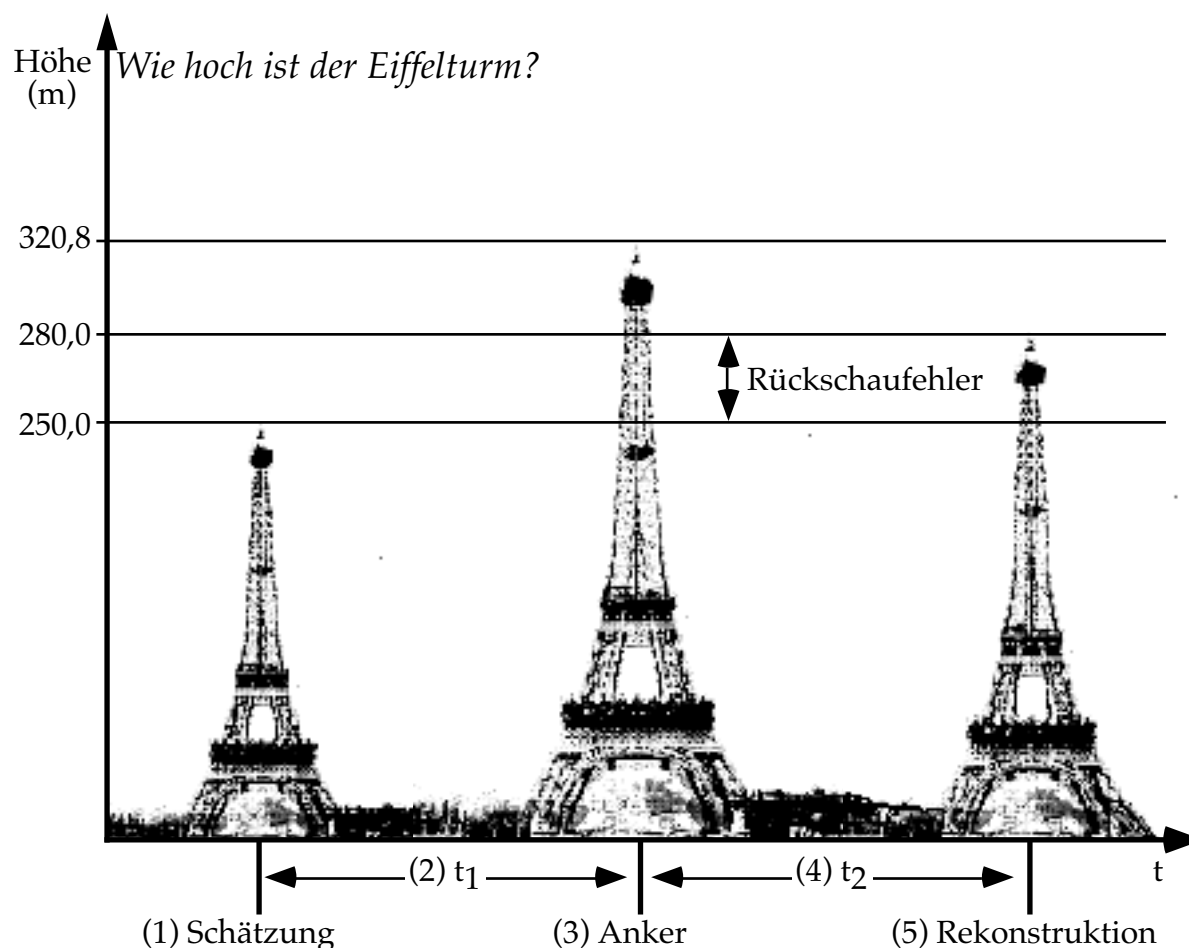


Abb. 1.1: Die Abbildung zeigt anhand einer Almanachfrage den typischen Ablauf im Gedächtnisdesign: Nach Abgabe einer Schätzung, anschließender Enkodierung eines Ankers ist die Rekonstruktion der Schätzung in Richtung des Ankers verschoben. Die Numerierungen in der Abbildung korrespondieren entsprechenden Erläuterungen im Text.

(1) In der ersten Phase werden typischerweise Antworten auf schwierige Wissensfragen (Almanachfragen²) verlangt. Die Fragen sind meist so gewählt, daß nur sehr wenig Probanden die Lösung kennen, aber möglichst viele genügend Wissen haben, um eine Schätzung abzugeben. In Abb. 1.1 wurde beispielsweise auf die Frage nach der Höhe des Eiffelturms 250 Meter geschätzt.

(2) Es folgt ein erstes Retentionsintervall. Je nach experimenteller Anordnung vergeht mehr oder weniger Zeit, bis der Anker präsentiert wird. Die Veränderung dieses Retentionsintervalls erlaubt es, gezielt den Einfluß von Vergessen auf die korrekten Erinnerungen an die Schätzungen zu manipulieren. Es zeigte sich der erwartete Effekt, daß ein Anstieg des Retentionsintervalls zu einem Rückgang korrekter Erinnerungen und zu einem stärkeren Rückschaufehler führte (siehe Fischhoff, 1975; Hell, Gigerenzer, Gauggell, Mall, & Müller; 1988; *Pennington, 1981*).

(3) Diese dritte Phase ist zugleich die erste Phase für Experimente im Urteils- und Ankerdesign. Es ist naturgemäß auch die wichtigste, entscheidende Phase. Schließlich wird hier die Information eingeführt, die den Effekt verursacht. Typischerweise werden den Probanden zu Experimentalitems Informationen präsentiert, die normalerweise als Lösung bezeichnet werden (zu Kontrollitems werden keine Lösungen präsentiert). Um die verschiedenen Bezeichnungen in Untersuchungen zum Rückschaufehler und zu Ankereffekten zu vereinheitlichen, möchte ich allgemein von Ankerinformation oder kurz *Ankern* sprechen. In Abbildung 1.1 wird als Antwort auf die Frage nach der Höhe des Eiffelturms der (korrekte) Anker 320,8 Meter präsentiert. Der Anker wird in dieser Phase zur Kenntnis genommen und je nach theoretischer Vorstellung enkodiert oder nicht. Wie bereits erwähnt, muß der Anker nicht notwendigerweise als Lösung bezeichnet werden, damit er als relevante Information genutzt wird. Auch als Schätzung eines anderen Probanden zeigte er seine Wirkung (Pohl, 1998), sogar als Ergebnis eines Münzwurfs (Connolly & Bukzar, 1990). Ähnlich erstaunlich war auch die Wirkung eines Glücksrades bei Tversky und Kahneman (1974) auf die Frage nach dem Anteil afrikanischer Staaten in der UNO.

Letzteres Experiment stellt zugleich auch den Musterfall für Experimente im Ankerdesign dar. Entsprechend dem Vorgehen im Urteilsdesign wird auch im

² Das Material, welches zur Überprüfung des Rückschaufehlers genutzt wird, ist durchaus unterschiedlich. Neben episodischem Material werden einfache Wissensfragen (Almanachfragen) verwendet. Auch der Antworttyp, ob es sich um eine Antwort auf einer festen Antwortskala handelt oder ob die Antwort keinerlei Beschränkungen des Antwortformats unterliegt, wird unterschieden (siehe dazu auch später in diesem Kapitel). Aus pragmatischen Gründen und weil es im deutschen Sprachraum das meist verwendete Antwortformat ist, sei das Design hier an einer Almanachfrage mit offener Antwortskala eingeführt.

Ankerdesign, zusammen mit der gestellten Frage, direkt der Anker präsentiert. Abweichend vom Urteilsdesign folgt im Ankerdesign dann oft noch ein komparatives Urteil: „*Liegt die Lösung höher oder niedriger als der Anker?*“ Dieses Urteil ist für unplausible Anker (wie etwa das Ergebnis des Glücksrades) entscheidend, damit eine systematische Verfälschung des Ankers beobachtet werden kann³.

(4) Der Präsentation des Ankers folgt ein zweites Retentionsintervall. Ähnlich, wie das erste haben experimentelle Manipulationen Auswirkungen auf das Vergessen. Das Vergessen betrifft in diesem Fall naturgemäß Schätzung und Anker. Es zeigte sich der erwartete Effekt, daß ein Anstieg des Retentionsintervalls zu einem schwächeren Rückschaufehler führte (siehe Erdfelder & Buchner, 1998; Hell et al. 1988). In den meisten Experimenten im Urteils- und Ankerdesign und ebenfalls in einigen Experimenten im Gedächtnisdesign fällt dieses zweite Intervall weg: Zum Zeitpunkt der Abgabe einer eigenen Schätzung (Urteils- und Ankerdesign) respektive der Rekonstruktion der Schätzung (Gedächtnisdesign) liegt der Anker vor. Die Konsequenz ist, daß der Anker nicht erinnert werden muß und der Effekt maximiert wird (z.B. Stahlberg, Eller, Romahn, & Frey, D., 1993). Diese experimentelle Situation hat aber einen entscheidenden theoretisch relevanten Nachteil. Sie erlaubt nicht zu entscheiden, ob es sich bei dem Effekt um eine Antwortstrategie oder einen Gedächtniseffekt handelt. Wenn man daher an einem großen Rückschaufehler interessiert ist, aber keinen direkten Konflikt zwischen Urteil und Anker erzeugen will, liegt es nahe, Ankerdarbietung und die Aufforderung sich an die eigene Schätzung zu erinnern nicht simultan, aber in kurzem Abstand voneinander darzubieten (z.B. gemeinsam in einer experimentellen Sitzung)

(5) In dieser letzten Phase folgt die Aufforderung an die Probanden sich an die eigene Schätzung zu erinnern. In Abbildung 1.1 wird die Höhe des Eiffelturms beispielsweise als 280 Meter rekonstruiert. Typischerweise findet bei Experimentalitems im Mittel eine Annäherung der Erinnerungen an den Anker. Die Erinnerungen der Kontrollitems liegen im Mittel um den gleichen Wert, wie die Schätzung. Dabei setzt man voraus, daß ohne die Ankerinformation die Erinnerungen unsystematisch um die Schätzungen streuen⁴.

³ Ohne das komparative Urteil bestünde kein semantischer Zusammenhang zwischen Anker und Schätzung. Die Frage, ob die eigene Schätzung unter oder oberhalb des zufälligen Ankers liegt, macht den Anker erst zu einer nützlichen Information. Von einer Zufallszahl wird der Anker zu einer Information mit dem Inhalt *der Wert liegt ober-/unterhalb des Ankers*.

⁴ Dies ist aber nicht immer der Fall. „*Nimmt man nämlich an, daß die Schätzungen der Probanden nicht zu 100% reliabel sind, so müssen die Erinnerungsverteilungen, wenn sich die Probanden an nichts mehr erinnern*

Zusammenfassend läßt sich festhalten, daß im Gedächtnisdesign alle beschriebenen fünf Phasen vorkommen (wobei hier Extremfälle, in denen eines der Retentionsintervalle den Wert Null hat, eingeschlossen sein sollen) und sowohl im Anker- als auch im Urteilsdesign die ersten beiden Phasen wegfallen und gleich mit der Ankerdarbietung begonnen wird. Der entscheidende theoretische Unterschied zwischen Urteils- (Anker-) und Gedächtnisdesign liegt in der Definition und Erklärung des Rückschaufehlers. Schließlich gibt es im Gedächtnisdesign im Gegensatz zum Urteilsdesign eine objektive Wahrheit (die tatsächlich gegebene Schätzung). Daraus folgt, daß im Gedächtnisdesign die eigenen Schätzungen als Kontrolle für die spätere Erinnerung oder Rekonstruktion dienen, während im Urteilsdesign nur gegen den Mittelwert einer Personengruppe ohne Ankerwissen getestet werden kann. Wie bereits angedeutet, wird der Rückschaufehler im Urteilsdesign eher als Urteilsphänomen verstanden und mit Antworttendenzen erklärt, während der Rückschaufehler im Gedächtnisdesign eher als Gedächtnisverzerrung verstanden und mit Gedächtniseffekten erklärt wird (siehe auch Hawkins & Hastie, 1990). Im Urteilsdesign kann man die beiden Einflußgrößen prinzipiell nicht unterscheiden. Nur das Gedächtnisdesign ermöglicht eine Trennung von Urteils- und Gedächtniseffekten. Es ist daher auch für die theoretische Fragestellung relevant, welchen der beiden Designtypen man zur Operationalisierung des Rückschaufehlers einsetzt.

Der Rückschaufehler ist im Gedächtnisdesign in der Regel kleiner als im Urteilsdesign (Campbell & Tesser, 1983; Davies, 1992; Fischhoff, 1977; Pennington, 1981; Pennington, Rutter, McKenna, & Morley, 1980; Powell, 1988; Wood, 1978, Exp. 2). Dieser Unterschied ist besonders stark bei kurzen Retentionsintervallen; bei längeren Zeitabständen wird er dagegen nivelliert (Fischhoff & Beyth, 1975). Da der Rückschaufehler im Gedächtnisdesign kleiner ausfällt, übernehmen hier Gedächtnisprozesse einen größeren Anteil bei der Rekonstruktion. Schon die häufigere korrekte Erinnerung der eigenen Schätzung oder auch nur des Bereichs, in dem diese lag, sorgt zwangsläufig für eine geringere mittlere Abweichung von der Schätzung. Im Urteilsdesign dagegen gibt es keine Erinnerungsspuren. Die Lösung übt ihre verfälschende Wirkung daher allein während des Urteilsprozesses aus.

und daher einfach noch mal schätzen, zum Mittelwert der Schätzverteilung regredieren. Da dieser Mittelwert in der Regel nicht extrem weit vom Anker entfernt liegt, regredieren die Erinnerungen nicht nur gegen den Mittelwert der Schätzverteilung, sondern auch gegen den Anker selbst“ (vgl. Kohnert, 1996, S.60).

2.2 Befunde

2.2.1 Material

Im Wesentlichen können drei verschiedene Arten von Materialien zur Untersuchung des Rückschaufehlers unterschieden werden: Beurteilung von Fallgeschichten, Prognosen bei echten Ereignissen und Wissensurteile bei Almanachfragen. Mit allen Materialarten ließ sich über einen weiten Bereich von Situationen ein Rückschaufehler zeigen.

Fallgeschichten. Bei Fallgeschichten wird Probanden zunächst eine realistische Geschichte mit offenem Ausgang zum Lesen gegeben. Am Ende der Geschichte folgen mehrere alternative Ausgänge. Im Urteilsdesign⁵ folgt dann die Behauptung, ein bestimmter Ausgang sei eingetreten, zusammen mit der Aufforderung, die Wahrscheinlichkeit der Ausgänge einzuschätzen und dabei die Ergebnisinformation zu ignorieren. Der Rückschaufehler zeigt sich durch einen Vergleich mit einer Kontrollgruppe, die keine Information über den Ausgang erhält.

Die bekannteste Untersuchung zum Rückschaufehler soll hier als Beispiel dienen: Fischhoff (1975) legte seinen Probanden vier Fallgeschichten vor (zwei historische Ereignisse aus Geschichtsbüchern, zwei Fallgeschichten aus der klinischen Psychologie). In der prominentesten Fallgeschichte wurde das Szenario eines Krieges der Gurkas gegen die Engländer an der nördlichen Grenze von Indien mit Bengalen im November 1814 berichtet. Die Probanden erfuhren etwas über die Kampfstärke der Gurkas, die Gelände- und klimatischen Verhältnisse usw. Nachdem die Probanden die Geschichte gelesen hatten, erhielt ein Teil von ihnen Information über den angeblichen Ausgang der Geschichte (z.B. *„Beide Seiten erreichten eine militärische Pattsituation, konnten sich jedoch nicht auf ein Friedensabkommen einigen“*). Im Anschluß daran sollten sie die Wahrscheinlichkeit für die vier möglichen Ausgänge des Konfliktes angeben, so als würden sie den Ausgang nicht kennen. Eine Kontrollgruppe erhielt keine Information über den Ausgang der Geschichte und sollte einfach die Wahrscheinlichkeit der Ausgänge einschätzen. Fischhoff konnte zeigen, daß Probanden der Experimentalgruppe den rückgemeldeten Ausgang systematisch wahrscheinlicher einschätzten als die Kontrollgruppe.

⁵ Die Mehrzahl der Experimente mit Fallgeschichten wurde im Urteilsdesign durchgeführt. Es existieren nur wenige Experimente mit Fallgeschichten im Gedächtnisdesign (siehe Eisenhauer, 1998).

Prognostische Urteile. Hier wird das gleiche Konzept verfolgt wie auch schon bei den Fallgeschichten. Der Realitätsgrad wird jedoch erhöht, indem ein prognostisches Urteil über ein bevorstehendes Ereignis verlangt wird. Im Unterschied zu den Fallgeschichten kann der Ausgang des Ereignisses hier nicht manipuliert werden. Nach dem Eintreten des Ereignisses bittet man die Probanden nochmals, ihre Prognose zu erinnern und vergleicht diese mit den abgegebenen Prognosen.

Wiederum sei zur Illustration eine klassische Studie von Fischhoff und Beyth (1975) zitiert: Sie ließen israelische Studenten vor dem historischen Staatsbesuch von US-Präsident Nixon in der Volksrepublik China 1972 die Wahrscheinlichkeit möglicher Ergebnisse des Besuchs einschätzen. Einige Zeit nach dem Staatsbesuch sollten sich die Probanden an ihre Einschätzungen erinnern. Bei tatsächlich über die Medien bekanntgewordenen Ergebnissen des Staatsbesuchs waren die Erinnerungen systematisch in Richtung der Resultate verzerrt.

Verplanken und Pieters (1988) befragten ihre Probanden vor und nach dem Kernreaktorunfall in Tschernobyl nach der Wahrscheinlichkeit für einen schweren Unfall. Überraschenderweise fanden Sie, daß die Wahrscheinlichkeit für schwere Unfälle unmittelbar nach dem Kernreaktorunfall geringer eingeschätzt wurde. Sie interpretierten ihre Befunde als einen umgekehrten Rückschleifer. Diese Interpretation ist allerdings mehr als zweifelhaft, da unmittelbar nach dem Unfall lediglich ein paar Dutzend Arbeiter starben und die Ergebnisse demnach auch so interpretiert werden können, daß sie einen Rückschleifer zeigen (siehe auch Arkes 1988). Wichtig ist daher festzuhalten, daß ohne eindeutige und explizite Definitionen der zu beurteilenden Ereignisse die Befunde ebenfalls nicht eindeutig interpretierbar sind.

Almanachfragen. Hier findet ein qualitativer Wechsel von episodischem, verknüpftem (oft auch kausal zusammenhängendem) Wissen zu eher unzusammenhängendem Faktenwissen, wie es etwa aus Wissensquiz bekannt ist, statt. Es werden eine Vielzahl von Fragen aus unterschiedlichen Bereichen gestellt. Als Anker wird entweder die korrekte Antwort auf die Wissensfrage, die Schätzung eines anderen Experten oder eines anderen Probanden rückgemeldet. Wie bereits berichtet, können sogar Zufallszahlen als Anker wirken. Auch diese Methode wurde zuerst von Fischhoff (1977) angewendet. Fischhoff legte seinen Probanden Aussagen wie „*Absinth ist ein Edelstein*“ vor. Aufgabe war, die Richtigkeit der Aussage auf einer Skala zu verifizieren. Danach erhielt ein Teil der Probanden die richtigen Antworten auf die Fragen, während ein anderer Teil keine Informationen bekam. Schließlich

hatten alle Probanden die Aufgabe, sich an ihre Schätzungen zu erinnern. Fischhoff konnte zeigen, daß bei Experimentalitems verglichen zu Kontrollitems eine Zunahme der Konfidenz für den Anker zu beobachten war.

Ein Problem bei Verifikationsaufgaben ist, daß die Sensibilität des Maßes relativ gering ist, da die Endpunkte der Skala zugleich die Anker sind und nicht über sie hinausgegangen werden kann. Die Wahrscheinlichkeit für einen Rückschleifer ist damit abhängig von der Güte der ersten Antworten. Liegen die Probanden mit ihrer ersten Antwort auf der Seite des Ankers, gibt es keinen (oder nur geringen) Spielraum für eine Annäherung an den Anker. Liegen sie dagegen auf der anderen Seite, gibt es mehr Möglichkeiten, sich in der Erinnerung dem Anker anzunähern als sich von ihm zu entfernen.

In vielen Studien werden Aufgaben gestellt, die Prozentangaben zwischen 0 - 100% verlangen (z. B. Leary, 1982; Stahlberg, Eller & Frey, 1990; Verplanken & Pieters, 1988). Hier kann der Anker prinzipiell überall im Bereich von 0 bis 100% liegen und ist damit nicht auf die Endpunkte der Skala beschränkt. Allerdings können Schätzungen und Erinnerungen – wie bei der Ratingskala – nicht über diese Endpunkte hinausgehen.

Im deutschen Sprachraum wird vor allem mit quantitativen Wissensfragen gearbeitet, bei denen die Probanden keine Wahrscheinlichkeits- oder subjektiven Sicherheitsurteile abgeben müssen, sondern direkt auf einer quantitativen Skala antworten können („Wie hoch ist der Eiffelturm?“). Die zur Verfügung stehenden Zahlenbereiche sind prinzipiell nach oben und unten offen. Dies bringt aber eine Reihe von Problemen mit sich. So ist zum einen die Varianz der Schätzungen (und dementsprechend auch die der Erinnerungen) oft sehr groß. Zum anderen ist der direkte Vergleich zwischen verschiedenen Items nicht mehr möglich. Die numerischen Angaben der Probanden müssen deswegen auf ein vergleichbares Maß transformiert werden (z. B. Hell et al., 1988; Kohnert, 1996).

Die wichtigste Unterscheidung im verwendeten Material betrifft, wie bereits erwähnt, die Unterscheidung, ob es sich um singuläre Fakten oder um zusammenhängende Texte, deren Informationen kausal miteinander verknüpft sind, handelt. Christensen-Szalanski und Willham (1991) kamen nach einem Vergleich von Almanachfragen und Episoden zu der Feststellung, daß die Effektgröße im ersteren Fall dreimal größer ist. Es ist anzunehmen, daß die Enkodierung komplexer Sachverhalte, wie z. B. die Beschreibung des Krieges zwischen Briten und Gurka in

Nepal (Fischhoff, 1975; Wasserman *et al.*, 1991) wesentlich elaborierter erfolgt als bei singulären Fakten. So wäre es denkbar, daß die möglichen alternativen Ausgänge schon während des Lesens der Geschichte antizipiert wurden und daher der Spielraum für eine verfälschende Wirkung des Ankers sehr viel kleiner wurde. Ergo sollte der Rückschaufehler klein ausfallen. Auch Fischhoffs (1975) Interpretation des Rückschaufehlers als einer sich „einschleichenden Gewißheit“ („*creeping determinism*“) zielt auf Enkodierungsprozesse: Die Wichtigkeit einzelner Informationen der Episode wird im Lichte des jeweiligen Ausgangs anders beurteilt und die Kausalstruktur der Geschichte ändert sich grundlegend.

Wasserman, Lempert und Hastie (1991) fanden weitere Evidenz für Fischhoffs „*creeping determinism*“. Sie fanden heraus, daß ein kritischer Faktor für die Stärke des Effektes der kausale Zusammenhang zwischen dem Ausgang und seinen möglichen Ursachen ist. Sobald ein solcher Zusammenhang herstellbar ist, findet nach Ansicht der Autoren eine Rekonzeptualisierung des Ereignisses derart statt, daß der Ausgang mit seinen kausalen Ursachen verknüpft wird und somit ein fester Bestandteil des Ereignisses selbst wird. Hingegen würden Zufallsfaktoren mit großer Wahrscheinlichkeit nicht zu einer solchen Rekonzeptualisierung führen. Wasserman *et al.* manipulierten das Material von Fischhoff (1975) dergestalt, daß sie als Begründung für den Sieg der Engländer über die Gurkas einmal die überlegene Disziplin der Truppen (deterministischer Faktor) oder einen plötzlichen, überraschenden Regensturm (Zufallsfaktor) angaben. Sie erzielten damit einen stärkeren Effekt für die deterministischen Begründungen. Pohl, Ludwig und Ganner (1999) fanden keinen Effekt der Elaboration: Eine Elaboration der Schätzung führte vermehrt zu korrekten Erinnerungen, zeigte aber keinerlei Einfluß auf die Höhe des Rückschaufehlers.

Da sowohl prognostische Urteile wie auch Fallgeschichten von oft sehr hoher praktischer Relevanz sind, folgt hier ein thematisch geordneter Einschub zur praktischen Bedeutung des Rückschaufehlers.

Politische Wahlen. Wie in der gerade erwähnten Studie von Fischhoff fand auch Leary (1981,1982) in einer Befragung vor und nach den Präsidentschaftswahlen von 1980 in den USA und zu einem Football Spiel die üblichen Effekte eines Rückschaufehlers. Der Effekt wurde in keinsten Weise durch das persönliche Engagement der Probanden an den Wahlen oder dem Spiel moderiert. Leary schloß daraus, daß motivationale Prozesse den Effekt nicht erklären könnten. Allerdings behaupten moderne Theorien der sozialen Motivation, daß ein vorteilhaftes Selbstbild aufrecht-

zuerhalten unabhängig von dem Engagement sei. Auch Synodinos (1986) fand in einer Untersuchung zu den Wahlen zum Gouverneursamt 1982 in Hawaii keinen Zusammenhang zu Selbstachtung und politischem Engagement. Er schloß daraus, daß auch er mit seiner Untersuchung nur die Robustheit des Effekts nachgewiesen habe und das motivationale Faktoren keine Rolle spielten. Powell (1988) fand in einer Befragung zu drei lokalen und nationalen Wahlen im November 1984 einen stabilen Rückschafehler. Es fiel ihm auf, daß besonders in den Konfidenzurteilen große Effekte zu finden waren. Er fand, daß auch die anderen Urteile der Probanden von deren Gefühl der Sicherheit abhingen. Tufte und Sun (1975) fanden heraus, daß auch professionelle Meinungsforscher einem Rückschafehler erliegen, wenn sie versuchen, die Wahlergebnisse durch Auswahl bestimmter Schlüsselbezirke nachträglich vorhersagbar zu machen. Eine statistische Analyse, die die vorhersagende Kraft zufällig ausgewählter Bezirke mit denen der Schlüsselbezirke verglich, fand keine Unterschiede.

Evaluation von Entscheidungen. Es ist wichtig, zwischen der Qualität und der Genauigkeit eines Urteils zu unterscheiden (Edwards, 1984). So sollten Informationen, die erst nach dem Urteil verfügbar wurden, nicht zur Beurteilung der Güte des Urteilsprozesses herangezogen werden (Baron & Hershey, 1988). Hingegen wiesen beispielsweise Mitchell und Kalb (1981) einen deutlichen Verstoß bei Prüfern nach, die die Tätigkeiten einer Krankenschwester beurteilen sollten: So wurden die gleichen Tätigkeiten unterschiedlich beurteilt, je nachdem, ob der Ausgang positiv oder negativ war. Ähnliche Ergebnisse fand auch Buchman (1985), der Probanden die Fakten einer Firma und die Wahrscheinlichkeit für einen Bankrott beurteilen ließ, bei einer Firma, die ihren Bankrott erklärt hatte: Auch sie beurteilten die Fakten entsprechend. Baron und Hershey (1988) fanden in fünf Experimenten einen Zusammenhang zwischen dem Ausgang und der Beurteilung der Qualität der zu beurteilenden Entscheidungen, der Kompetenz der Entscheidungsträger und der Bereitschaft, sich von ihnen in monetären Angelegenheiten beraten zu lassen. Den Probanden war zwar klar, daß sie die Informationen über den Ausgang nicht berücksichtigen sollten, konnten es aber dennoch nicht vermeiden. Die Erklärung von Baron und Hershey lautet, daß durch den Ausgang sich die Salienz kongruenter Informationen stark erhöht habe und daher diese Informationen auch eine ungleich höhere Gewichtung erhielten.

Soziale Urteilsphänomene. In der sozialpsychologischen Literatur finden sich ebenfalls viele Befunde, die eine Verwandtschaft mit dem Rückschafehler erkennen lassen. So wird zum Beispiel die Schuldfrage bei Vergewaltigung klar durch den

Ausgang beeinflusst. Es zeigte sich nach Janoff-Bulman, Timko und Carli (1985), daß sowohl die Wahrscheinlichkeiten verschiedener Ereignisse als auch die moralische Bewertung des Opfers und Täters stark vom Ergebnis beeinflusst werden. Snyder und Uranowitz (1978) fanden, daß das Wiedererkennen biographischer Informationen aus dem früheren Leben eines fiktiven Charakters durch seine sexuelle Präferenzen in seinem späteren Leben beeinflusst wurden (alternative Erklärung dieser Befunde bei: Belleza & Bower, 1981; Clark & Woll, 1981). L. Ross, Lepper, und Hubbard (1975; siehe auch: L. Ross, Lepper, Strack & Steinmetz, 1977; Sherman, Skow, Herwitz, & Stock, 1981) fanden, daß einmal gebildete Eindrücke eine erstaunliche Perseveranz zeigen: Probanden beurteilten die Indizien von Abschiedsbriefen und erhielten dazu (falsche) Rückmeldungen. Nachdem die Rückmeldungen als falsch gekennzeichnet wurden, verharrten ihre Selbst-Evaluierungen weiterhin an ihrer alten, diskreditierten Rückmeldung. Die Erklärungen aus der sozialen Urteilsforschung ähneln sehr Fischhoffs (1975) *creeping determinism*: Die Probanden tendierten dazu, die ursprüngliche Rückmeldung mit einem Netzwerk von Inferenzen über Selbstkonzepte und Fähigkeiten zu elaborieren. Ein weiteres Beispiel aus der Sozialpsychologie stellt die Vorhersage von Aktienkursen dar (Fama, 1970). Man kann die Kursentwicklung als einen Random-Walk-Prozeß darstellen und daher spielt die vergangene Preisentwicklung der Aktien keine systematische Rolle. Das allgemeine Vertrauen auf vorherige Aktienkurse kann als ein Beispiel für den Rückschaufehler interpretiert werden. So schreiben Marktbeobachter, die den endgültigen Kurs einer Aktie kennen, fälschlicherweise bestimmten Mustern aus der bisherigen Entwicklung der Aktie eine Bedeutung zu. Die Bedeutung dieser Muster stellt jedoch ein Artefakt des Wissens um den Ausgang dar.

Medizinische Urteile. Eine wichtige praktische Bedeutung hat der Rückschaufehler auf retrospektive und prospektive medizinische Diagnosen. Es könnte auch einer der Gründe für den offensichtlich langsamen Zuwachs medizinischer Expertise sein, daß Diagnostiker nach postdiagnostischem Feedback glauben, es schon immer gewußt zu haben und daher weniger aus der Erfahrung lernten als ihnen möglich wäre. Außerdem werden üblicherweise vor gravierenden Eingriffen unabhängige Zweitdiagnosen eingeholt, deren Unabhängigkeit bei Vorliegen einer Erstdiagnose zumindest stark angezweifelt werden kann. Pennington, et al. (1980) befragten Frauen beim Frauenarzt bezüglich der Wahrscheinlichkeit, daß sie schwanger seien, vor und nach einem Schwangerschaftstest. Es zeigte sich keinerlei Effekt des nunmehr bekannten Ergebnisses. Pennington et al. führen dafür zwei mögliche Gründe an: Zum einen könnten die Erinnerungen an ihre erste Schätzung zu gut

gewesen sein und zum anderen haben die Frauen eine gewisse Ahnung, ob sie schwanger sind. Arkes, Wortmann, Saville und Harkness (1981) fanden einen Rückschafehler bei Medizinerinnen, wenn sie ihnen Fallberichte vorlegten, die mit einem diagnostischen Titel oder ohne einen Titel versehen waren: Die Mediziner ließen sich durch die Titel beeinflussen. Arkes, Faust, Guilmette und Hart (1988) replizierten dieses Ergebnis in einer weiteren Untersuchung mit Neuropsychologen. Es gelang ihnen jedoch, mit einer zweiten Gruppe den Effekt der Rückmeldung ganz zu reduzieren, indem sie ihre Probanden anwies, zu jeder möglichen Diagnose ein Indiz aus der Fallgeschichte anzufügen. Diese Technik des „*counterfactual reasoning*“ hatten zuvor schon Koriat, Löwenstein und Fischhoff (1980) angewendet.

Juristische Urteile. In diesem Bereich hat der Rückschafehler wichtige praktische Implikationen: So ist sowohl den Anwälten als auch den Zeugen einer Gerichtsverhandlung immer der Ausgang eines Falles schon bekannt. Auch sollen beispielsweise Richter oder Geschworene vor Gericht gewisse Informationen bewußt ignorieren. Besonders deutlich konnte Neisser (1981) im Falle von John Deans Zeugenaussagen zur Watergate-Affäre beim Vergleich mit den Bändern einen Rückschafehler nachweisen. Sue, Smith und Caldwell (1973) präsentierten Probanden ein belastendes Telefongespräch, welches eine Gruppe ignorieren sollte, da es als Beweis vor Gericht nicht anerkannt wurde. Sie fanden einen bedeutenden Rückschafehler bei allen Probanden, die das Telefongespräch gehört hatten. Ähnliche Befunde ergaben sich auch in zahlreichen anderen Experimenten (Carrette & Moreland, 1983; Thompson, Fong & Rosenhan, 1981; Werner, Kagehiro, & Strube, 1982; Wolf & Montgomery, 1977).

Der nicht zu ignorierende Einfluß von außergerichtlichen Medienberichten ist in zahlreichen Experimenten nachgewiesen worden (Hoiberg & Stires, 1973; Padawer-Singer & Barton, 1975; Sue, Smith & Gilbert, 1974; Sue, Smith & Perdroza, 1975 und Devine & Ostrom, 1985). Casper, Benedict und Kelley (1988) ließen Probanden einen Fall einer Beschwerde eines Bürgers gegen einen Polizisten wegen unerlaubter Haussuchung beurteilen. Sie fanden, daß abhängig davon, ob sie mitgeteilt bekamen, daß in dieser Wohnung Beweisstücke für illegale Aktivitäten gefunden wurden - etwa Drogen oder Ähnliches - die Probanden auch in ihrer nachträglichen Erinnerung der Fakten sowie in ihrem Urteil beeinflußt wurden. So interpretierten und erinnerten Probanden, welche die Rückmeldung erhalten hatten, daß in der Wohnung tatsächlich Beweisstücke für illegale Aktivitäten gefunden wurden, die Fakten in Konsistenz mit diesem Ausgang, während Probanden, denen mitgeteilt

worden war, es hätten sich bei der Haussuchung keine Beweisstücke finden lassen, die Fakten nun konsistent zu diesem Ausgang interpretierten und erinnerten.

Ähnliche Befunde sind auch aus der Forschung zur Textverarbeitung bekannt, wo z.B. die Aktivierung entsprechend unterschiedlicher Schemata zu entsprechend unterschiedlichen erinnerten Details führte (z. B. Pichert & Anderson, 1978). Auf die Nähe zu Annahmen der Schematheorie etwa, der zufolge Enkodierungsprozesse zwangsläufig eine Abstraktion der zu enkodierenden Inhalte bedingen, haben schon Alba und Hasher (1983) und Kohnert (1996) hingewiesen. Eine spätere Erinnerung bedingt zwangsläufig Rekonstruktionsprozesse, bei denen fehlende Details ergänzt werden. Dazu stehen aber nur allgemeine Heuristiken zur Verfügung, die zu unvermeidlichen Fehlern und Verzerrungen führen. Das hieße, der Rückschaufehler und ähnliche Phänomene markieren die Kehrseite unserer Fähigkeit zum adaptiven Lernen. Sie sind gewissermaßen der Preis für die große Flexibilität unseres Gedächtnisses.

2.2.2 Zur Robustheit des Rückschaufehlers

Fast alle Experimente, die sich mit der Robustheit des Effekts befaßten, benutzten auch Almanachfragen als Material. Fischhoff (1977) interessierte sich, ob Warnungen die Stärke des Effekts modulieren können: Sowohl die Instruktion, die Aufgabe sei extrem wichtig und ihr solle soviel Aufmerksamkeit wie möglich gewidmet werden, als auch der Hinweis, daß in vorangegangenen Experimenten die Tendenz zu erkennen gewesen wäre, daß Personen ihr Wissen angesichts der Lösung stark überschätzten, verbunden mit der Bitte, in dieser Aufgabe diese Tendenz so gut wie möglich zu vermeiden, resultierten in einem gleich starken Rückschaufehler, wie eine Bedingung ohne diese Instruktionen.

Instruktion und Aufklärung Auch die Aufforderung, die Lösungen zu ignorieren, blieb wirkungslos (Bukszar & Connolly, 1988; Fischhoff, 1975; Hasher *et al.*, 1981, Exp. 1; Hennessey & Edgell, 1991). Wie bereits erwähnt, konnten Pohl und Hell (1996) zeigen, daß der Effekt sogar dann auftritt, wenn jedem Probanden individuell der persönliche Bias genau erklärt wurde und man sie aufforderte, in einem folgenden Experiment stärker dagegen anzugehen. Koriat *et al.* (1980) erzielten hingegen mit der Instruktion, Gründe für und wider der eigenen Antwort zu generieren, eine Reduktion des Rückschaufehlers. Davies (1987) fand, daß die Annäherung an den ursprünglichen Gedächtniszustand zum Zeitpunkt der Abgabe

der ersten Schätzung den Rückschafehler deutlich reduziert. Davies stellte seinen Probanden zum Zeitpunkt der Erinnerung Notizen zur Verfügung, die sie bei der Abgabe ihrer Schätzungen angefertigt hatten.

Schul und Burnstein (1985), Wyer und Budesheim (1987) und Wyer und Unversagt (1985) präsentierten ihren Probanden Beschreibung von Eigenschaften und Verhalten einer Person mit der expliziten Instruktion, sich einen Eindruck von dieser Person zu bilden. Die Probanden wurden anschließend gebeten, bestimmte Informationen, die sie erhalten hatten, zu ignorieren. Genau wie bei Bukzar und Connolly (1988) hatten Probanden damit arge Schwierigkeiten. Das Selbstkonzept von Personen ist durch ein starkes Gefühl von Konsistenz bestimmt (Greenwald, 1980). Indes ist dieser Eindruck zeitlicher Kontinuität übertrieben. So sind sich Personen nicht bewußt, in welchem Grad sich ihre Einstellungen über die Zeit hinweg verändern (Bem & McConnel, 1970; Goethals & Reckman, 1973; M. Ross, McFarland, & Fletcher, 1981). Die Personen integrieren einstellungsrelevante Informationen in ihr Anschauungssystem. Anschließend werden ihre Urteile und Meinungen durch ihre gegenwärtigen Anschauungen bestimmt (siehe M. Ross & Conway, 1986). Hier sind die Parallelen zu Fischhoffs (1975) *creeping determinism* deutlich.

Wahrheit. Hasher, Attig und Alba (1981) diskreditierten die Lösung. In einem ersten Experiment erfuhren die Probanden, daß die Versuchsleitung sich vertan hätte und sie soeben ein falsches Lösungsblatt erhalten hätten: Nicht die ersten 10 Aussagen wären falsch und die zweiten 10 wahr, sondern genau umgekehrt, die ersten 10 seien wahr und die zweiten 10 falsch! Die *vertauschten* Lösungen waren aber genauso wirksam wie in der Kontrollgruppe die *nicht-vertauschten* Lösungen. Im zweiten Experiment wurde den Probanden entweder mitgeteilt, daß die Lösungen in einigen Fällen anders als angegeben wären (d. h. wahr statt falsch und umgekehrt), oder es wurde ihnen lediglich gesagt, daß die dargebotenen Lösungen falsch gewesen wären. Im Vergleich zu einer Kontrollbedingung (ohne Diskreditierung) war der Rückschafehler im ersten Fall reduziert, im zweiten Fall sogar völlig eliminiert. Die Wirksamkeit dieses Vorgehens spricht gegen eine automatische Assimilation der Lösung. Das Experiment wurde von Erdfelder und Buchner (1998; Experiment 3) mit den gleichen Befunden repliziert. Konsistent zu den Schlußfolgerungen Hashers et al. schließen auch Erdfelder und Buchner, daß die Diskreditierung des Ankers zu einer Reduktion der Interferenz führt und damit die Erinnerungsleistung ansteigt.

Ein kürzlich durchgeführtes eigenes Experiment (Eisenhauer, 2000), in dem die Diskreditierung des Ankers (wahre oder falsche Lösung eines anderen Probanden) mit und ohne Zeitlimit untersucht wurde, bestätigte die Befunde von Hasher et al. und Erdfelder und Buchner. Auch hier zeigte sich, daß bei als falsch bezeichneten Anker der Effekt ausblieb. Zudem zeigte sich keine Interaktion mit der Zeitbeschränkung. Dies gibt einen Hinweis darauf, daß der evaluative Prozeß relativ wenig Zeit in Anspruch nimmt und eventuell automatischen Charakter hat. Gemeint ist, daß hier eventuell eine sehr schnelle Entscheidung über die Nützlichkeit, der Enkodierung des Ankers vorausgehen könnte. Diese Schlußfolgerungen müssen jedoch noch als reichlich spekulativ gelten, da das Intervall von der Präsentation des Ankers bis zur Warnung, lediglich 2 Sekunden betrug. Möglicherweise war die Zeit zu kurz und der Anker wurde bis zum Zeitpunkt der Warnung noch nicht enkodiert.

In der Metaanalyse von Christensen-Szalanski und Willham (1991) wird die Wahrheit der zu beurteilenden Aussage als eine der beiden wichtigsten Moderatorvariablen identifiziert. Die Autoren geben eine Korrelation von $r = .34$ zwischen der Höhe des Rückschaufehlers und dem Wahrheitswert ($\alpha < .001$) an. Dieses Ergebnis belegt die Bedeutung der Lösungsevaluation: Warum sollte man sich Aussagen merken, die falsch sind?

Schwierigkeit Die zweitwichtigste Moderatorvariable nach der Wahrheit stellt laut Christensen-Szalanski und Willham (1991) die Schwierigkeit des Materials dar. Die Autoren sprechen von der Vertrautheit (*familiarity*) des Materials, betonen aber die verwandte Bedeutung: Je vertrauter ein Item, desto leichter sollte seine Beantwortung sein (und umgekehrt). Sie ermittelten eine Korrelation von $r = .31$ zwischen Rückschaufehler und Vertrautheit ($\alpha < .001$). Der Rückschaufehler ist auch tatsächlich bei leichten Items kleiner als bei schwierigen (Dawson, Arkes, Siciliano, Blinkhorn, Lakshmanan, & Petrelli, 1988; Eller, 1984; Fischhoff, 1975, 1977; Fischhoff & Beyth, 1975; Hoch & Loewenstein, 1989; Wood, 1978, Exp. 2). Hier ist allerdings einzuwenden, daß dieser Befund ein Artefakt der statistischen Auswertung sein kann, denn bei leichten, d. h. eher richtig beantworteten Items ist der Spielraum für Verfälschungen natürlich geringer als bei schwierigen, eher falsch beantworteten (vgl. Hoch & Loewenstein, 1989, S. 611).

Überraschung Ein besonderer Fall liegt bei sehr schwierigen Items vor: Hier kann es zu Überraschungen kommen, wenn der Anker bekannt gegeben wird. Hoch und Löwenstein (1989) stellten einen Zusammenhang zur Schwierigkeit der Aufgabe her:

Je schwieriger das Material, desto geringer wurde der Effekt des Ankers, bis zur Umkehrung. Hoch und Löwenstein interpretieren ihre Befunde so, daß überraschende Anker zur Reaktion „Ich kann es unmöglich gewußt haben“ und somit zu einem entgegengesetzten Rückschleifer führen. Eine wichtige und subtile Frage bezieht sich auf den Zusammenhang zwischen dem Gefühl der Überraschung und der Evaluation und Modellveränderung. In welchem Ausmaß ist Überraschung ein Nebeneffekt der Schwierigkeit, den Ausgang in die personenspezifischen generellen Wissensstrukturen zu integrieren?

Plausibilität Eine weitere wichtige Metakognition stellt die Plausibilität des Ankers dar. So zeigte die subjektive Plausibilität der Anker starke Effekte (Pohl, 1998, 1999; Strack & Mussweiler, 1997): Während subjektiv gute Anker zu maximalen Ankereffekten führten (die erinnerten Schätzungen orientierten sich nahezu ausschließlich am Anker), traten bei subjektiv schlechten Ankern keine oder sogar Kontrasteffekte auf (die erinnerten Schätzungen waren weiter vom Anker entfernt als die tatsächlichen; s. auch Schwarz & Bless, 1992).

Die Befunde lassen sich folgendermaßen zusammenfassen, daß erstens die Angabe einer Lösung drastisch die Suche nach Alternativen einschränkt, zweitens, die Diskreditierung der Lösungen als falsch den Ankereffekt aufhebt, und drittens Metaprozesse den Effekt moderieren.

3 Theorien zum Rückschaufehler

3.1 Adaptives Lernen und Motivation

3.1.1 Rückschaufehler und adaptives Lernen

Für den theoretischen Psychologen betreffen die bedeutsamsten Auswirkungen der Forschung zum Rückschaufehler das Verhältnis zwischen dem Rückschaufehler und adaptiven Lernen aus Erfahrung. Die drei Mechanismen, die Hawkins und Hastie (1990) als die primären Quellen der Verfälschung ansehen, selektives Gedächtnis, Evaluation und Integration und Änderungen des Modells, stellen auch die Hauptmechanismen adaptiven Lernens und kompetenten Urteilens dar. Auf eine Weise stellt der Rückschaufehler die Kehrseite erfolgreichen Lernens und Urteilens dar.

Erstens sind selektive Aufmerksamkeit für Ereignisse aus der externen Welt und *selektiver Abruf* aus dem Gedächtnis notwendige Charakteristika eines adaptiven Systems in einer komplexen Umgebung. Auch die optimistischste Vision moderner Computer-Forscher bestreitet nicht, daß das menschliche Gehirn ohne dramatische Selektion, Abstraktion und Vereinfachung die Flut an Informationen verarbeiten könnte. Hawkins und Hastie vermuten, daß in den meisten natürlichen Kontexten mit engen Restriktionen auf der Menge an Informationen, die tief verarbeitet werden kann, eine durch den Ausgang begrenzte Informationsaufnahme, wie man sie in Experimenten zum Rückschaufehler und in der Entscheidungsforschung beobachten kann, eine effiziente und erfolgreiche Strategie ist, um Informationen zu selektieren. Möglicherweise ist diese Strategie sogar der normative Regelfall in den meisten natürlichen Urteilsaufgaben (siehe Anderson, 1983). Diese Überlegung stellt auch den zentralen Ausgangspunkt vorliegender Arbeit und zweier kognitiver Erklärungsmodelle des Rückschaufehlers dar: Eine Bevorzugung von Informationen durch Erhöhung ihrer späteren Abrufwahrscheinlichkeiten wurde zuerst von Pohl (1993) und Pohl und Eisenhauer (1997) als zentrale Erklärung des Rückschaufehlers eingeführt. Die *selektive Aktivierung* ist das Grundprinzip, nach dem in SARA (Pohl, Hardt & Eisenhauer, in press.; siehe auch Kap. 3) die verzerrende Wirkung eines Ankers erklärt wird. Eine ähnliche Konzeption – die *selektive Erhöhung der Zugänglichkeit* – wurde von Strack und Mussweiler (1997) zur Erklärung von Ankereffekten vorgeschlagen.

Zweitens wird ein Entscheidungsträger angesichts einer großen wechselseitig voneinander abhängigen, beziehungsreichen, ambigen und widerspruchshaltigen Informationsmenge danach trachten, diese zu vereinfachen und zu ordnen. Dieser Evaluationsprozeß beinhaltet das Bilden eines Modells der kausalen Zusammenhänge. Dieses Modell stellt den Mediator zwischen den Rohdaten und der endgültigen Entscheidung dar. Auch Fischhoffs (1975) *creeping determinism* kann verstanden werden als ein Verstehens- und Inferenzprozeß, um Informationen über einen Ausgang in das mentale Modell des Entscheidungsträgers zu integrieren. Der Rückschaufehler und die Unmöglichkeit, das ursprüngliche mentale Modell vor Bekanntgabe der Lösung wiederzuerlangen, ist damit ein Nebeneffekt der gleichen Prozesse, die im Dienste der Erlangung von Genauigkeit, Reliabilität und Expertise in komplexen Entscheidungssituationen stehen.

Drittens ist die Veränderung des mentalen Modells vermutlich der primäre Mechanismus induktiven Lernens (z.B.: Holland, Holyoak, Nisbett, & Thagard, 1986). Mit wachsender Expertise in komplexen Entscheidungssituationen wird auch eine generische Urteilsstrategie gebildet (z.B.: Slovic & Lichtenstein, 1971), welche zur Gewichtung und Kombination von Cues genutzt wird, um akkurate und reliable Entscheidungen in spezifischen Fällen treffen zu können. Der Rückschaufehler ist wiederum ein Nebenprodukt der adaptiven Anpassung der Urteilsstrategie.

3.1.2 Motivationale Faktoren beim Rückschaufehler

Selbstdarstellung. Die Vermutung, daß der Effekt zumindest teilweise auf den Motiven beruht, intelligent, kenntnisreich und scharfsinnig zu erscheinen, lag auf der Hand. So wurden Fischhoffs ursprüngliche Experimente (1975) dahingehend interpretiert, daß die Probanden ihre Schätzungen derart korrigiert hätten, um so zu erscheinen, als ob sie die Zukunft vorhersagen könnten. Hawkins und Hastie (1990) glauben, daß den Selbstwert steigernde Motive unter bestimmten Konditionen eine Rolle spielen (z.B.: wenn die ursprüngliche Güte der Schätzung im Fokus steht) und daß dieser Prozeß erst zu einem sehr späten Zeitpunkt bei der Generierung einer Antwort zum Tragen kommt. Allerdings läßt sich der Rückschaufehler nicht als Artefakt darstellen. Der beobachtete Effekt motivationaler Faktoren ist zwar nicht vernachlässigbar, aber sehr klein. Leary (1981, 1982) findet beispielsweise keinen Zusammenhang zwischen der Beteiligung des Selbstwertgefühls und dem Rückschaufehler. Fischhoff und Beyth (1975) und Wood (1978) versuchten, motivationale Faktoren durch die Instruktion auszuschließen (z.B.: „schätze die

Antworten von Probanden, die keine Lösung erhalten haben“) und fanden dennoch einen Rückschaufehler.

Einige Forscher haben die Rolle selbstwertdienlicher Motive untersucht: Fischhoff (1977) instruierte seine Probanden, „hart zu arbeiten“, um ihre ersten Schätzungen zu reproduzieren, ohne jedoch den Rückschaufehler dadurch schmälern zu können. Hell et al. (1988) stellten Unterschiede in der Memorierbarkeit der eigenen Schätzung in Abhängigkeit vom Intervall zwischen Schätzung und Prüfung fest. Sie fanden außerdem eine komplexe Interaktion zwischen Memorierbarkeit und Motivation und stellten fest, daß schlechte Memorierbarkeit eine Grundvoraussetzung für den Effekt sei, aber als zweiter Faktor Motivation den Effekt durchaus moderieren kann. Ein hoch motivierter Proband wird auch einen verringerten Effekt des Rückschaufehlers zeigen. Campbell und Tesser (1983) fanden eine positive Korrelation zwischen dem Bedürfnis nach Vorhersagbarkeit und dem Bedürfnis nach Selbstdarstellung und dem Ausmaß des Rückschaufehlers. Allerdings erklären diese individuellen Differenzen in der Motivation nur sechs Prozent der Varianz. Bisher wurde nur in Studien mit Almanachfragen als Material ein Effekt von motivationalen Faktoren gefunden, wenn auch ein nur sehr geringer.

Motivation versus Kognition Eine scharfe Trennung motivationaler und kognitiver Zugänge lehnen Hawkins und Hastie (1990) ab. So führen beispielsweise auch Campbell und Tesser (1983) individuelle Unterschiede in der Selbstdarstellungsmotivation und der Motivation zur Vorhersagbarkeit und zusätzliche kognitive Mechanismen an, um das Verhalten ihrer Probanden zu deuten. Hell et al. (1988) schließen ebenfalls, daß sowohl motivationale als auch kognitive Faktoren in Betracht gezogen werden müssen. Hawkins und Hastie denken, daß zusammengesetzte, multifaktorielle Erklärungen notwendig sind, um die ganze Breite der bisher betrachteten Verhaltensvariationen zu erfassen.

Zum anderen sind alle bisherigen Zugänge ein Hybrid motivationaler und kognitiver Faktoren. Alle kognitiven Zugänge beinhalten explizite und implizite Annahmen über die Verarbeitungsziele: Zum einen stellt kognitive Ökonomie ein wichtiges Ziel dar - Personen versuchen, die Belastung ihres Gedächtnisses zu minimieren. Oder Widerspruchsvermeidung (Meinungskonsistenz) - Personen streben eine Maximierung der Konsistenz zwischen verknüpften Anschauungen an, d.h. eine Minimierung der Widersprüche bei gleichzeitiger Maximierung der Gemeinsamkeiten verknüpfter Anschauungen. Schließlich ist auch externale Übereinstimmung ein denkbare und plausibles Ziel kognitiver Ansätze - Personen

streben eine hohe Übereinstimmung zwischen ihren Anschauungen und den tatsächlichen Zuständen in der realen Welt an. Auf der anderen Seite beinhalten auch die motivationalen Faktoren Selbstdarstellung, Bedürfnis nach Vorhersagbarkeit Annahmen über kognitive Prozesse. Um den Bedürfnissen nachzukommen, muß eine Person Entscheidungen treffen, wie sie sich verhalten muß, um ihr Bedürfnis zu befriedigen. Das erfordert zweifellos kognitive Verarbeitung.

Das Konzept der Motivation hat schon immer eine prominente Rolle in der kognitiven Forschung gespielt (z.B. Newell & Simon, 1972). Die Beschreibung der Mechanismen sollte verfeinert werden, Hypothesen sollten gebildet werden, wann die zahlreichen Kombinationen der Mechanismen in Urteilsaufgaben zur Anwendung kommen, und es sollten Untersuchungsmethoden entwickelt werden, um zwischen den verschiedenen Mechanismen mit empirischen Daten differenzieren zu können.

3.2 Kognitive Erklärungen

Kognitive Erklärungen des Rückschaufehlers suchen die Ursache des Rückschaufehlers in der Struktur oder den Prozessen des kognitiven Systems. Die verschiedenen kognitiven Ansätze unterscheiden sich wesentlich darin, ob der Rückschaufehler durch Prozesse erzeugt wird, die zur Zeit der Verarbeitung der Ankerinformation ablaufen, oder durch Prozesse, die erst zur Zeit der Abgabe des Rückschaurteils vor sich gehen. Die eine Gruppe von Erklärungsansätzen betrachtet den Rückschaufehler und auch den Ankereffekt hauptsächlich als ein Gedächtnisphänomen (z.B. Fischhoff, 1975, 1977; Hell et al., 1988; Loftus, 1975; Loftus & Loftus 1980; Pohl 1993; Pohl et al. in press). Die anderen Erklärungsansätze sehen vor allem in Urteilsprozessen den Schlüssel zur Erklärung des Effekts (z.B. Erdfelder & Buchner, 1998; Hell 1993; Jacowitz & Kahneman, 1995; Kahneman, 1992; McCloskey & Zaragossa, 1985; Stahlberg et al., 1993; Stahlberg & Maass, 1998; Strack & Mussweiler, 1997).

3.2.1 Der Rückschaufehler als Gedächtnisphänomen

ERINNERUNGSVERMÖGEN

Die einfachste Antwortstrategie wäre eine Suche des Langzeitgedächtnisses. Fischhoff (1977) und Wood (1978) schätzten, daß ihre Probanden etwa zwei Drittel ihrer ursprünglichen Antworten erinnern konnten. Hasher et al. (1981) berichtet, daß ein sehr überzeugender Grund zur Ignorierung der Rückmeldungen, zu unbeeinflusster Erinnerung führt. Auch Verplanken und Pieters (1988) interpretieren ihr Ergebnis, daß Personen mit einem stärkeren Erkenntnisbedürfnis (high need for cognition) gleiche Ratings in Vorhersage und Rückschau zeigen, so, daß eine Suche im LZG keinen Rückschaufehler verursachen kann. Hell et al. betonen, daß ein Vergessen der ursprünglichen Schätzung eine notwendige Bedingung sei, um einen Rückschaufehler zu erhalten.

Hawkins und Hastie (1990) glauben, daß die direkte Suche im LZG unter bestimmten Umständen auftritt. Beispielsweise, wenn ein vorhergehendes Urteil explizit geäußert wurde und wo die Rückmeldung keine kausalen Elaborationen fördert, die diese mit der ursprünglichen Schätzung verbindet (z.B.: Almanachfragen). Oder, wenn das erste Urteil sehr distinkt und von großer persönlicher Relevanz ist (z.B. bei Schwangerschaften; Pennigton et al., 1980). Schwierig ist, daß eine Übereinstimmung zwischen erster Schätzung und Erinnerung zwar eine notwendige Bedingung zum Nachweis von Erinnerungen aus dem LZG sind, aber nicht ausreichend ist, um zweifelsfrei nachzuweisen, daß diese Übereinstimmung auch auf eine Erinnerung zurückzuführen ist.

UNMITTELBARE ASSIMILATION

Die erste kognitive Erklärung des Rückschaufehlers wurde von Fischhoff (1975) unter der Bezeichnung *unmittelbare Assimilation (immediate assimilation)* geprägt. Fischhoff erklärt den Rückschaufehler damit, daß Personen die Ankerinformation unmittelbar nach ihrem Bekanntwerden mit ihrem Wissen über das relevante Ereignis assimilieren. Diese Assimilation geschieht automatisch und ist damit für die Probanden nicht vermeidbar. In der Situation, sich nochmals an sein ursprüngliches Urteil erinnern zu müssen, kann der Proband nur auf das bereits durch den Anker veränderte Wissen zurückgreifen. Es kommt zwangsläufig zu einer Verfälschung.

Der Beurteiler kann sich nur schwer vorstellen, wie es hätte anders ausgehen können. Für diesen Eindruck eines vermeintlich unausweichlichen Fortschreitens in Richtung auf den tatsächlichen Ausgang hat Fischhoff die Bezeichnung der sich „*einschleichenden Gewißheit*“ (*creeping determinism*) geprägt, im Gegensatz zu dem philosophischen Determinismus, dem bewußten Glauben, daß was auch immer passiert, so kommen mußte. Fischhoff (1975) befürwortet eine Bewertung von Informationen im Lichte des Ausgangs: Wird beispielsweise der britische Sieg über die Gurkas propagiert, könnte ein Proband den Ausgang der Geschichte integrieren, indem er semantische Verbindungen in Form von kausalen Relationen in seine Gedächtnisstruktur der Episode aufnimmt (z.B.: verursachte, verhinderte, erlaubte, etc.; siehe z.B.: Miller & Johnson-Laird, 1976; Schank & Abelson, 1977). Der Proband schreibt die Geschichte neu, so daß Anfang und Ausgang logisch miteinander verknüpft werden (siehe Pennington & Hastie, 1988; Trabasso & Van den Broek, 1985).

Die ersten Experimente Fischhoffs (1975) zielten auf die Beeinflussung der Wahrscheinlichkeitseinschätzung alternativer Ausgänge von Ereignissen. Post factum erhöht das Wissen um den Ausgang die Wahrscheinlichkeit des rückgemeldeten Ereignisses. Zur Erklärung greift Fischhoff auf drei Heuristiken Tverskys und Kahnemans (1974) zurück: (a) der Repräsentativitäts-, (b) der Verfügbarkeits- und (c) der Ankerheuristik.

(a) Nach der Repräsentativitätsheuristik führt das mit dem Anker verschmolzene Wissen zu einer konsistenten Informationssuche mit konfirmativer Suchstrategie: Es werden vor allem Informationen gefunden, die den rückgemeldeten Ausgang bestätigen. Dies führt natürlich auch zu einer Verschiebung der Wahrscheinlichkeitseinschätzung dieses Ausgangs. Der rückgemeldete Ausgang wird jetzt als wesentlich wahrscheinlicher wahrgenommen.

(b) Die Verfügbarkeitsheuristik führt zu einem ähnlichen Effekt. Angesichts des mit dem Anker verschmolzenen Wissens sind konfirmative Informationen wesentlich leichter verfügbar und werden demnach auch als erste gefunden. Das Ergebnis ist eine Konsolidierung des rückgemeldeten Ausgangs und ebenfalls eine Erhöhung der eingeschätzten Wahrscheinlichkeit.

(c) Die Ankerheuristik führt direkt zu einer Überschätzung der Wahrscheinlichkeit des rückgemeldeten Ausgangs. Ausgehend von einer maximalen Wahrscheinlichkeit von 1 für den rückgemeldeten Ausgang muß dieser Wert in Richtung auf den

ursprünglichen Wert adjustiert werden. Das mit dem Anker verschmolzene Wissen sorgt dann dafür, daß diese Adjustierung zu früh endet und ein zu hoher Wert genannt wird. Fischhoff (1977) fand allerdings eine mit der Ankerheuristik unvereinbare Asymmetrie des Rückschaufehlers bei wahren versus falschen Aussagen. Als Folge der Ergebnisse lehnte er die Ankerheuristik als Erklärung für seine früheren Befunde ab: Fischhoff schlägt als Erklärung eine unterschiedlich tiefe Verarbeitung vor, bei der die zugehörigen Inferenzprozesse je nach Verarbeitungstiefe viel oder wenig Einfluß zeigen.

Damit wird deutlich, daß Fischhoffs Erklärung aus zwei Komponenten besteht: Erstens einer *automatischen Integration* der Ankerinformation in die Wissensbasis und zweitens einem *selektiven Informationsabruf*, für die Fischhoff die Heuristiken Tverskys und Kahnemans (1974) vorsieht. In einer weiteren Arbeit präziserte Fischhoff (1977) seine Vorstellung des *selektiven Informationsabrufs*. Fischhoff nimmt an, daß die Integration des Ankers neben einer Reinterpretation der Informationen auch zu einer Stärkung assoziativer Verknüpfungen ankerkonsistenter Informationen führt. Damit liefert Fischhoff die Grundlagen für die Theorien der *selektiven Aktivierung* (Pohl, 1993; Pohl & Eisenhauer, 1997), *selektiven Zugänglichkeit* (Strack & Mussweiler, 1997) und die grundlegenden Prozesse in SARA (Pohl et al., in press; siehe Kap. 3).

Nicht erklärbar durch die Annahme einer unmittelbaren Assimilation ist allerdings, daß es nach wie vor separate (wenn auch verzerrte) Erinnerungen an die eigene Prognose und an die Lösung gibt (Pohl, et al., 1999) oder daß eine nachträgliche Diskreditierung der Lösung den Rückschaufehler aufheben kann (Eisenhauer, 2000; Erdfelder & Buchner, 1998; Hasher, et al., 1981).

THEORIEN ZUM FALSCHINFORMATIONSEFFEKT

Die hier vorzustellenden Erklärungsansätze stammen ursprünglich aus der Forschung zum Falschinformationseffekt. Daher folgt hier zunächst ein kurzer Exkurs zum Zeugenaussagen-Paradigma, aus dem der Falschinformationseffekt stammt. Auf die ursprüngliche Formulierung der Überlegung zum Falschinformationseffekt soll dann der Zusammenhang zum Rückschaufehler hergestellt werden.

Falschinformationseffekt. Experimentelle Studien zu Zeugenaussagen folgen meist demselben typischen Ablauf. Er sei hier an einem der bekanntesten Experimente

von Loftus, Miller und Burns (1978) erläutert: Die Probanden im Experiment sahen zunächst eine Diaserie über ein bestimmtes Ereignis (z.B. Dias über den Hergang eines Unfalls). Im Anschluß daran bekamen die Probanden der Experimentalbedingung – eingebettet in eine schriftliche Darstellung des Ereignisses – eine falsche Information über ein Detail, das in der Diaserie zu sehen war (war im Dia ein „Vorfahrt-achten-Schild“ zu sehen, so wurde im Text von einem „Stopp-Schild“ geredet). Probanden der Kontrollgruppe bekamen diese falsche Information nicht. Am Ende wurde die Gedächtnisleistung durch einen Zwei-Alternativen-Wiedererkennenstest geprüft, wobei in Bezug auf die kritischen Items zwischen Original- („Vorfahrt-achten-Schild“) und inkonsistenter Information („Stopp-Schild“) entschieden werden mußte („war in der Diaserie ein „Stopp-Schild“ oder ein „Vorfahrt-achten-Schild“ zu sehen?“). Probanden der Experimentalgruppe wählten signifikant häufiger das falsche, inkonsistente Item („Stopp-Schild“). Loftus et al. erklärten diesen Befund damit, daß zwei Informationen zum selben Ereignis nicht in getrennten Gedächtnisspuren gespeichert, sondern zum Zeitpunkt der Verarbeitung der zweiten Information integriert und als integrierte Information gespeichert würden. Zum Testzeitpunkt sei dann nur noch eine Information (die bei der Integration entstanden ist) abrufbar, die frühere Spur sei verloren. Dieses Ergebnis wird als *misleading postevent information effect* oder kurz als *misinformation effect* bezeichnet (Belli, 1989). Im deutschen Sprachraum hat sich der Begriff „Falschinformationseffekt“ (Sporer, 1993) etabliert.

Substitutionshypothese. Loftus und Loftus (1980) bezogen einen noch extremeren Standpunkt zur Erklärung des Falschinformationseffekts. In ihrem *update and erase*-Modell wird die Erinnerung an die ursprüngliche Information durch die nachfolgende Falschinformation überschrieben, wobei die ursprüngliche Gedächtnisspur vollständig gelöscht wird. Eine anschließende Gedächtnissuche kann dann nur noch die Falschinformation finden, die dann auch gewählt wird. Besonderes Augenmerk wurde deshalb auf den Zugang zur ursprünglichen Information gelegt, da ein solcher Nachweis die Substitutionshypothese falsifizieren würde. In der Tat fanden sich hier erfolgreiche Manipulationen und Loftus, Korf und Schooler (1989) modifizierten ihr *update and erase* – Modell, indem sie nun annahmen, daß keineswegs in jedem Fall eine Substitution durch die Falschinformation stattfinden müsse. Hasher et al. (1981) wendeten die *Substitutionshypothese* als erste auf den Rückschaufehler an: Beim Rückschaufehler werden als Anker in der Regel objektiv korrekte Antworten verwendet. Diese stehen meist im Widerspruch zu den falschen und unsicheren Schätzungen. Die Probanden sollten bestrebt sein, ihr altes falsches Wissen durch diese Informationen zu ersetzen. Beim Abruf können sie dann nur

noch auf ihr Wissen um die Ankerinformationen zurückgreifen. Dies erzeugt den Rückschaufehler.

Hasher et al. (1981) lieferten auch gleichzeitig mit ihren Experimenten zur Diskreditierung des Ankers einen Befund, der durch die *Substitutionshypothese* nicht erklärt werden kann. Ein erfolgreicher Abruf der Schätzung bzw. eine nicht mehr zum Anker verschobene Verteilung von Erinnerungen, nachdem ein Anker präsentiert wurde, ist durch die *Substitutionshypothese* nicht zu erklären. Hell et al. (1988) weisen darauf hin, daß im Gegensatz zum Falschinformationseffekt beim Rückschaufehler kein Widerspruch zwischen der Schätzung und dem Anker herrschen muß. So muß die Schätzung, der Eiffelturm sei 250 Meter hoch, keinesfalls im Widerspruch zu der tatsächlichen Höhe von 320 Metern stehen. Ein weiterer zentraler Unterschied zwischen dem Falschinformationseffekt und Rückschaufehler betrifft den Fokus der Aufmerksamkeit. Während beim Falschinformationseffekt ein nicht zentrales Detail verändert wird, betrifft der Effekt des Rückschaufehlers die zentrale Information. Es ist plausibel anzunehmen, daß die Gedächtnisspur für eine zentrale Information schwerer zu beeinflussen ist als die Spur für ein kaum beachtetes Detail. Pohl (1993) weist weiter darauf hin, daß wenn man - wie es die *Substitutionshypothese* verlangt - die Schätzung durch den Anker ersetzt, es bei numerischen Informationen, wie etwa Schätzungen und Ankern zu Almanachfragen, zum maximalen Fehler käme. Probanden müßten dann immer behaupten, sie hätten exakt die Lösung gewußt! Das ist allerdings sehr unwahrscheinlich!

Fusionshypothese. Die Vorstellung der Fusionshypothese geht davon aus, daß beide Informationen zu einem einzigen Wert integriert werden und dabei ihre spezifischen Charakteristika verlieren. Eine spätere Trennung ist dann nicht mehr möglich. Im Gegensatz zur *Substitutionshypothese* besitzt die verbliebene Information jetzt aber Charakteristika sowohl der originalen als auch der Falschinformation. Von zentraler theoretischer Bedeutung ist jedoch, wann diese Fusion stattfindet. Manche Autoren gehen von einer Fusion zum Zeitpunkt der Enkodierung der Falschinformation aus und sprechen von *representation blends*, während andere die Fusion erst zum Testzeitpunkt während der Antwortformulierung im Gedächtnistest vermuten und von *recollection blends* sprechen (z.B. Metcalfe & Bjork, 1991; Schooler & Tanaka, 1991). Findet die Fusion bereits zum Zeitpunkt der Präsentation der Falschinformation statt, ist prinzipiell eine Trennung der Informationen nicht mehr möglich. Ist diese Trennung prinzipiell möglich, sollte man eventuell noch an die originale Information gelangen können. So unterscheiden Schooler und Tanaka (1991) *composite blends* von *compromise blends*. Erstere enthalten

Informationen, die aus zwei Quellen zusammengesetzt wurden und die prinzipiell wieder trennbar sind, z. B. eine Straßenszene aus einem Dia und ein Verkehrsschild aus einem Fragebogen, während letztere dagegen nur noch einen gemeinsamen, fusionierten Wert enthalten. So konnte beispielsweise Loftus (1977) zeigen, daß Mischerinnerungen, die zwischen Original- und Falschinformation liegen, möglich sind. Probanden sahen als originale Information ein blaues Auto, während die Falschinformation lautete, es handele sich um ein grünes Auto. In der Testphase, in der die Probanden die originale Farbe des Autos auf einem Farbkreis anzeigen sollten, wurde die Farbe des Autos deutlich in Richtung „grün-blau“ verschoben erinnert. Die Übertragung der *Fusionshypothese* auf den Rückschaufehler ist einfach. Die Vorstellung, daß auch beim Rückschaufehler Schätzung und Anker zu einem Wert fusionieren, bedeutete, daß auch hier eine spätere Trennung zwischen Schätzung und Anker nicht mehr möglich wäre.

Auch die *Fusionshypothese* kann nicht die Befunde von Hasher et al. (1981) zur erfolgreichen Diskreditierung des Ankers erklären. Eine nicht mehr zum Anker verschobene Verteilung von Erinnerungen, nachdem ein Anker präsentiert wurde, ist auch durch die *Fusionshypothese* nicht zu erklären. Hell et al. (1988) haben aus der *Fusionshypothese* abgeleitet, daß der Rückschaufehler in seiner Größe nicht von der Länge des Retentionsintervalls zwischen Ankerenkodierung und Erinnerungstest abhängen sollte. D.h. er sollte bei längeren Retentionsintervallen nicht bedeutend kleiner werden. Hell et al. (1988) fanden allerdings eine sehr starke Abhängigkeit des Rückschaufehlers vom Retentionsintervall. Bei einem Intervall von einer Woche zwischen Ankervorgabe und Test war der Rückschaufehler erheblich kleiner als bei einem Intervall von nur wenigen Minuten.

Koexistenzhypothese. Die Koexistenzhypothese besagt, daß sowohl die originale und die Falschinformation nebeneinander im Gedächtnis repräsentiert sind (z.B. Loftus & Loftus, 1980). Der Falschinformationseffekt resultiert dann aus einer unterschiedlichen Zugänglichkeit der beiden Spuren beim Abruf: Die Falschinformation ist leichter zugänglich als die originale Information, da sie auch später enkodiert wurde und wird deshalb im Test auch häufiger gewählt. Auch hier fällt eine Übertragung der Annahmen auf den Rückschaufehler nicht schwer. Die Idee der Koexistenz von Gedächtnisspuren wird auch im *Modell der relativen Spurstärke* von Hell et al. (1988) aufgegriffen.

Der gravierendste Unterschied zwischen dem Falschinformationseffekt und dem Rückschaufehler betrifft jedoch die Testphase des Experiments (siehe Pohl, 1993):

Während es den Probanden beim Rückschaufehler im Gedächtnisdesign eindeutig klar ist, daß sie eine eigene Schätzung abgegeben haben und später dazu eine Information bekommen haben, sind sich die Probanden beim Falschinformationseffekt nicht bewußt, daß ihnen in der schriftlichen Beschreibung des Ereignisses eine Falschinformation untergeschoben wurde. Die Konsequenz ist, daß im Falschinformationseffekt keinerlei Gedächtnissuche stattfindet, solange den Probanden nicht klar ist, daß überhaupt eine zweite Information eingeschleust wurde. Beim Rückschaufehler ist dagegen eindeutig klar, daß es zwei Informationen gab und es kann dann auch eine Gedächtnissuche stattfinden.

MODELL RELATIVER SPURSTÄRKE

Hell et al. (1988) haben als Alternative zu den Modellvorstellungen von Loftus und Loftus (1980) und Fischhoff (1975) eine Erklärung des Rückschaufehlers postuliert, die auf dem Konzept der relativen Stärke von Gedächtnisspuren beruht. Sie übernahmen von Fischhoff die Annahme, daß die vorgegebene Ankerinformation und die Spur der ursprünglichen Antwort automatisch und unmittelbar integriert (assimiliert) werden. Allerdings gehen die beiden Teilinformationen nicht gleich stark in die integrierte Gedächtnisspur ein, sondern in Abhängigkeit der relativen Spurstärke. D.h. wenn die Spurstärke der ursprünglichen Antwort gering ist, z.B. weil sie zeitlich weit zurückliegt, wird die integrierte Spur im Wesentlichen von der Ankerinformation determiniert werden. Dies wird in der Abrufsituation zu einem vergleichsweise starken Rückschaufehler führen. Ist die Spur der ursprünglichen Antwort hingegen stark, dann wird ihr Anteil an der integrierten Spur groß sein und der Rückschaufehler daher gering. Die Studie von Hell et al. (1988) bestätigt diese Vorhersage in zweifacher Weise: Zum einen durch den schon erwähnten Effekt, daß der Rückschaufehler bei einem großen Behaltensintervall zwischen Originalantwort und Ankerinformation stärker war als bei einem kleinen. Zum anderen hatten Hell et al. die Spurstärke der ursprünglichen Antwort dadurch manipuliert, daß die Versuchsperson in einer Bedingung dazu aufgefordert wurde, ihre Urteile zu begründen. Dies sollte zu tieferer Verarbeitung und damit besserem Behalten der Originalantwort führen (vgl. Craik & Lockhart, 1972). Tatsächlich war der Rückschaufehler bei den vorher begründeten Antworten deutlich schwächer als bei den nicht begründeten Antworten. Dieses Ergebnis von Hell et al. entspricht einer ganzen Reihe von Befunden, die zeigen, daß bei tieferer Enkodierung des ursprünglichen Urteils geringere Rückschaufehler auftreten (Arkes et al., 1988; Davies, 1987; Slovic & Fischhoff, 1977) und daß umgekehrt bei stärkerer

Enkodierung der korrekten Antworten der Rückschaufehler größer ausfällt (Wood, 1978).

Das Modell relativer Spurstärke kann zwar einen großen Teil der Befunde zum Rückschaufehler erklären (vgl. Pohl, 1992), scheitert allerdings auch an den Ergebnissen der Studie von Hasher et al. (1981), da es keine explizite Möglichkeit beinhaltet, die einmal integrierte Gedächtnisspur nachträglich zu ändern. Auch die tatsächlichen Verteilungen der rekonstruierten Schätzungen bereiten dem Modell der relativen Spurstärke erhebliche Schwierigkeiten. Aufgrund der gewichteten Mittelung müßten nämlich alle rekonstruierten Schätzungen zwischen der Originalschätzung und der Lösung liegen. Das aber ist nicht der Fall. In typischen Untersuchungen befinden sich bei den Experimentalitems im Mittel nur 36% der Daten in diesem Bereich (siehe Pohl, 1993).

3.2.2 Der Rückschaufehler als Urteilsphänomen

Dem Modell der relativen Spurstärke sehr nahestehend sind auch Überlegungen zur Spureninterferenz (z.B. Hasher et al., 1981; Kohnert, 1996). Nach der Interferenztheorie koexistieren die Gedächtnisspuren von Schätzung und Anker im Gedächtnis. Erst zum Zeitpunkt des Abrufs treten die beiden Spuren in Konkurrenz zueinander. Einer der stabilsten Befunde der Forschung zur Interferenztheorie ist, daß der Einfluß zwischenzeitlicher Information - des Ankers - auf die Erinnerung der ursprünglichen Information - der Schätzung - mit zunehmendem zeitlichen Abstand zwischen dem Zeitpunkt der Bekanntgabe und dem Test sinkt. Dies entspricht exakt den Ergebnissen von Hell et al. (1988). Nach Tulvings Theorie des episodischen Gedächtnisses (Tulving, 1983) können auch die Befunde zu Effekten tieferer Enkodierung problemlos eingebunden werden. Die Vorhersagen aus dem Modell relativer Spurstärke von Hell et al. und aus der Interferenzhypothese sind im Wesentlichen die gleichen. Die beiden Erklärungen unterscheiden sich darin, zu welchem Zeitpunkt die beiden konkurrierenden Informationen miteinander in Interaktion treten: entweder schon während der Verarbeitung des Ankers oder erst beim Abruf. Aufgrund der vorliegenden empirischen Befunde der Rückschaufehlerforschung kann keiner der beiden Ansätze eindeutig vorgezogen werden.

ANCHORING AND ADJUSTMENT

Im klassischen Experiment von Tversky und Kahneman (1974) wurde der *Anker-effekt* demonstriert. Personen sollten eine Antwort auf die Frage geben, wieviele afrikanische Staaten Mitglieder der UNO seien. Bevor sie antworten durften, sollten die Probanden aber ein Glücksrad betrachten, das gespielt wurde und bei einer Zahl zwischen 0 und 100 stehen blieb. Tversky und Kahneman (1974) hatten das Glücksrad so manipuliert, daß es für eine Teilnehmergruppe stets bei der Zahl „10“ stehen blieb, bei der zweiten Gruppe jedoch stets bei „65“. Daraufhin wurden die Probanden gefragt, ob der Prozentsatz afrikanischer Staaten in der UNO über oder unter dem „zufällig“ erhaltenen Wert liege. Anschließend mußten die Probanden eine exakte Schätzung abgeben. Die erste Gruppe nannte im Mittel einen niedrigeren Schätzwert (25%) als die zweite Gruppe (45%), wobei sich beide Gruppen offenbar an dem durch das Glücksrad vorgegebenen Anker orientierten.

Tversky und Kahneman (1974) schlugen als Ursache für die beobachtbare systematische Verschiebung der Schätzungen zum vorgegebenen Anker die *Ankerheuristik* (anchoring and adjustment) vor. Tversky und Kahneman interpretierten die beobachtete systematische Verschiebung als ungenügende Adjustierung der Schätzung von den Ankerinformationen. Demnach gehen Personen vom vorgegebenen Anker aus und adjustieren in die ihrer Meinung nach plausible Richtung. Dabei stoppen sie jedoch zu früh, nämlich am ersten, subjektiv plausiblen Wert (Jacowitz & Kahneman, 1995). Auf den ersten Blick scheint diese Interpretation angemessen, theoretisch ist sie allerdings nicht zufriedenstellend, da sie eher eine Umschreibung des Befunds denn eine Beschreibung seiner Ursachen darstellt (Gigerenzer, 1996). Es bleibt im Unklaren, was einen plausiblen Wertebereich ausmacht. Pohl (1998) konnte außerdem zeigen, daß Ankereffekte auch bei bereits plausiblen Ankern auftraten, wo eine Adjustierung gar nicht hätte auftreten dürfen, während bei völlig unplausiblen Ankern keine Effekte auftraten. Diese Befunde können nicht im Rahmen der Heuristik erklärt werden.

ANTWORTTENDENZMODELL

Auch dieser Erklärungsansatz stammt wie die Substitutions-, Fusions- und Koexistenzhypothesen ursprünglich aus der Forschung zum Falschinformationseffekt. Daher gehe ich auch hier zunächst auf die ursprüngliche Formulierung der

Überlegung zum Falschinformationseffekt ein und erkläre dann den Zusammenhang zum Rückschaufehler.

Nach dem Modell von McCloskey und Zaragoza (1985) sind die Gedächtnisverzerrungen in Experimenten zum Falschinformationseffekt im Wesentlichen das Resultat einer einfachen rationalen Urteilsstrategie der Probanden. Das Gedächtnis für die erste Information wird durch die Falschinformation in keiner Weise beeinträchtigt. Den beobachteten Effekt bezeichneten sie demzufolge als „Artefakt“, da er nichts mit verfälschten oder falsch abgerufenen Gedächtnisinhalten zu tun habe.

Ausgangspunkt ist das Standarddesign des Falschinformationseffekts, in dem als Testalternativen die originale Information und die Falschinformation zur Auswahl vorgegeben werden. McCloskey und Zaragoza machen jedoch die Annahme, daß die Falschinformation die Erinnerung an die originale Information in keinerlei Weise beeinträchtigt. Demzufolge sollte auch der Prozentsatz korrekt erinnelter Information in Kontroll- und Experimentalbedingung identisch sein. Nehmen wir beispielsweise an, es werden sowohl in der Kontroll- als auch in der Experimentalbedingung 50% der Informationen korrekt erinnert. In der Kontrollbedingung, in der keinerlei inkonsistente Information eingeschleust wurde, raten Probanden, d. h. sie wählen mit gleicher Wahrscheinlichkeit die richtige bzw. die falsche Alternative. Damit erreichen sie insgesamt 75% korrekter Antworten. In der Experimentalbedingung allerdings könnte es sein, daß sich ein Teil der Probanden - beispielsweise 20% - an die Falschinformation erinnert und diese dann wählt. Die restlichen 30% der Probanden aus der Experimentalgruppe werden dann wieder raten und sich jeweils zur Hälfte für die richtige Antwortalternative entscheiden. Damit erreicht die Experimentalgruppe insgesamt nur 65% korrekter Antworten, d. h. es liegt ein Falschinformationseffekt von 10% vor! Dieser „Effekt“ beruht aber ausschließlich auf den unterschiedlichen Ratewahrscheinlichkeiten. McCloskey und Zaragoza bevorzugen den Begriff „Antworttendenzen“ (*demand characteristics*), die durch das Testverfahren hervorgerufen werden. Die schlechtere Leistung in der Experimentalbedingung bedarf somit keiner kognitiven Interpretation.

Hell et al. (1988) haben als erste vorgeschlagen, das McCloskey und Zaragoza-Modell auf den Rückschaufehler anzuwenden. In der Experimentalbedingung erinnern sich die Probanden für einen Teil der Items nur ungefähr an ihre Schätzung und rekonstruieren diese mehr oder minder genau. Die rekonstruierten Schätzungen ergeben dann ungefähr eine Normalverteilung um die tatsächlichen Schätzungen. Für einen anderen Teil der Items erinnern sie sich aber nur an die Lö-

sung, ohne zu wissen, ob ihre Schätzung darunter oder darüber lag. Diese von den Lösungen ausgehenden Neuschätzungen bilden dann eine Normalverteilung um die Lösungen herum. Die Summe aus beiden Verteilungen – und das sind die empirisch beobachtbaren Daten – suggeriert eine homogene Verteilung, die zu den Lösungen hin verschoben ist und die mithin als „Rückschaufehler“ interpretiert wird. Es kam aber an keiner Stelle zu irgendeiner Interferenz zwischen Schätzung und Lösung.

Die Konsequenzen aus dieser Überlegung sind evident. Wenn die Gesamtverteilung der Experimentaldaten durch die Addition der beiden beschriebenen Verteilungen erklärt werden kann, ist die Annahme von systematischen verfälschten Gedächtnisinhalten überflüssig. Es ist daher von entscheidender Bedeutung herauszufinden, ob die beobachtbare Verteilung der Experimentaldaten außer auf den beiden erwähnten unbeeinflussten Verteilungen noch auf einer dritten Verteilung beruhen könnte, die aus systematisch durch den Anker verfälschten Erinnerungen besteht. Eine allgemeine (mittlere) Verzerrung der Erinnerungen zu diagnostizieren, reicht also nicht aus. Es müssen die zugrundeliegenden Fälle separiert werden. Um eine Separierung dieser theoretisch postulierten, praktisch aber nicht identifizierbaren Erinnerungstypen zu ermöglichen, wurden multinomiale Modelle verwendet (z.B. Erdfelder & Buchner, 1998; Kohnert, 1996). Das in diesen Untersuchungen verwendete Modell trennt allerdings lediglich zwischen Verteilungen um die Schätzung und solche um die Lösung, wobei die Verteilungen nicht symmetrisch um die Schätzung bzw. Lösung liegen müssen. Dadurch wird die möglicherweise vorhandene dritte Verteilung überdeckt. In einem anderen Modell (z.B. Pohl, 1993; Pohl & Gawlik, 1995) fanden sich dagegen substantielle Hinweise auf eine Verteilung, die auf verfälschten Erinnerungen basierte. Die multinomiale Analyse stellt allerdings nur ein recht grobes Werkzeug zur Separierung bereit. Die Daten werden in nur wenige Kategorien eingeteilt (in Erinnerungen unterhalb der tatsächlichen Schätzung, zwischen tatsächlicher Schätzung und Anker oder oberhalb des Ankers). Die genaue Verteilung der Abweichungen bei den erinnerten Schätzungen geht dabei verloren. So ist es viel plausibler anzunehmen, daß die Rekonstruktion der früheren Schätzungen stets zu variablen Anteilen (von jeweils 0 - 100%) von der Erinnerung an die Schätzung und der an den Anker beeinflusst wird (siehe Pohl et al., in press).

Das *Antworttendenzmodell* ist mit den Daten von Hasher et al. (1981) vereinbar. Eine Diskreditierung der Ankerinformation kann Probanden, die zwar den Anker, aber nicht mehr ihr Urteil abrufen können, dazu ermutigen, den Anker bei der Rekonstruktion aktiv zu vermeiden. In dieser Sichtweise zeigt die Möglichkeit,

Ankerinformation zu diskreditieren, daß es sich beim Rückschaufehler um das Ergebnis eines strategischen Urteilsprozesses handelt. Auch der Befund von Hell et al. (1988), daß der Rückschaufehler sinkt, wenn finanzielle Anreize für exakte Erinnerungen gegeben werden, kann als eine Möglichkeit der Probanden angesehen werden, ihren Urteilsprozeß sinnvoll zu kontrollieren. Das Antworttendenzmodell ist konsistent mit den wesentlichen empirischen Befunden. Es kann sowohl der Einfluß des zeitlichen Abstands zwischen Urteil, Anker und Test wie auch der Effekt tieferer Enkodierung von Urteil oder Anker vorhergesagt werden. Diese Faktoren sollten direkt die Wahrscheinlichkeit der verschiedenen Gedächtniszustände und mittelbar die Wahrscheinlichkeit (oder den Anteil) der Kategorie beeinflusster Erinnerungen verändern.

REKONSTRUKTIONSTHEORIE

Die *Rekonstruktionstheorie* (Stahlberg, et al., 1993; Stahlberg & Maass, 1998) basiert auf dem Antworttendenzmodell von McCloskey und Zaragoza (1985) und postuliert ebenfalls, daß keinerlei Beeinträchtigung des Gedächtnisses angenommen werden muß, um verzerrte Erinnerungen an frühere Urteile zu erklären. Der beobachtete Rückschaufehler entsteht dadurch, daß Probanden, denen die Lösung bekannt ist, die aber vergessen haben, wo ihre Schätzung lag, ihre Rekonstruktion allein auf die Information der Lösung stützen. In der Kontrollbedingung kann dieser Fall nicht auftreten. Stahlberg und Maass (1998) nehmen an, daß der Anker als Abrufschlüssel zum Auffinden der Schätzung genutzt wird. Probanden adjustieren ausgehend vom Anker in Richtung auf die Schätzung und stoppen diesen Adjustierungsprozeß zu früh, da sie ihr Wissen als Folge der Ankerpräsentation überschätzen. Die *Rekonstruktionstheorie* sagt also richtig voraus, daß auch hier der Anteil korrekter Erinnerungen (im Gedächtnisdesign) nicht von der Darbietung eines Ankers beeinflusst ist.

Die genauen theoretischen Vorstellungen über die beteiligten Prozesse in der *Rekonstruktionstheorie* bleiben sehr vage. Dieser Ansatz berücksichtigt auch nicht, wie sich die falsch erinnerten Urteile verteilen, d. h. ob sie sich ausschließlich um die Lösung gruppieren oder ob sie überzufällig häufig zwischen Lösung und Originalschätzung liegen.

SELEKTIVE ZUGÄNGLICHKEIT

Strack und Mussweiler (1997) berufen sich in ihrer Theorie der *selektiven Zugänglichkeit* gezielt auf die theoretische Nähe des Ankereffekts zum Rückschaufehler. Sie gehen davon aus, daß für die Entstehung des Ankereffekts bei dem klassischen Experiment Tverskys und Kahnemans (1974) zur Frage, wieviele afrikanische Staaten Mitglied der UNO seien, zwei Urteilsprozesse verantwortlich zeichnen: Zunächst müssen Probanden in einem *komparativen Urteil* entscheiden, ob ihre Schätzung ober- oder unterhalb des Ankers (der Zufallszahl des Glücksrades) liegt, dann müssen sie in einem *absoluten Urteil* den genauen Wert schätzen.

Strack und Mussweiler übertragen die Situation direkt auf den Rückschaufehler und behaupten, wenn der Anker im Rückschaufehler präsentiert werde, löse dieser zunächst ein *komparatives Urteil* aus. Im nachfolgenden *absoluten Urteil* werde dann die Schätzung genannt. Entscheidend für die Theorie ist, daß mit diesen Urteilsprozessen gleichzeitig auch zwei kognitive Prozesse ausgelöst werden: Zum einen erhöht der Anker die Zugänglichkeit ankerkonsistenter Informationen, zum anderen sorgt der Anker dafür, daß im Sinne einer *positiven Teststrategie* die Hypothese getestet wird, daß der Anker auch plausibel ist. Der Versuch, den Anker positiv als eine plausible Antwort auf eine Frage zu testen, führt dazu, daß selektiv ankerkonsistente Informationen gesammelt werden, die dann bei der Beantwortung des absoluten Urteils dazu führen, daß dieses systematisch in Richtung auf den Anker verfälscht ist. Die Verfälschung findet nach Meinung von Strack und Mussweiler zum Zeitpunkt der Rekonstruktion statt. Der Anker wird als plausibler Wert angesehen und es werden zum Zeitpunkt des absoluten Urteils zum einen Informationen gesammelt, die diese Hypothese stützen, zum anderen bedingt der Anker aber auch als *Abrufreiz* selber, daß vor allem ankerkonsistente Informationen gefunden werden. Beide Prozesse sorgen für eine *selektive Zugänglichkeit* ankerkonsistenter Informationen, die dann zu der Verfälschung in Richtung auf den Anker führen.

Die theoretischen Vorstellungen dieser Theorie sind sehr detailliert und können eine Vielzahl der Befunde zum Rückschaufehler erklären. Die grundlegende Vorstellung selektiver Zugänglichkeit besitzt eine sehr große Nähe zu der Theorie der *selektiven Aktivierung* (Pohl, 1993; Pohl & Eisenhauer, 1997), welche sowohl die Grundlage für das kognitive Prozeßmodell SARA (Pohl et al.; in press) als auch für diese Arbeit darstellt.

Problematisch an der Theorie der *selektiven Zugänglichkeit* ist die unhinterfragte Gleichsetzung des Ankereffekts und des Rückschaufehlers. Sicherlich besitzen beide Phänomene ein sehr hohes Maß an Gemeinsamkeiten, aber es gibt auch wichtige Unterschiede. Zum einen läßt sich der Ankereffekt direkt nur mit dem Rückschaufehler im Urteilsdesign vergleichen, da beim Ankereffekt keine vorherige unabhängige Schätzung verlangt wird. Zum anderen ist das komparative Urteil für die Entstehung des Ankereffekts geradezu konstituierend, während es das für den Rückschaufehler keineswegs ist. Es ist sehr fraglich, ob ein komparatives Urteil überhaupt stattfindet. Die Theorie der selektiven Zugänglichkeit kommt daher nicht ohne zusätzliche Annahmen und Präzisierungen aus.

SELEKTIVE AKTIVIERUNG

Die Theorie soll hier nur ganz kurz dargestellt werden, da sie die zentrale Annahme des Prozeßmodell SARA (Pohl et al.; in press) ist, welches im nächsten Kapitel ausführlich beschrieben werden soll. Im Rahmen dieser Modellvorstellung werden dann auch die Annahmen zur selektiven Aktivierung ausführlich dargelegt und präzisiert. Die Idee der *selektiven Aktivierung* (Pohl, 1993; Pohl & Eisenhauer, 1997) ist in diesem Kapitel schon bei mehreren Theorien angeklungen und fand in der Theorie der *selektiven Zugänglichkeit* (Strack & Mussweiler, 1997) ihren bisher elaboriertesten Niederschlag.

Die Annahmen der *selektiven Aktivierung* beruhen auf grundlegenden Konzepten der kognitiven Psychologie. Der Ausgangspunkt ist die Annahme, daß die beobachtete Verschiebung der Erinnerung in Richtung auf den Anker durch eine *selektive Aktivierung* bestimmter Informationen des pro Item und Person postulierten Wissens verursacht werden. Demnach entfaltet der Anker eine aktivierende Wirkung auf mit ihm assoziierten Informationen im Gedächtnis. Die Aktivierung assoziierter Informationen im Gedächtnis geschieht durch eine automatische Aktivationsausbreitung im semantischen Netzwerk (Anderson, 1983). Die stärker aktivierten Informationen sind in einer anschließenden Gedächtnissuche leichter zugänglich bzw. schneller verfügbar (McNamara, 1992), was zu einem vermehrten Abruf dieser Gedächtnisinhalte und als Konsequenz letztlich zu einer Verschiebung der Antwort in Richtung der dargebotenen Information führt. Die beobachtbare Verteilung der Erinnerungen ist somit erklärbar, ohne daß der Anker als Abrufhinweis wirken muß. Der Rückschaufehler beruhte dann auf einer Veränderung von Gedächtnisinhalten. Zusätzlich kann aber auch der Anker als

Abrufreiz genutzt werden, wenn er gefunden wird, und es wäre somit möglich, den Effekt auch als Urteilsprozeß zu verstehen. Zusätzliche Annahmen wie etwa ein notwendiges *komparatives Urteil* sind nicht nötig

3.3 Abschließende Anmerkung

Der Rückschaufehler kommt im Alltag häufig vor. Die Anziehungskraft dramatischer Literatur beruht zum großen Teil auch auf dem Rückschaufehler. Die klassischen Tragödien von Ödipus, König Lear etc. würden viel an Tragik verlieren, wenn die Darsteller und das Publikum nicht daran glauben würden, *daß sie es schon immer gewußt haben sollten*. Die Tragik der Theorien jedoch ist, daß unabhängig von ihrer fraglichen Adäquatheit allen Theorien eine genaue Beschreibung der kognitiven Prozesse, die zu einer systematisch verfälschten Rekonstruktion führen, fehlt. Wie werden Schätzungen generiert, Anker enkodiert und Erinnerungen an Schätzungen rekonstruiert? In nächsten Kapitel wird ein Modell vorgestellt, das diese Prozesse und die Wissensrepräsentation, auf der sie operieren, detailliert beschreibt: SARA (Pohl et al., in press).

4 SARA

4.1 Einführung

Alle bisher diskutierten Theorien lassen die Frage nach den beteiligten Prozessen, die für die systematische Verschiebung zum Anker verantwortlich zeichnen, offen. Das Ziel dieses Kapitels ist es, dieses theoretische Vakuum zu schließen und mit der Vorstellung von SARA eine detaillierte Beschreibung der beteiligten kognitiven Prozesse zu liefern. Dieses Kapitel greift dabei natürlich auf vieles zurück, was in vorangegangenen Kapiteln bereits ausführlich dargelegt wurde. Es werden daher bereits vorgestellte Überlegungen nur so knapp dargestellt, wie sie zum Verständnis des Modells SARA, um das es hier ja gehen soll, notwendig ist.

Um eine detailliertere Beschreibung leisten zu können, ist es erforderlich, die zu erklärende systematische Verschiebung zum Anker in kleinere Teilprozesse aufzugliedern. Beim Rückschaufehler ergibt sich eine natürliche Gliederung durch das experimentelle Paradigma. Betrachten wir den komplexeren Fall eines Experimentes im Gedächtnisdesign, so wird schon aus dem Ablauf klar, daß sich das Experiment in drei Abschnitte unterteilen läßt: Abgabe der Schätzung; Enkodierung des Ankers und Erinnerung der Schätzung. Es ist daher sinnvoll, jeden dieser Abschnitte einer genaueren Untersuchung zu unterziehen. Wie wird eine Schätzung generiert? Auf welchen Informationen beruht sie? Wie wird die Lösung enkodiert? Welchen Einfluß übt sie auf die Schätzung aus? Und schließlich, wie wird die Schätzung rekonstruiert? Gibt es Gemeinsamkeiten in den beteiligten Prozessen? Gibt es noch weitere Prozesse, die berücksichtigt werden müssen?

4.2 Organisation des Wissens

Ablaufschema Betrachten wir zunächst den Prozeß, der zur Generierung einer Schätzung führt. In einem typischen Experiment mit Almanachfragen sollen Antworten auf schwierige Wissensfragen gegeben werden. Grundsätzlich gibt es jetzt drei Möglichkeiten (siehe Abb. 4.1):

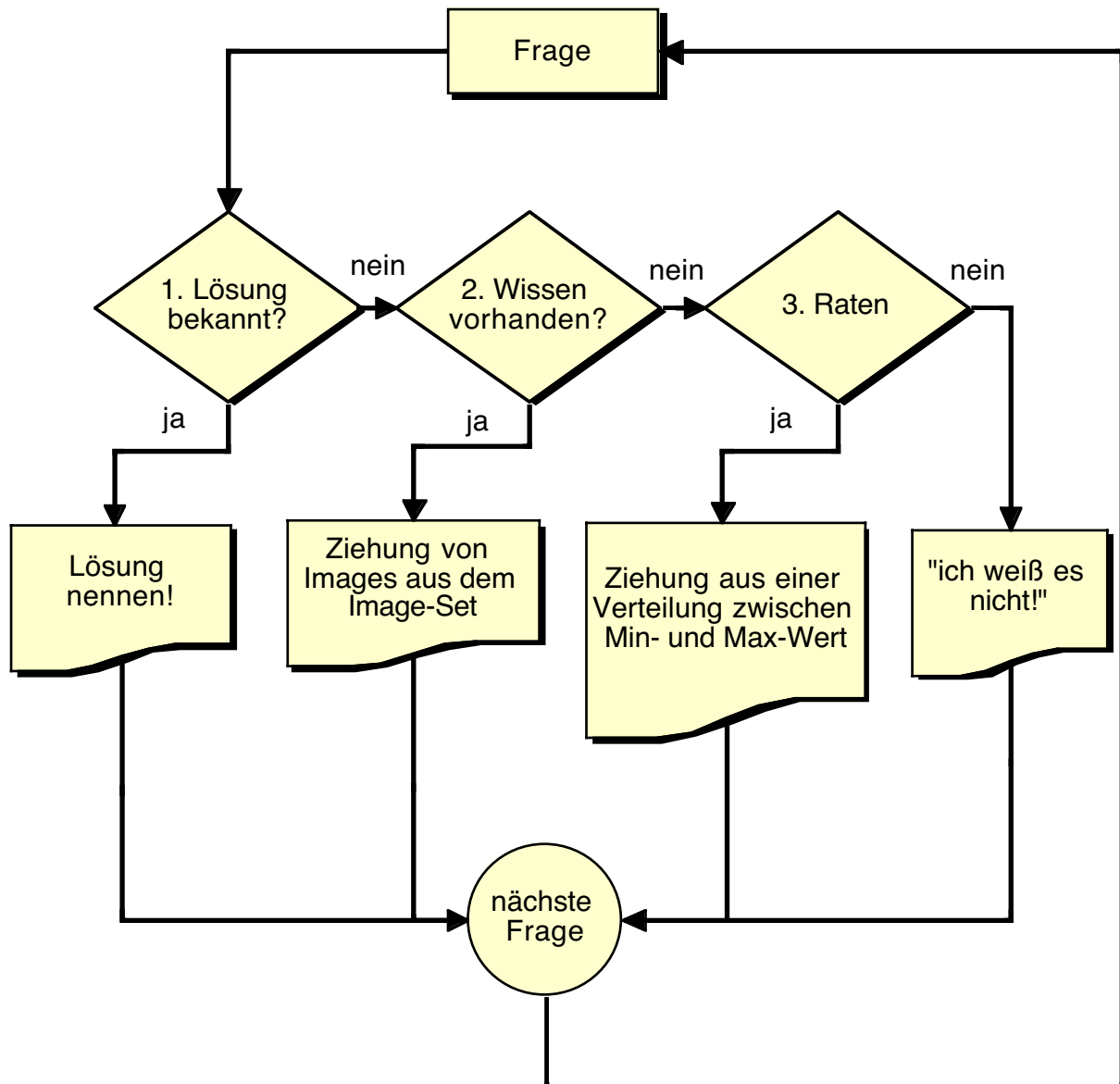


Abb. 4. 1 Schema zur Beantwortung einer Frage.

1. Die Suche nach einer Antwort ist erfolgreich. Die Antwort auf die Frage ist bekannt oder es besteht zumindest die subjektive Gewißheit, daß die abgerufene Information die Antwort auf die Frage ist. In diesem Fall wird natürlich die Antwort genannt und zur nächsten Frage weitergegangen.

2. Die Suche nach einer Antwort ist erfolglos, aber es sind grundsätzlich mehrere inhaltliche Images vorhanden, die helfen, die Antwortsuche einzugrenzen. In diesem Fall wird eine Anzahl Images gezogen werden. Die Images kommen ins Arbeitsgedächtnis und, wie bereits erwähnt, werden in jedem Abrufzyklus die Assoziationsstärken zwischen den Images bzw. zwischen ihnen und den Cues um einen bestimmten Betrag erhöht. Anschließend werden die Werte der abgerufenen

Images zu einer Schätzung integriert und genannt. Dieser Fall ist inhaltlich gesehen der interessanteste Fall.

3. Die Suche nach einer Antwort ist erfolglos, weil entweder keine inhaltlichen Images zur Frage vorhanden sind oder kein Image abgerufen werden kann. In diesem Fall würde man auf die Frage mit „Ich weiß es nicht!“ antworten. Dieser Fall ist aber in Experimenten zum Rückschafehler nicht vorgesehen. Probanden sollen zu jeder Frage eine Antwort generieren. Die Konsequenz ist, daß in diesem Fall geraten wird. Dieses Raten geschieht nach festgelegten Regeln. So ist anzunehmen, wenn Probanden zumindest den Sinn der Frage erfasst haben, sie eine Zahl aus einem plausiblen Bereich nennen werden. In SARA nehmen wir daher an, daß es zu jeder Frage ein Minimal- und Maximalwert gibt, die als Images im Image-Set vorkommen. Dann erfolgt eine Ziehung aus einer Normalverteilung über diesem Bereich.

Organisation Die Grundlage für das Prozessmodell SARA bildet die Annahme von der Organisation des Wissens in einem semantischen Netzwerk (Anderson, 1983) und automatischer Aktivationsausbreitung. Die Organisation des zu betrachtenden Teilbereiches des Netzwerkes geschieht durch die gestellte Aufgabe. Im Falle von Ankereffekt und Rückschafehler bestimmt die Frage, welches Wissen zur Beantwortung oder zur Abgabe einer Schätzung zur Verfügung steht. Für diesen Teilbereich des Wissens, welcher potentiell und interindividuell zur Beantwortung der Frage zur Verfügung steht, wird in SARA der Begriff Image-Set verwendet. Auf die Wissenseinheiten (Images) innerhalb des Image-Sets können elementare kognitive Prozesse zurückgreifen. Die zentrale Annahme von SARA und dieser gesamten Arbeit ist, daß die beobachteten Verzerrungseffekte durch *selektive Aktivierung* bestimmter Informationen verursacht werden (Pohl, 1993; Eisenhauer, 1996). Entscheidend für das Modell ist die aktivierende Wirkung enkodierter Informationen (z. B. die Lösung) auf mit ihnen stark assoziierte Images des Image-Sets. Die so aktivierten Images erhalten eine größere Wahrscheinlichkeit, bei einer nachfolgenden Gedächtnissuche aufgefunden zu werden, (McNamara, 1992). Dies führt zu einer erhöhten Wahrscheinlichkeit für den vermehrten Aufruf von Images, die mit der Lösung assoziiert sind und damit letztendlich zu einer Verschiebung in Richtung auf die dargebotene Information.

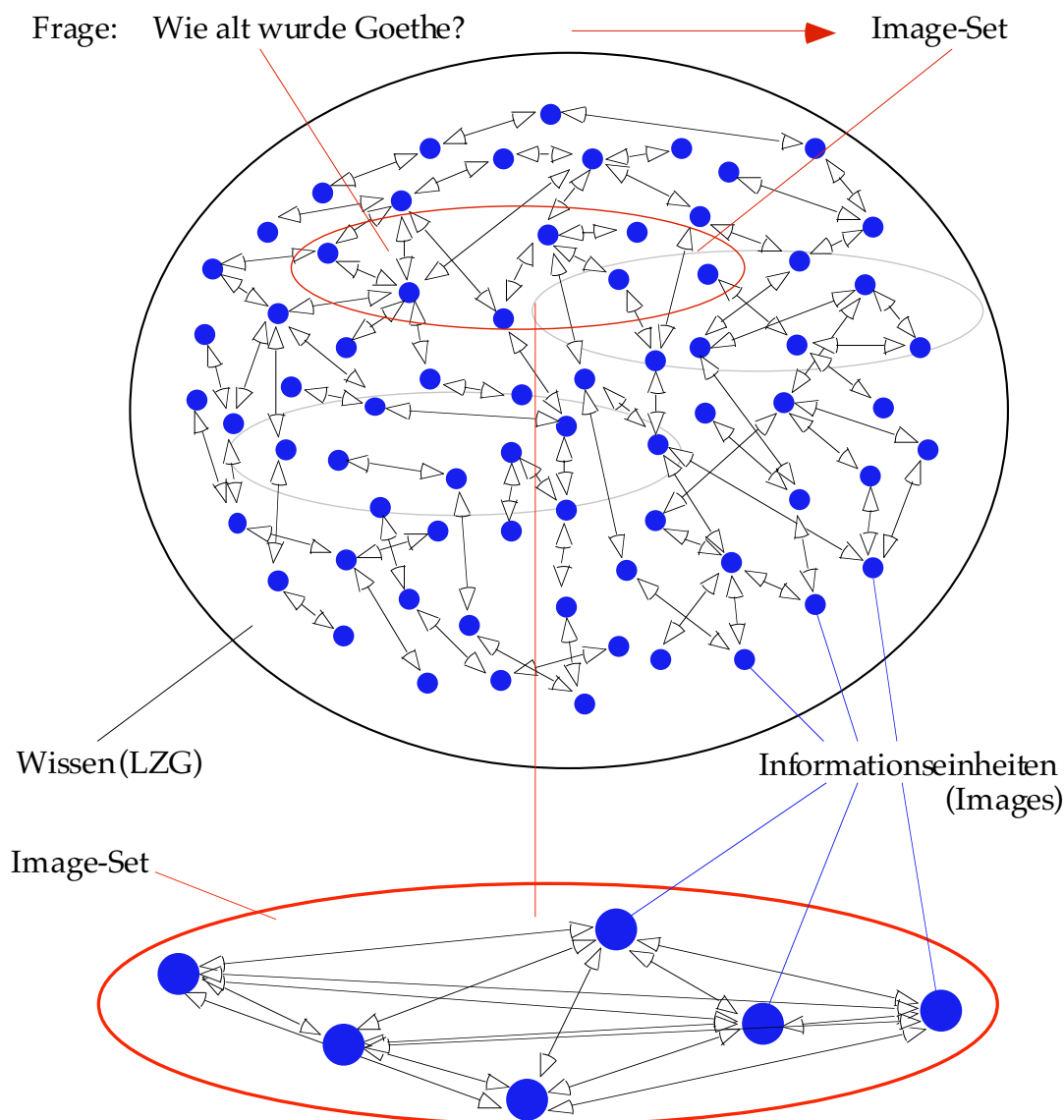


Abb 4.2: Organisation der Informationseinheiten (Images) durch die Frage als Image-Set im LZG.

Architektur In der Organisation des Gedächtnisses lehnt sich SARA an die klassischen Annahmen der kognitiven Psychologie. SARA unterscheidet zwischen einem Langzeitgedächtnis (LZG), in dem sämtliche Informationen gespeichert sind, und einem Arbeitsgedächtnis, in dem erfolgreich abgerufene bzw. zu enkodierende Informationen weiterverarbeitet werden. Das Wissen im LZG wird durch semantische Ähnlichkeit organisiert. Die Kapazität des Langzeitgedächtnisses gilt als unbeschränkt, während die Kapazität des Arbeitsgedächtnisses dagegen auf einige wenige Einheiten limitiert ist. Hier übernimmt SARA die klassischen Annahmen von Miller (1956) und erlaubt nicht mehr als 5 - 9 Informationseinheiten (7 ± 2 Chunks). Im Arbeitsgedächtnis wird zusätzlich die zeitliche Reihenfolge des Eintreffens von Wissensselementen berücksichtigt.

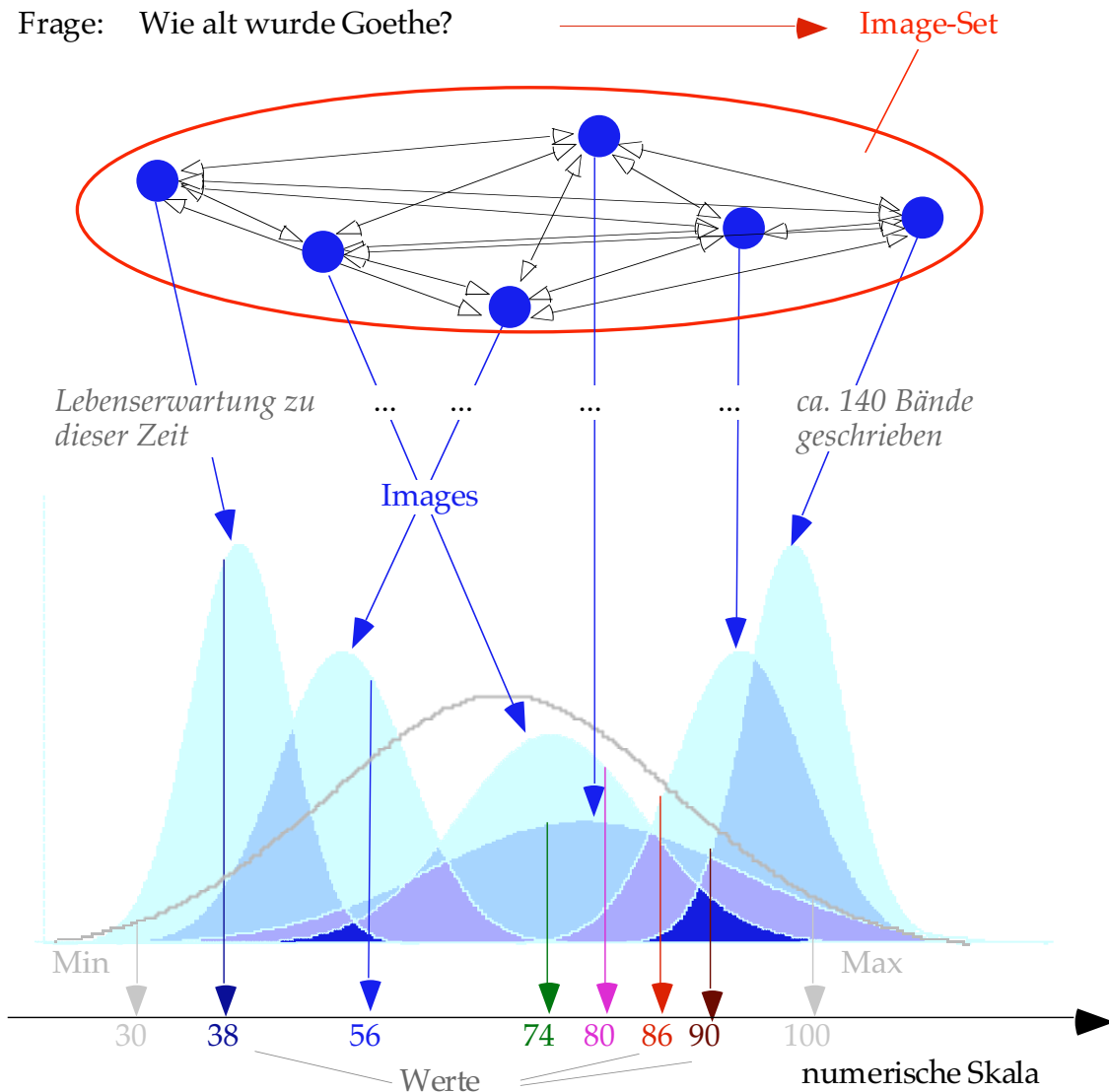


Abb. 4.3: Ziehung der Werte aus den Verteilungen der Images.

Informationseinheiten Wie bereits erwähnt, organisiert in SARA die gestellte Aufgabe das Wissen, auf dem die Prozesse operieren. Die Frage, etwa nach dem Sterbealter Goethes, oder die Aufgabe zu entscheiden, ob Absinth ein Edelstein oder ein Getränk ist, ob die UNO mehr oder weniger afrikanische Mitgliedsländer als eine Zufallszahl hat, bestimmt jeweils den Ausschnitt des Wissens, auf dem die von SARA angenommenen Prozesse operieren sollen. In Abbildung 4.2 wird exemplarisch an der Frage nach dem Todesalter Goethes die Organisation des Wissens durch die Frage verdeutlicht. In SARA wird für diesen Teilbereich des Wissens, welcher potentiell und interindividuell zur Beantwortung der Frage zur Verfügung steht, der Begriff Image-Set verwendet (siehe Abb. 4.2). Die individuell verfügbaren Informationseinheiten des Image-Sets werden als Images bezeichnet. Der Image-Set beinhaltet mindestens subjektiv plausible oder durch die Fragestellung vorgegebene Minimal- und Maximalwerte, die den Lösungsbereich einschränken (siehe Abb. 4.3).

SARA kann nur numerisch verwertbare Information bearbeiten. In Abbildung 4.3 werden exemplarisch die mögliche Verteilung von numerischen Substraten für die durchschnittliche Lebenserwartung zum Ende des 18. Jahrhunderts oder der Information, daß Goethe über 140 Bände verfasst hat, veranschaulicht. Prinzipiell lassen sich mehrere Arten von numerischen Informationen unterscheiden, die ein Image enthalten kann, wie z. B. Schätzwerte mit einem Unsicherheitsintervall, Extremwerte mit Minimum und Maximum oder Plausibilitätsintervalle mit unterer und oberer Grenze. Aus Gründen der Vereinfachung wird angenommen, daß jede dieser Informationen eines Images durch einen konkreten numerischen Wert repräsentiert werden kann. Dies könnte im Falle eines Images, das ein Unsicherheitsintervall enthält, ein zufällig aus dieser Verteilung gezogener Wert sein. Alle Prozesse in SARA arbeiten nur mit diesen konkreten, numerischen Inhalten (im folgenden „Werte“ genannt) (siehe Abb. 4.3).

4.3 Prozesse

Sampling Der grundlegende Prozeß in SARA ist Sampling. Mit Sampling ist ein zyklischer Such- und Abrufprozeß von Images aus dem Image-Set gemeint. Es werden gezielt Informationen im Langzeitgedächtnis gesucht und ins Arbeitsgedächtnis gerufen. Jeder Hinweisreiz (Cue) im Arbeitsgedächtnis ist mit allen Images verknüpft, d. h. jeder Cue nimmt Einfluß auf die Organisation des Image-Sets und damit auf die Resultate nachfolgender Verarbeitungsprozesse. Jeder Verarbeitungsprozeß, den SARA annimmt, beruht auf Sampling. Sowohl bei der Generierung einer Schätzung als auch bei der Enkodierung einer Information (Lösung, Anker) und wieder bei der Rekonstruktion der Schätzung findet Sampling statt.

Abruf von Images Jedes Image im Image-Set besitzt eine bestimmte Assoziationsstärke zur Frage. Die Informationssuche wird durch die jeweils vorhandenen Cues determiniert. Die Wahrscheinlichkeit eines Images, bei der Suche gefunden und abgerufen zu werden, bestimmt sich aus dessen Gesamtaktivierung. Dies ist das Produkt aller Assoziationen dieses Images zu den vorhandenen Cues. Üblicherweise ist die Frage das erste Cue, das ins Arbeitsgedächtnis gelangt. Die Assoziationsstärken, die die Cues zu den einzelnen Images des Image-Sets aufweisen, bestimmen die Wahrscheinlichkeiten der Images, beim Suchprozeß gefunden zu werden. Wenn die relative Gesamtaktivierung hoch genug ist, erhöht sich die Wahrscheinlichkeit für ein Image, gefunden zu werden, und nur wenn sie auch

absolut hoch genug ist, d. h. über einer definierten Abrufschwelle liegt, kann es abgerufen werden (siehe Anderson, 1983). Kommt ein Image ins Arbeitsgedächtnis, wird es seinerseits zu einem weiteren Cue für die mit ihm assoziierten Images des Image-Sets. Dadurch verändern sich die Abrufwahrscheinlichkeiten für die übrigen Images. Das abgerufene Image besitzt zu den ihm numerisch nahestehenden Images eine relativ höhere Assoziationsstärke. Dadurch werden diese beim nächsten Such- und Abrufzyklus bevorzugt gefunden und abgerufen.

Welches Aktivierungsniveau ein einzelnes Image hat, ist nicht vorherzusagen. Es hängt von der Lerngeschichte jedes einzelnen Individuums ab. Beispielsweise könnte ein Besuch der Stadt Weimar viele Images aktiviert haben, die für die Beantwortung der Frage zu Goethes Todesalter wichtig sind. In diesem Fall wäre sogar ein Großteil der Informationen aktiviert worden. Denkbar wäre aber auch die Aktivierung eines einzelnen Images, beispielsweise durch den Besuch einer Ausstellung in der zufällig auch ein bekanntes Porträt Goethes ausgestellt war. Es ist daher einleuchtend, daß keine genaue Aussagen über den Aktivierungsgrad einzelner Images im Gedächtnis getroffen werden können. Die Konsequenz für SARA ist die Annahme zufälliger Aktivierungsstärken für die einzelnen Images.

Lernen Lernen entsteht durch den Abruf von Images in das Arbeitsgedächtnis. Dies führt zum systematischen Anstieg der Assoziationsstärken einzelner Images im Image-Set. Die Assoziationsstärkenmatrix verändert sich während des Samplings für diejenigen Images, die sich gemeinsam im Arbeitsgedächtnis befinden. In jedem Zyklus werden die Assoziationsstärken zwischen den Images bzw. zwischen ihnen und den Cues um einen bestimmten Betrag erhöht. Dadurch werden spezifische Konstellationen von Images gelernt. Dies führt zu einer den Samplingprozeß überdauernden Reorganisation des Netzwerks. Wieviele Such- und Abrufzyklen im Einzelfall tatsächlich stattfinden, variiert sowohl von Frage zu Frage bzw. von Samplingprozeß zu Samplingprozeß. Außerdem ist die intendierte Tiefe der Verarbeitung, die zur Verfügung stehende Zeit sowie die Leistungsmotivation der jeweiligen Person entscheidend. Je tiefer die Verarbeitung, je länger die zur Verfügung stehende Zeit bzw. je leistungsmotivierter eine Person ist, desto mehr Zyklen finden statt. Mit der Zahl der Zyklen in einem Samplingprozeß steigt das Ausmaß der Veränderungen in der Assoziationsstruktur des Image-Sets. Dieses Prinzip der Anhebung der Assoziationsstärken sorgt für eine erhöhte Abrufwahrscheinlichkeit der betroffenen Images. Dies ist das Grundprinzip, nach dem die systematische Wirkung einer Rückmeldung (Anker, Lösung) erklärt wird.

Im Anschluß an das Sampling werden die Werte der ins Arbeitsgedächtnis abgerufenen Images zu einer Schätzung integriert. Die einfachste Annahme ist hier die Berechnung des arithmetischen Mittels aus den Werten der Images (Anderson, 1981, 1986). Der so ermittelte Wert unterliegt abschließend noch Rundungs- bzw. Korrekturprozessen, damit inkompatible Schätzungen (wie z. B. Nachkommastellen bei Jahreszahlen) vermieden werden und eher „runde“ Werte entstehen.

Vergessen Vergessen wird als ein dem Lernen entgegengesetzter Prozeß verstanden. Assoziationen, die in früheren Prozessen gestärkt wurden (weil die Images, die sie verbinden, gemeinsam im Arbeitsgedächtnis waren), können auch wieder als Funktion der verstrichenen Zeit abnehmen (decay). SARA nimmt an, daß durch Decay nicht mehr als das „verlernt“ wird, was vorher durch Samplingprozesse aufgrund selektiver Aktivierung neu erlernt wurde. Dabei nähern sich die Assoziationsstärken asymptotisch den präexperimentellen Ausgangswerten. Dieser Annahme liegt die Überlegung zugrunde, daß die anfänglichen Assoziationsstärken aufgrund der numerischen Ähnlichkeit dauerhaft bestehen. Es wäre außerdem unsinnig, sämtliche Assoziationsstärken bis auf Null abnehmen zu lassen, denn ein solches Langzeitgedächtnis bestünde dann aus lauter isolierten Informationseinheiten, von denen keines jemals wieder abgerufen werden könnte. Ein vollständiges Vergessen ist allerdings für neu enkodierte Informationen (wie z. B. für die eigene Schätzung oder für den Anker) möglich.

Neben diesem „regulären“ Vergessen werden in SARA auch unsystematische Veränderungen der Assoziationsstärken angenommen (fluctuation; Mensink & Raaijmakers, 1988, 1989). Die der Frage verfügbaren Images sind vielfältig verbunden mit Informationen des Langzeitgedächtnisses, die nicht Bestandteil des momentanen Image-Sets sind. Jeder Informationsverarbeitungsprozeß kann potentiell Images des aktuellen Image-Sets betreffen. Dadurch können sich alle Assoziationsstärken jederzeit verändern. Diese schwer erfaßbaren Aktivitäten sowie das „Rauschen im System“ werden hier durch Fluktuationen abgebildet, die sämtliche Assoziationsstärken geringfügig und unsystematisch verändern können.

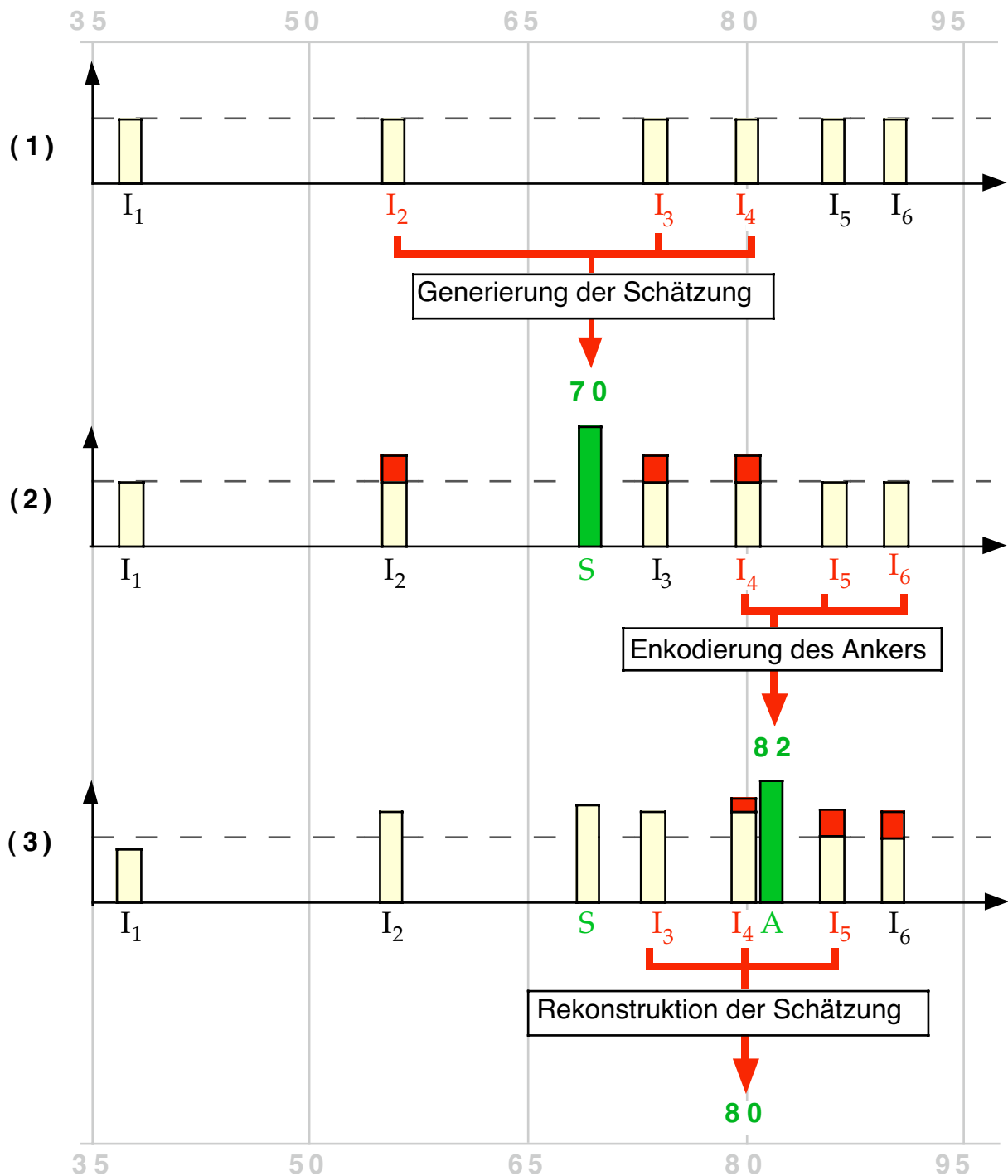


Abb. 4.4: Veränderung der Assoziationsstärken der Images im Image-Set durch Samplingprozesse bei der (1) Generierung der Schätzung, (2) Enkodierung des Ankers und der (3) Rekonstruktion der Schätzung. Die Balken stehen für angenommene Images auf die Frage nach dem Alter Goethes. Die Höhe der Balken soll die Assoziationsstärke verdeutlichen. Die Lage repräsentiert den numerischen Wert der Images. Zu sehen ist außerdem der Anstieg der Assoziationsstärke nach Sampling.

4.4 Spezifische Verarbeitungsprozesse

Der folgende Abschnitt zu den spezifischen Verarbeitungsprozessen orientiert sich an einen typischen Versuchsablauf zum Rückschaufehler im Gedächtnisdesign und verfolgt den inhaltlich interessanten Fall, daß ein Image-Set mit Images zur Abgabe einer Schätzung vorhanden ist..

Generierung einer Schätzung Nehmen wir beispielsweise an, es wird nach dem Alter von Goethe gefragt. Ist das exakte Alter nicht bekannt, wie in den meisten Fällen, muß eine Schätzung abgegeben werden. Es wird eine Suche im Langzeitgedächtnis initiiert nach Informationen (Images), die potentiell nützlich für die Abgabe einer Schätzung sein könnten. Beispielsweise könnte die Information nützlich sein, wie das durchschnittliche Todesalter von Personen zur damaligen Zeit war (circa 38 Jahre). Die Information, daß Goethe über 140 Bände veröffentlicht hat, daß er an die 60 Jahre an Faust (Urfaust, Faust I und II) gearbeitet hat, daß er Zeitgenosse und Bewunderer Napoleon Bonapartes war (Napoleon wurde 52 Jahre alt), Schillers (46 Jahre), Alexander von Humboldts (90 Jahre), Mozarts (35 Jahre), Schopenhauers (72 Jahre). Auch die Erinnerung an einen Stich von ihm in hohem Lebensalter, oder daß er als „Olympier“ bezeichnet wurde, stellt eine potentiell nützliche Information dar (siehe auch Abb.4.2 und 4.3). Die Idee ist einfach: In jedem Fall erfolgt eine Gedächtnissuche nach potentiell nützlichen Informationen. In einem typischen Experiment, in dem zwischen 10 bis 60 Fragen gestellt werden, bleibt wenig Zeit zur Beantwortung dieser Fragen. Die Suche wird also schon aus diesem Grunde nicht erschöpfend ausfallen.

Der Einfachheit halber betrachten wir ein Image-Set aus sechs Images zur Frage nach Goethes Todesalter. Die mit dem Fragecue assoziierten Informationen (I_1 bis I_6) sind entsprechend ihrer Inhalte numerisch aufsteigend angeordnet. Als Inhalte wurden hier die Werte 38, 56, 74, 80, 86 und 90 gewählt. So könnte beispielsweise der Wert „38“ ein Image repräsentieren, das die durchschnittliche Lebenserwartung zu Ende des 18. Jahrhunderts enthält. Durch Sampling werden nun probabilistisch in mehreren Such- und Abrufzyklen Images ins Arbeitsgedächtnis geladen. Als Ausgangspunkt gelangt zunächst die Frage als Fragecue ins Gedächtnis.

Tabelle 4.1: Generierung einer Schätzung

| Image | 38 | 56 | 74 | 80 | 86 | 90 |
|---------------|-----------|-----------|-----------|-----------|-----------|-----------|
| Fragecue | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 |
| Produkt | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 |
| 1. ▼ | | | . | | | |
| Fragecue | 0,50 | 0,50 | 0,65 | 0,50 | 0,50 | 0,50 |
| Cue 74 | 0,49 | 0,74 | | 0,91 | 0,83 | 0,77 |
| Produkt | 0,24 | 0,37 | | 0,46 | 0,41 | 0,39 |
| 2. ▼ | | . | | | | |
| Fragecue | 0,50 | 0,65 | 0,75 | 0,50 | 0,50 | 0,50 |
| Cue 74 | 0,49 | 0,82 | | 0,91 | 0,83 | 0,77 |
| Cue 56 | 0,74 | | 0,74 | 0,66 | 0,57 | 0,51 |
| Produkt | 0,18 | | | 0,30 | 0,24 | 0,20 |
| 3. ▼ | | | | . | | |
| Fragecue | 0,50 | 0,75 | 0,82 | 0,65 | 0,50 | 0,50 |
| Cue 74 | 0,49 | 0,87 | | 0,94 | 0,83 | 0,77 |
| Cue 56 | 0,74 | | 0,82 | 0,76 | 0,57 | 0,51 |
| Cue 80 | 0,40 | 0,66 | 0,91 | | 0,91 | 0,86 |

Die Abbildung 4.4 zeigt die Images angeordnet auf einer numerischen Skala, ihre Assoziationsstärken und ihre Änderungen während des Sampling. Die anfänglichen Assoziationsstärken zwischen Images und Fragecue wurden auf .50 gesetzt, um ihre Entwicklungen leichter verfolgen zu können (1). Sämtliche Veränderungen in der Abbildung sind rot unterlegt. Die Abbildung 4.4 gliedert sich in drei Abschnitte, die den jeweiligen spezifischen Verarbeitungsprozeß darstellen. Unter (1) wird die Generierung einer Schätzung veranschaulicht. Abschnitt (2) veranschaulicht die Enkodierung eines Ankers und Abschnitt (3) schließlich die Rekonstruktion der Schätzung. Die Tabelle 4.1 enthält die zugehörige Assoziationsmatrix und ihre Änderungen während des Sampling. In der ersten Zeile ist der Wert der postulierten Images einer fiktiven Person angegeben. Vereinfachend wurde hier von drei Zyklen ausgegangen, bei dem jeder Zyklus erfolgreich ein Image ins Arbeitsgedächtnis ab-

ruft. Sämtliche Veränderungen in der Tabelle sind grau unterlegt. Alle Samplingprozesse laufen nach demselben Schema ab: Zunächst werden die Gesamtaktivierungen der Images als Produkt aller im Arbeitsgedächtnis vorhandenen Cues berechnet. Dann wird versucht, eines dieser Images abzurufen. Das erfolgreich abgerufene Image (•) wird zu einem weiteren Cue im Arbeitsgedächtnis und bestimmt so die weitere Suche mit.

In einem ersten Zyklus (1. ▼) wird beispielsweise das Image I_3 mit dem Wert 74 gezogen. Dies könnte ein Image repräsentieren, welches ein Porträt eines relativ alt aussehenden Goethes enthält. Dieses Image fungiert jetzt als weiterer Cue im Arbeitsgedächtnis. Dazu werden aufgrund numerischer Ähnlichkeit Assoziationsstärken zu den restlichen Images des Image-Sets aufgebaut: Je ähnlicher, je stärker sind sie assoziiert. Außerdem steigt die Assoziation zwischen ihm und dem Fragecue von anfänglich .50 auf .65. Dadurch erhöht sich die Wahrscheinlichkeit, daß dieses Image in einem späteren Prozeß bei Vorliegen des Fragecues erneut gefunden wird (siehe Zeile 5, Spalte 4 der Tabelle 4.1 und I_3 in (2); Abb. 4.4). Vor dem nächsten Zyklus werden wieder die Gesamtaktivierungen der Images als Produkt aller im Arbeitsgedächtnis vorhandenen Cues berechnet. Dann wird versucht, ein weiteres Image abzurufen. Dieser Prozeß wiederholt sich bei jedem weiteren Zyklus. Im zweiten Zyklus wird das Image mit dem Wert „56“ abgerufen und im dritten Zyklus das Image mit dem Wert „80“. Am Ende werden die abgerufenen Images integriert und als Schätzung auf die Frage nach Goethes Todesalter genannt. Im Beispiel wäre dies „70“ ($= (74+56+80)/3$). Zusätzlich wird die Schätzung als neues Image enkodiert (dem Image-Set hinzugefügt) und mit dem Fragecue verknüpft (s. Tab. 4.2 und (2); Abb. 4.4).

Tabelle 4.2: Enkodierung des Ankers

| Image | 38 | 56 | 70 | 74 | 80 | 86 | 90 |
|---------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Fragecue | 0,50 | 0,75 | 0,80 | 0,82 | 0,65 | 0,50 | 0,50 |
| Ankercue (82) | 0,37 | 0,63 | 0,83 | 0,89 | 0,97 | 0,94 | 0,89 |
| Produkt | 0,19 | 0,47 | 0,66 | 0,73 | 0,63 | 0,47 | 0,44 |
| 1. ▼ | | | | | | . | |
| Fragecue | 0,50 | 0,75 | 0,80 | 0,82 | 0,65 | 0,65 | 0,50 |
| Ankercue | 0,37 | 0,63 | 0,83 | 0,89 | 0,97 | 0,96 | 0,89 |
| Cue 86 | 0,31 | 0,57 | 0,77 | 0,83 | 0,91 | | 0,94 |
| Produkt | 0,06 | 0,27 | 0,51 | 0,60 | 0,57 | | 0,42 |
| 2. ▼ | | | | | | | . |
| Fragecue | 0,50 | 0,75 | 0,80 | 0,82 | 0,65 | 0,75 | 0,65 |
| Ankercue (82) | 0,15 | 0,50 | 0,77 | 0,85 | 0,96 | 0,97 | 0,92 |
| Cue 86 | 0,31 | 0,57 | 0,77 | 0,83 | 0,91 | | 0,96 |
| Cue 90 | 0,26 | 0,51 | 0,71 | 0,77 | 0,86 | 0,94 | |
| Produkt | 0,01 | 0,11 | 0,34 | 0,45 | 0,49 | | |
| 3. ▼ | | | | | . | | |
| Fragecue | 0,50 | 0,75 | 0,80 | 0,82 | 0,75 | 0,82 | 0,75 |
| Ankercue (82) | 0,15 | 0,50 | 0,77 | 0,85 | 0,97 | 0,98 | 0,94 |
| Cue 86 | 0,08 | 0,42 | 0,69 | 0,77 | 0,94 | | 0,97 |
| Cue 90 | 0,26 | 0,51 | 0,71 | 0,77 | 0,90 | 0,96 | |
| Cue 80 | 0,40 | 0,66 | 0,86 | 0,91 | | 0,91 | 0,86 |

Enkodierung eines Ankers In der Regel wird in dieser Phase zunächst die Frage zusammen mit einem spezifischen Anker dargeboten (hier der Lösung „82 Jahre“). Der Anker wird dabei meist als bedeutsame Information für die gestellte Frage betrachtet. Diese Bedeutsamkeit erlangt der Anker durch die mit seiner Präsentation verbundene Kennzeichnung (z. B. als „Lösung“ oder als „Schätzung einer anderen Person“) und aufgrund der relativen Unwissenheit des Probanden, die jede

fragerelevante Information wichtig erscheinen läßt. Aus diesen Gründen wird in SARA angenommen, daß die Darbietung des Ankers zu seiner Enkodierung führt. Die Enkodierung des Ankers wird dadurch geleistet, daß ähnliche Informationen im Langzeitgedächtnis gefunden und mit diesen Hinweisreizen verknüpft werden. Dazu dient erneut der beschriebene Samplingprozeß. Die Stärke der Assoziationen zwischen Anker und Images richtet sich wiederum nach der numerischen Nähe zwischen dem Ankerwert und den Inhalten der Images. Die vorhandenen Cues bestimmen die Gesamtaktivierung der Images und damit deren Wahrscheinlichkeit, abgerufen zu werden. Die dem Anker (numerisch) naheliegenden Images haben aufgrund ihrer stärkeren Assoziation zum Anker eine höhere Abrufwahrscheinlichkeit als diejenigen Images, die eine größere (numerische) Distanz zum Anker aufweisen. Dadurch gelangen beim Sampling eher dem Anker naheliegende Images ins Arbeitsgedächtnis.

Bei der Enkodierung des Ankers (Tab. 4.2 und (2); Abb. 4.4) startet der Samplingprozeß mit einem zusätzlichen Cue, dem Ankercue. Die Assoziationen zwischen den Images und dem Fragecue sind jetzt nicht mehr bei .50, sondern aufgrund der Lerngeschichte während der Schätzungsgenerierung verändert. Die Assoziationsstärken der Cues richten sich wie sonst auch nach der numerischen Ähnlichkeit: Je ähnlicher, je stärker sind sie assoziiert. Ansonsten läuft der Samplingprozeß unverändert ab. In diesem Fall werden die Images „86“, „90“ und „80“ erfolgreich abgerufen. Ihre Assoziationen zum Frage- und zum Ankercue werden deshalb erhöht, wiederum je nach gemeinsamer Aufenthaltsdauer im Arbeitsgedächtnis. Am Ende wird der Anker („82“) als weiteres Image dem Image-Set hinzugefügt (s. Tab. 4.3 und (3); Abb. 4.4). Dies ist der für den Rückschaufehler entscheidende Prozeß, durch den die Images, die dem Anker nahestehen, selektiv aktiviert werden. Wird das derart reorganisierte Wissen später benutzt, um eine Schätzung zu generieren (Urteilsdesign) oder sich an eine ursprüngliche Schätzung zu erinnern (Gedächtnisdesign), haben die durch die Ankerenkodierung selektiv aktivierten Images sowie der Anker selbst eine höhere Wahrscheinlichkeit, abgerufen zu werden. Die Suche und deren Ergebnis sind dann in Richtung des Ankers verzerrt.

Tabelle 2.3: Rekonstruktion einer Schätzung

| | | | | | | | | |
|---------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Image | 38 | 56 | 70 | 74 | 80 | 82 | 86 | 90 |
| Fragecue | 0,50 | 0,75 | 0,80 | 0,82 | 0,75 | 0,70 | 0,82 | 0,75 |
| Kontextcue | 0,50 | 0,75 | 0,80 | 0,82 | 0,65 | 0,70 | 0,50 | 0,50 |
| Produkt | 0,25 | 0,56 | 0,64 | 0,68 | 0,48 | 0,49 | 0,41 | 0,38 |
| 1. ▼ | | | | | . | | | |
| Fragecue | 0,50 | 0,75 | 0,80 | 0,82 | 0,82 | 0,70 | 0,82 | 0,75 |
| Kontextcue | 0,50 | 0,75 | 0,80 | 0,82 | 0,75 | 0,70 | 0,50 | 0,50 |
| Cue 80 | 0,40 | 0,66 | 0,86 | 0,91 | | 0,97 | 0,91 | 0,86 |
| Produkt | 0,10 | 0,37 | 0,55 | 0,62 | | 0,48 | 0,38 | 0,32 |
| 2. ▼ | | | | . | | | | |
| Fragecue | 0,50 | 0,75 | 0,80 | 0,88 | 0,88 | 0,70 | 0,82 | 0,75 |
| Kontextcue | 0,50 | 0,75 | 0,80 | 0,88 | 0,82 | 0,70 | 0,82 | 0,75 |
| Cue 80 | 0,40 | 0,66 | 0,86 | 0,94 | | 0,97 | 0,91 | 0,86 |
| Cue 74 | 0,49 | 0,74 | 0,94 | | 0,91 | 0,89 | 0,83 | 0,77 |
| Produkt | 0,05 | 0,27 | 0,52 | | | 0,42 | 0,31 | 0,25 |
| 3. ▼ | | | | | | | . | |
| Fragecue | 0,50 | 0,75 | 0,80 | 0,91 | 0,91 | 0,70 | 0,88 | 0,75 |
| Kontextcue | 0,50 | 0,75 | 0,80 | 0,91 | 0,88 | 0,70 | 0,65 | 0,75 |
| Cue 80 | 0,40 | 0,66 | 0,86 | 0,96 | | 0,97 | 0,94 | 0,86 |
| Cue 74 | 0,49 | 0,74 | 0,94 | | 0,94 | 0,89 | 0,88 | 0,77 |
| Cue 86 | 0,31 | 0,57 | 0,77 | 0,83 | 0,91 | 0,94 | | 0,94 |

Rekonstruktion einer Schätzung Wird nun die Aufgabe gestellt, sich an die ursprüngliche Schätzung zu erinnern, führt dies zu der Suche nach der entsprechenden Information im Image-Set. Auch diese Suche wird durch einen Samplingprozeß realisiert. Als Hinweisreize fungieren neben der Frage der Aufgabenkontext („Generiere eine Schätzung!“). Im Prinzip kann auch die ursprüngliche Antwort gefunden und abgerufen werden. Wenn ihre Aktivierung

jedoch zu gering ist, wird sie nicht als die gesuchte Information erkannt. In diesem Fall wird sie wie jedes andere Image behandelt. Dieser Einschränkung liegt die Annahme zugrunde, daß die Stärke der Quelleninformation mit der Gesamtaktivierung eines Images kovariiert. Ist die Gesamtaktivierung der abgerufenen Schätzung ausreichend hoch, wird die Suche abgebrochen und die Originalschätzung genannt, d. h. es kommt zu einer korrekten Erinnerung. In Experimenten zum Rückschaufehler finden sich im Mittel 27% korrekte Erinnerungen (gemittelt über 25 Datensätze bei einer Streuung von 11%; Pohl, 1993). Dieser Wert hängt vom Retentionsintervall und damit vom Ausmaß des Vergessens ab. In SARA werden diese Veränderungen durch die beschriebenen Vergessensprozesse abgebildet.

Dieselben Überlegungen gelten auch für den Anker, wenn er bei der Suche im Gedächtnis gefunden wird. Bei genügend hoher Gesamtaktivierung wird er als Anker erkannt, sonst nicht. In vielen Untersuchungen findet die Erinnerungsaufgabe jedoch parallel zur Lösungsdarbietung statt, d. h. der Anker ist noch präsent und muß nicht erinnert werden. In diesem Fall wirkt er als weiterer Cue im Samplingprozeß.

Die Menge der bei der Rekonstruktion abgerufenen Images unterscheidet sich jedoch systematisch von denjenigen bei der Generierung der Originalschätzung. Dafür sind im Wesentlichen zwei Prozesse verantwortlich: (1) Die Ankerenkodierung führt aufgrund einer selektiven Aktivierung zu Veränderungen in der Assoziationsstruktur des Image-Sets und damit zu veränderten Abrufwahrscheinlichkeiten der einzelnen Images. (2) Der Anker, falls er präsent ist bzw. gefunden, aber nicht erkannt wird, fungiert als zusätzlicher Cue und beeinflusst so die Suche im Gedächtnis. Der erste Fall spiegelt eher Gedächtnisveränderungen (wie in der Assimilationstheorie von Fischhoff, 1975), der zweite eine Beeinflussung des aktuellen Abrufs (wie in der Rekonstruktionstheorie von Stahlberg & Maass, 1998) wider. Das Ergebnis ist in jedem Fall eine systematisch verzerrte Auswahl von Images.

Bei der Rekonstruktion der Schätzung schließlich (Tab. 4.3 und (3); Abb. 4.4) werden im Laufe des Samplingprozesses die Images „80“, „74“ und „86“ abgerufen. Die Originalschätzung („70“) wurde hier nicht gefunden. Als „erinnerte“ Schätzung wird deshalb der rekonstruierte Wert „80“ ($= (80+74+86)/3$) genannt, ein Wert, der gegenüber der Originalschätzung zum Anker hin verschoben ist, wie es für den Rückschaufehler typisch ist.

4.5 FAZIT

Mit SARA liegt erstmals eine umfassende Theorie mit breitem Anwendungsbereich für den Rückschaufehler und ganz allgemein für Ankereffekte vor. Viele der bekannten Befunde zum Rückschaufehler können mit den Annahmen des Modells leicht erklärt werden. Ein Vorteil des Modells ist es, daß es einerseits weitgehend auf klassischen Annahmen der kognitiven Psychologie beruht, andererseits aber flexible Prozeßmodule postuliert, die für viele Arten experimenteller Abläufe nutzbar gemacht werden können. Die Annahmen des Modells SARA beziehen sich einerseits auf die Ebene des Wissens, das Personen zur Beantwortung von Fragen nutzen, andererseits auf die kognitiven Prozesse, die eine Schätzung zu der gestellten Frage generieren, einen Anker enkodieren oder eine vergessene Schätzung rekonstruieren. Mit dem Modell wird versucht, die Veränderungen des Wissens in jeder Phase eines experimentellen Ablaufs zu erfassen, um so vorherzusagen, zu welchen Ergebnissen grundlegende Informationsverarbeitungsprozesse kommen werden. Der Rückschaufehler bzw. der Ankereffekt werden in diesem Rahmen durch Veränderungen in den Konstituenten des Wissens als Folge der selektiven Aktivierung von Wissenseinheiten erklärt. SARA erreicht dadurch einen wesentlich höheren Präzisionsgrad als bisherige Theorien zum Rückschaufehler.

Die *selektive Aktivierung* ist das Grundprinzip, nach dem in SARA die verzerrende Wirkung eines Ankers erklärt wird. Eine ähnliche Konzeption, allerdings beschränkt auf selektivem Abruf – die selektive Erhöhung der Zugänglichkeit – wurde von Strack und Mussweiler (1997; siehe Kap. 3.2.2) zur Erklärung von Ankereffekten vorgeschlagen. Die systematische Veränderung der Assoziationsstruktur des Image-Sets entweder durch die Enkodierung des Ankers oder aber auch dadurch, daß der vorliegende (oder erfolgreich abgerufene) Anker als zusätzlicher Abrufreiz wirkt, führt zum bevorzugten Abruf ankernaher Images, da ihre Gesamtaktivierungen höher sind als die anderer Images. Auf diese Weise werden spezifische Konstellationen von Images und neue Zusammenhänge gelernt. Damit enthält SARA sowohl das Konzept der Assimilation (Fischhoff, 1975, siehe Kap. 3.2.1) als auch das der verzerrten Rekonstruktion (Stahlberg et al., 1993; Stahlberg & Maass, 1998; siehe Kap. 3.2.2). Durch beide Prozesse ändern sich die Abrufwahrscheinlichkeiten der einzelnen Images. Eine solche Bevorzugung von Informationen durch Erhöhung ihrer späteren Abrufwahrscheinlichkeiten muß experimentell überprüft werden. Die Vorhersagen, die sich aus der selektiven Aktivierung des Image-Sets ergeben, sollen in fünf Experimenten eingehend geprüft werden.

Empirische Arbeiten

5 Experimente mit Wiedererkennenstest

EINLEITUNG

Für eine konkrete Frage (z.B.: „Wie lang ist die Elbe?“) ist es kaum möglich anzugeben, welche Informationen zur Generierung einer Schätzung herangezogen werden. Daher wurden in den hier vorzustellenden Experimenten die Image-Sets experimentell vermittelt.

Die Experimente folgten dem Gedächtnisdesign und fanden mit Ausnahme eines Experiments am Computer statt. Die Aufgabe bestand darin, in 18 Verifikationsaufgaben zu beurteilen, ob die dort beschriebenen Sachverhalte in den letzten Jahren eher gestiegen oder gefallen sind, und sich zum Testzeitpunkt, genau eine Woche später, wieder an das eigene Urteil zu erinnern. Zu jedem Sachverhalt wurden acht von insgesamt zwölf Argumenten dargeboten, wobei jeweils vier von sechs Argumenten dafür sprachen, daß der Sachverhalt gestiegen, und vier von sechs dafür, daß er gefallen war. Die restlichen vier Argumente dienten als neue Argumente für einen späteren Wiedererkennenstest. Ein Beispiel soll das verdeutlichen (kursiv sind jeweils zusammenfassende Stichworte zu den Argumenten angegeben):

Die Stadtwerke führten jüngst eine Umfrage zur Benutzung der Busse unter den Bürgern der Stadt durch. Hat die Benutzung der öffentlichen Busunternehmen im vergangenen Jahr zugenommen oder abgenommen?

Sie hat abgenommen,

weil schlechte Anschlußverbindungen bestehen. Anschlüsse schlecht

weil immer mehr Bürger bei Fahrten innerhalb der Stadt ihr eigenes Auto benutzen. Auto benutzt

weil die Busse häufig Verspätung haben. Verspätung gehabt

weil die Busse zu unregelmäßig fahren. unregelmäßig gefahren

weil preiswerte Reiseangebote reduziert wurden. Angebote reduziert

*weil Billigtarife für Kinder, Studenten und Senioren abgeschafft wurden.
Billigticket abgeschafft*

Sie hat zugenommen,

weil viele Bürger es aus Umweltschutzgründen vermeiden, bei Fahrten innerhalb der Stadt ihr Auto zu benutzen. *Auto vermieden*

weil durch den Ausbau des Streckennetzes alle Stadtbezirke gut erreichbar sind. *Streckennetz ausgebaut*

weil die Busunternehmen seit einiger Zeit Städtereisen anbieten. *Städtereisen angeboten*

weil die Busse in letzter Zeit immer voll sind. *Busse überfüllt*

weil ein Wochenendticket eingeführt wurde. *Wochenendticket eingeführt*

weil die Fahrpreise verringert wurden. *Fahrpreise verringert*

Die Probanden sollten dann ihre Sicherheit angeben, ob der entsprechende Sachverhalt gestiegen oder gefallen ist. Dazu konnten sie einen Regler auf in einer analogen Skala, deren Extrempunkte mit „gestiegen“ bzw. „gefallen“ bezeichnet waren, verschieben. Der jeweilige Abstand zu den Extrempunkten der Skala sollte die subjektive Sicherheit des Urteils anzeigen. Entsprechend sollte eine Angabe nahe dem Extrempunkten eine hohe Sicherheit, und eine Antwort in der Mitte der Skala Unsicherheit ausdrücken.

Zum zweiten Termin, genau eine Woche nach dem ersten, bekamen die Probanden erneut alle 18 Sachverhalte dargeboten (allerdings ohne Argumente), wobei für jeweils sechs der Sachverhalte die Rückmeldung erfolgte, der Sachverhalt wäre „gestiegen“, für weitere sechs, der Sachverhalt wäre „gefallen“ und für sechs keine Rückmeldung gegeben wurde. Die sechs Sachverhalte ohne Rückmeldung dienten als Kontrollitems. Die Aufgabe war, möglichst genau das Urteil aus der ersten Sitzung zu erinnern.

In den Experimenten wurde zusätzlich das Wiedererkennen der Argumente geprüft. In einem Alt-Neu-Paradigma wurden die Reaktionszeiten für die acht alten gegenüber den vier neuen Argumenten erhoben.

Die Erinnerung an die Konfidenzurteile sollte sich in den Experimentalbedingungen systematisch von der Kontrollbedingung unterscheiden: Die Rückmeldung, ein Sachverhalt wäre gestiegen, sollte die Sicherheit für „steigen“ erhöhen und für „fallen“ vermindern, während umgekehrt die Rückmeldung, daß ein Sachverhalt gefallen wäre, die Sicherheit für „fallen“ erhöhen und für „steigen“ mindern sollte. Ein solches Ergebnis entspräche einem typischen Rückschaufehler.

Im Alt-Neu-Paradigma sollte sich ebenfalls ein Zusammenhang mit der Rückmeldung zeigen lassen. Argumente, die für „fallen“ sprechen, sollten häufiger und schneller erkannt werden als solche, die für „steigen“ sprechen, wenn die Rückmeldung „fallen“ lautet und entsprechend umgekehrt. Dieses Ergebnis spräche für eine selektive Aktivierung des Image-Sets.

5.1 Experiment 1: Wiedererkennen der Argumente

METHODE

Design. Der Versuchsplan (siehe Abb. 5.1) bestand in einem 3×2 Design mit den Faktoren „Rückmeldung“ (gestiegen/gefallen/keine) und „Argument“ (fallen/steigen). Beide Faktoren wurden als Meßwiederholungsfaktoren realisiert. Die Zuordnung der Fragen zu dem Faktor „Rückmeldung“ wurde durch systematische Permutation einer einmaligen Zufallsreihenfolge realisiert, so daß alle Fragen über alle Bedingungen gleich häufig verwirklicht wurden. Eine mit „gestiegen“ bezeichnete Rückmeldung zeigte an, daß der zu beurteilende Sachverhalt in Wirklichkeit gestiegen wäre. Entsprechend zeigte eine mit „gefallen“ bezeichnete Rückmeldung an, daß der zu beurteilende Sachverhalt in Wirklichkeit gefallen wäre. Die Bedingung „keine“ Rückmeldung diente als Kontrollbedingung. Der Faktor „Argument“ gruppiert die Wiedererkennensleistung der Argumente in Argumente, die dafür sprachen, daß der entsprechende Sachverhalt gestiegen („steigen“) und Argumente, die dafür sprachen, daß er gefallen („fallen“) war. Die Argumente sprachen jeweils zur Hälfte dafür, daß der entsprechende Sachverhalt gestiegen respektive gefallen war. Die Darbietung erfolgte blockweise, wobei die Reihenfolge der Argumente als auch der Blöcke selbst einer Zufallsreihenfolge folgte, so daß alle Argumente gleich häufig an jeder Position präsentiert wurden.

Als abhängige Maße wurden neben der Differenz der Konfidenzurteile zum Zeitpunkt der ersten Einschätzung und dem Zeitpunkt der Erinnerung ($\Delta\%$) die Antworten und Reaktionszeiten in einer Alt-Neu-Entscheidung erhoben.

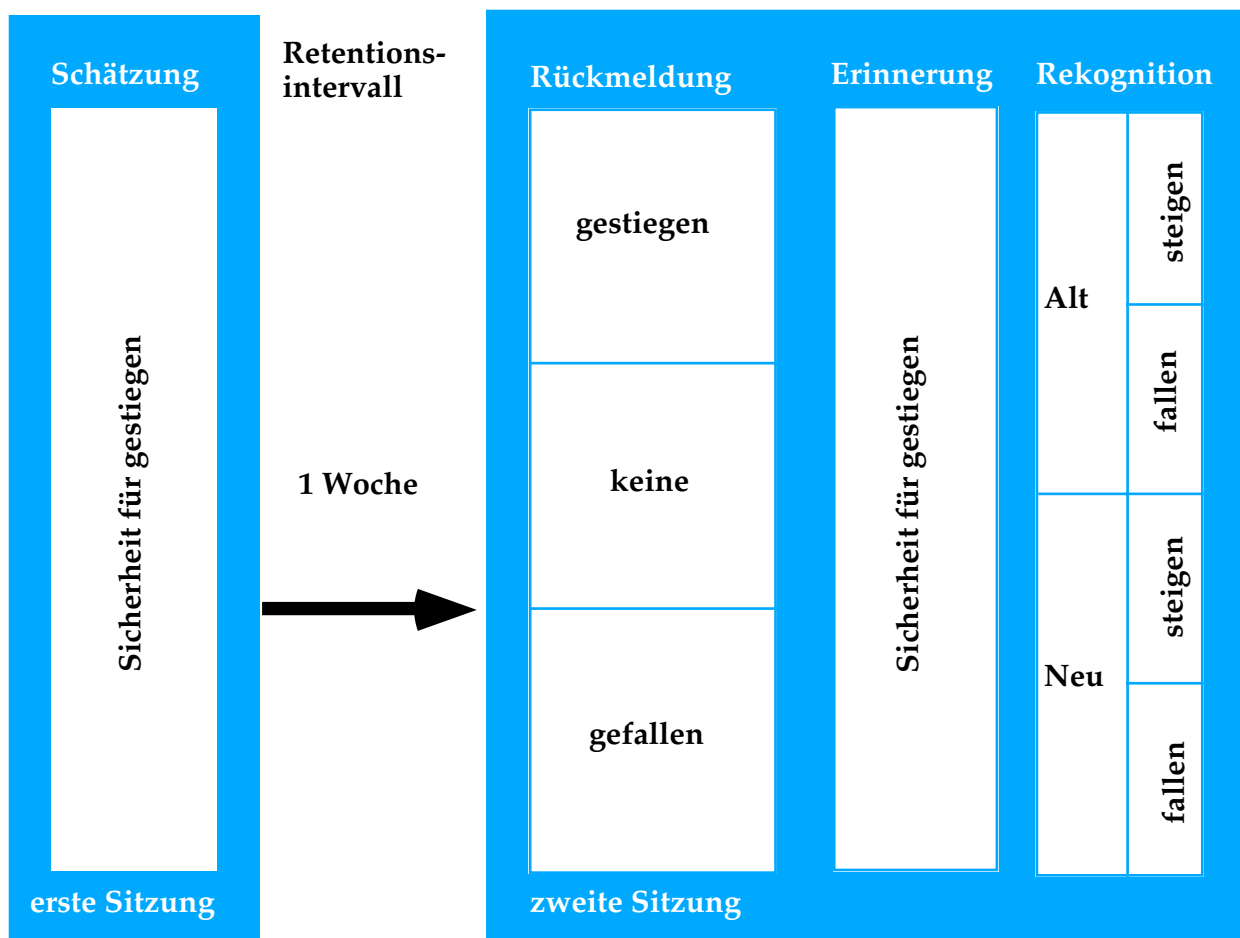


Abbildung 5.1: Versuchsplan von Experiment 1. Gedächtnisdesign mit einer Woche Retentionsintervall und Meßwiederholung auf den Faktoren „Rückmeldung“ und „Argumente“.

Probanden An der Untersuchung nahmen 64 Personen, vorwiegend Studierende aus verschiedenen Fachbereichen der Universität Trier, teil. Davon waren 39 weiblichen und 25 männlichen Geschlechts. Das Altersspektrum reichte von 19 bis 34 Jahren. Der Altersdurchschnitt lag bei 23.6 Jahren. Die Probanden erhielten 12 DM für die Teilnahme am Experiment.

Apparatur Die Experimente wurden alle, mit Ausnahme von Experiment 4, rechnergestützt am Computer durchgeführt. Es standen jeweils vier Macintosh PowerMacs 7100 mit 17 Zoll Bildschirmen zur Verfügung. Die Versuchssteuerungsprogramme wurden in HyperCard 2.4 programmiert.

Material 18 Verifikationsaufgaben, die danach beurteilt werden sollten, wie sicher die Probanden jeweils waren, ob ein betreffender Sachverhalt gestiegen oder gefallen war. Die Beurteilung der Sachverhalte erfolgte jeweils durch Verschiebung einer Markierung auf einer analogen Skala auf dem Computerbildschirm, die von

„gefallen“ zu „gestiegen“ reichte. Die Probanden bekamen keine Rückmeldung des konkreten numerischen Skalenwertes.

Bei der Auswahl der Sachverhalte (siehe Anhang) wurde großer Wert darauf gelegt, möglichst neutrale Sachverhalte zu präsentieren, bei denen wenig präexperimentelle Präferenzen zu erwarten waren. Dies sollte gewährleisten, daß die präsentierten Argumente zur Veränderung der Sachverhalte in die jeweilige Richtung, „gestiegen“ respektive „gefallen“, maximal wirken sollten. Die Sachverhalte bezogen sich alle auf die fiktive Stadt „Wilmerhausen“ und betrafen, wie in dem Beispiel bereits angedeutet, Veränderungen der Zahl öffentlicher Nahverkehrsmittel, Schülerzahlen einer Musikschule, Umsatzzahlen von örtlichen Geschäften und ähnliche Themen.

Um die Präsentation übersichtlicher zu gestalten und weil das Platzangebot auf dem Bildschirm begrenzt war, wurden die Argumente auf dem Bildschirm in Stichworten zusammengefaßt. Zunächst wurde den Probanden jedes Argument vollständig ausformuliert zusammen mit einem Stichwort dargeboten. Die Probanden sollten sich das Argument durchlesen und anschließend die Maustaste betätigen, damit das nächste Argument auf dem Bildschirm erscheinen konnte. Nachdem alle acht Argumente präsentiert worden waren, blieben auf dem Bildschirm, in der Reihenfolge ihrer Darbietung, nur die Stichwörter zu den Argumenten sichtbar. Die Probanden konnten aber jederzeit durch Klicken auf ein Stichwort das entsprechende Argument nochmals lesen. Um zu gewährleisten, daß alle Argumente genügend beachtet wurden, mußten die Probanden die Wichtigkeit aller Argumente einschätzen, bevor sie ihre Beurteilung des Sachverhaltes eingeben konnten. Die Wichtigkeit konnte durch Verschiebung einer Markierung auf einer analogen Skala hinter jedem Stichwort eingeschätzt werden. Die Endpunkte der Skalen waren mit „unwichtig“ bzw. „wichtig“ bezeichnet.

Ablauf Die Durchführung folgte dem Gedächtnisdesign und fand in zwei Sitzungen mit einer Woche Abstand statt. Der gesamte Versuch erfolgte am Rechner. Es bestand jeweils für vier Probanden gleichzeitig die Möglichkeit teilzunehmen.

In der ersten Sitzung wurden die Probanden zunächst in einer ausführlichen Instruktion (siehe Anhang) am Bildschirm aufgeklärt, daß sie im folgenden 18 Sachverhalte danach beurteilen sollten, ob der betreffende Sachverhalt gestiegen oder gefallen wäre. Sie wurden informiert, daß sie ihre Antwort in einer analogen Skala am unteren Bildschirmrand eintragen sollten, wobei der jeweilige Abstand zu den Extrempunkten der Skala, die mit „gestiegen“ bzw. „gefallen“ bezeichnet waren, die subjektive Sicherheit ihres Urteils anzeigen sollte. Es wurde dabei betont,

daß eine Antwort nahe einem Extrempunkt eine hohe Sicherheit ausdrücke, wohingegen eine Antwort in der Nähe der Mitte der Skala Unsicherheit ausdrücke. Sie wurden informiert, daß sie zu jedem Sachverhalt acht Argumente präsentiert bekämen. Sie wurden angehalten, diese sorgfältig zu lesen, da sie für den weiteren Verlauf des Experimentes sehr wichtig seien und ihnen ihre Entscheidung erleichtern sollten. Die Probanden wurden darüber informiert, daß aus Platzgründen auf dem Bildschirm jeweils nur noch Stichworte zu den Argumenten erschienen, sie aber jeweils durch Klicken auf ein entsprechendes Stichwort die Gelegenheit hätten, das zugehörige Argument nochmals zu sehen. Weiterhin wurden die Probanden instruiert, daß es zur Evaluation der Argumente notwendig wäre, jedes Argument in seiner Wichtigkeit einzuschätzen. Dazu erschien nach jedem Stichwort eine analoge Skala, auf der sie durch Mausklick einen Regler so einstellen könnten, daß er ihre Einschätzung der Wichtigkeit dieses Arguments wiedergäbe. Vor der Präsentation der ersten Frage wurde ein Übungsdurchgang (mit nicht ausgewerteten Sachverhalten) durchgeführt.

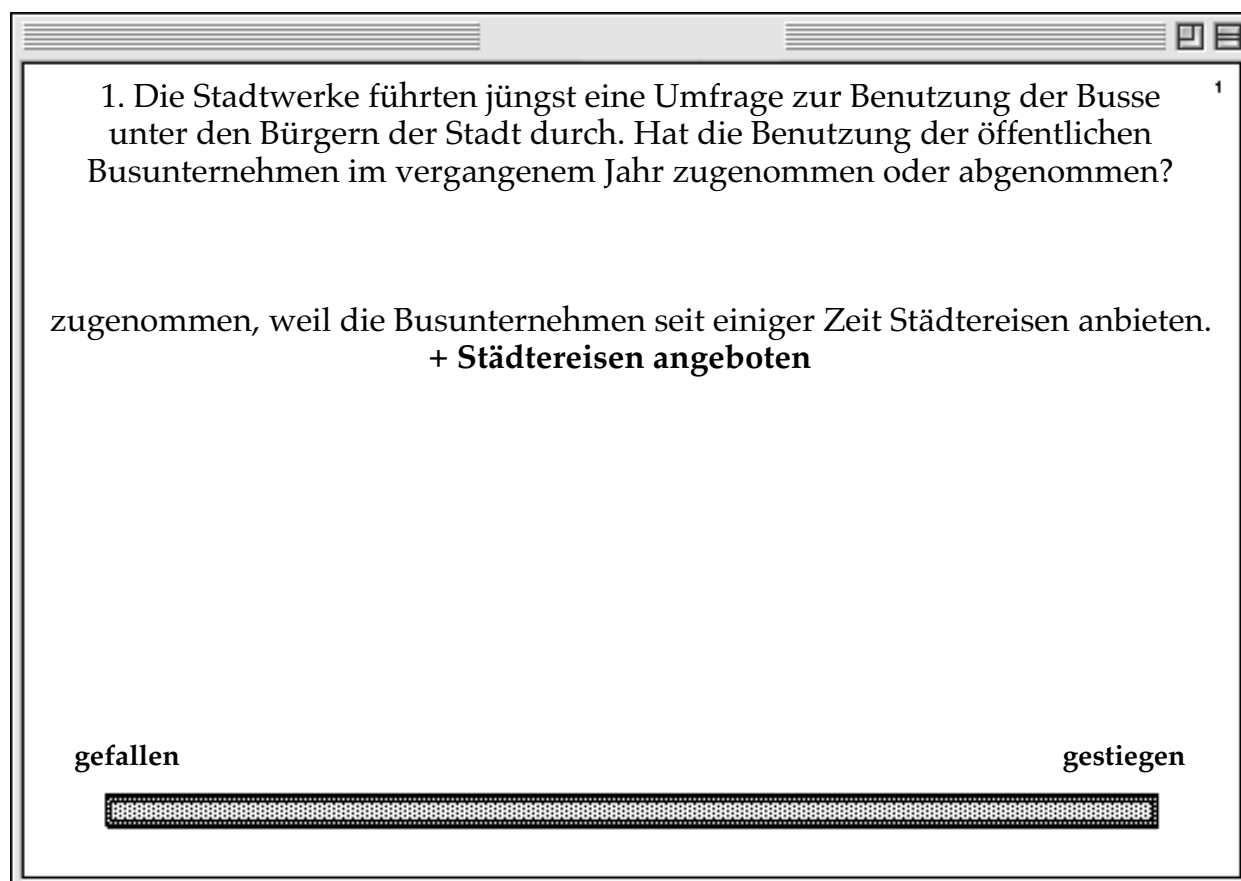


Abbildung 5.2: Bildschirm in Experiment 1 bei der Präsentation eines Arguments mit zugehörigem Stichwort (unterhalb des Arguments durch fetten Schriftschnitt hervorgehoben).

Zu Beginn eines Versuchsdurchgangs erschien am oberen Teil des Bildschirms jeweils ein Sachverhalt. Anschließend wurden nacheinander, jeweils einzeln, die acht Argumente dargeboten. Es wurde jeweils ein Block von vier Argumenten präsentiert, die dafür sprachen, daß der jeweilige Sachverhalt gestiegen war und ein Block von vier Argumenten, die dafür sprachen, daß der Sachverhalt gefallen war. Sowohl die Reihenfolge der Blöcke als auch die Reihenfolge innerhalb der Blöcke war randomisiert. Es erschien am Bildschirm jeweils das Argument, gefolgt von einem textlich abgehobenen Stichwort (siehe Abb. 5.2). Durch Klicken der Maus wurde das nächste Argument präsentiert. Nach dem letzten Argument erschien neben jedem Stichwort und am unteren Bildschirmrand eine analoge Skala, deren jeweilige Endpunkte mit „unwichtig“, „wichtig“ bzw. „gestiegen“ und „gefallen“ gekennzeichnet waren. Auf dem Bildschirm war zu diesem Zeitpunkt der Sachverhalt und in der Reihenfolge ihres Erscheinens die Stichworte zu den Argumenten und die Skalen zu sehen. Den Probanden war es freigestellt, durch Klicken auf ein Stichwort nochmals das ausformulierte Argument zu sehen (siehe Abb. 5.3).

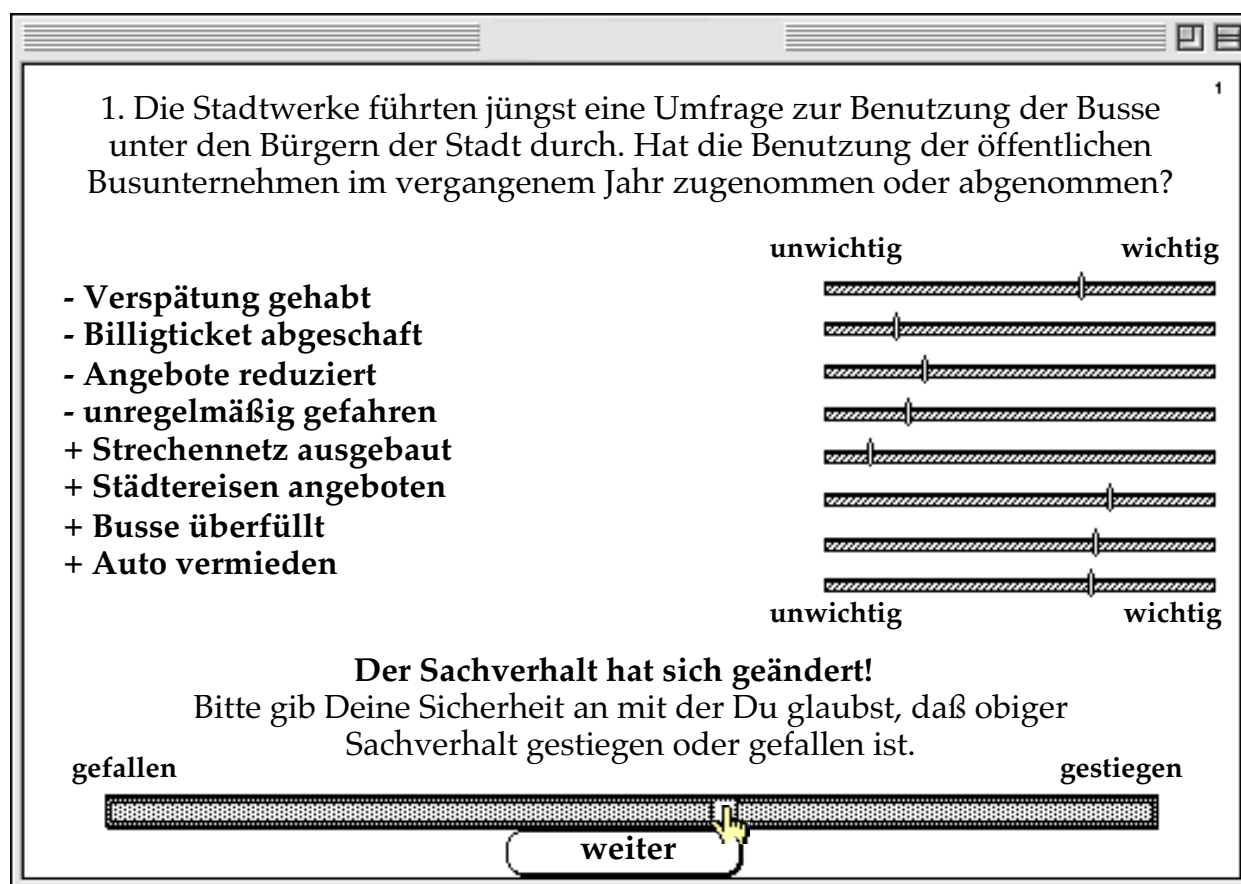


Abbildung 5.3: Bildschirm in Experiment 1 mit den Stichworten zu den Argumenten und den Skalen zur Einschätzung der Wichtigkeit der Argumente und der Änderung des Sachverhaltes.

Bei allen Skalen wurde durch einen Mausklick auf die Skala an dieser Stelle ein Schieberegler sichtbar, der mit der Maus verschoben werden konnte. Die Probanden hatten die Aufgabe zunächst die Wichtigkeit der Argumente zu beurteilen und anschließend auf der unteren großen Skala ihre Sicherheit anzugeben, ob der betreffende Sachverhalt gestiegen oder gefallen sei. Gleichzeitig erschien auf dem Bildschirm eine Taste mit der Aufschrift „weiter“, mit der die Probanden zum nächsten Sachverhalt gelangen konnten. Es bestand keinerlei Zeitlimit. Die Probanden konnten solange bei einem Sachverhalt verweilen, wie es ihnen beliebte. Dieser Teil des Experimentes dauerte im Durchschnitt 35 Minuten pro Proband.

In der zweiten Sitzung nach einer Woche wurden die Probanden instruiert, daß nun die gleichen Sachverhalte wie in Sitzung 1 nochmals dargeboten würden und es darum ginge, sich jetzt möglichst genau an die eigenen Konfidenzurteile aus der vorangegangenen Sitzung zu erinnern. Sie wurden informiert, daß ihnen zu einigen Fragen das tatsächliche Ergebnis, ob der Sachverhalt gestiegen oder gefallen wäre, präsentiert werden würde (siehe Abb. 5.4).

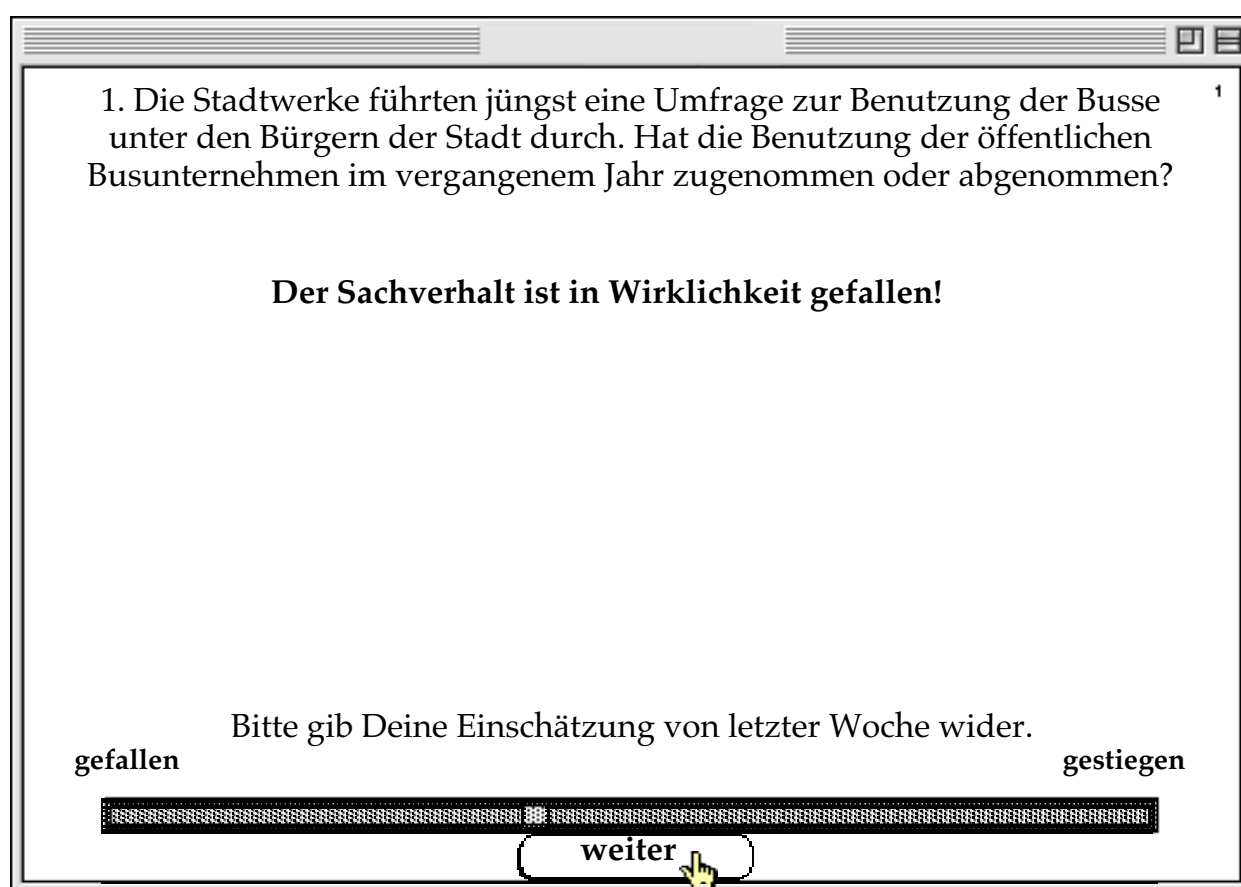


Abbildung 5.4: Bildschirm in Experiment 1 mit der Rückmeldung, obiger Sachverhalt sei in Wirklichkeit gefallen. Die Erinnerung an den Sachverhalt von vor einer Woche wurde mittels Schieberegler eingetragen.

Die Reihenfolge der Sachverhalte und der Bildschirmaufbau wurden beibehalten, im Gegensatz zur ersten Sitzung wurden jedoch keine Stichworte und Argumente präsentiert. Nachdem die Probanden zum jeweiligen Sachverhalt ihr Urteil von der ersten Sitzung zu replizieren versucht hatten, wurden ihnen in zufälliger Reihenfolge die acht alten und vier neue Argumente präsentiert. Anschließend gelangten sie zum nächsten Sachverhalt.

Hatten sich die Probanden an ihre Einschätzung des Sachverhalts aus der ersten Sitzung erinnert, erschien eine zweite Instruktion, die sie informierte, daß ihnen im folgenden jeweils nach Abgabe ihrer Antworten alte und neue Stichworte zu den Sachverhalten präsentiert werden würden (siehe Abb. 5.5). Ihre Aufgabe wäre es, möglichst schnell und korrekt zu entscheiden, ob es sich dabei um ein altes, im Zusammenhang mit diesem Sachverhalt schon genanntes Stichwort, oder um ein neues Stichwort handelte. Die Probanden wurden instruiert, ein kleines Fadenkreuz in der Mitte des Bildschirms zu fixieren und, sobald ein Stichwort erschiene, durch Druck auf eine von zwei deutlich gekennzeichneten und weit voneinander entfernten Tasten der Tastatur ihr Urteil abzugeben. Die mit „alt“ beschriftete „Ctrl-Taste“ und mit „neu“ beschriftete „Enter-Taste“ dienten der Eingabe der Entscheidung der Probanden. Analog zur ersten Sitzung wurde vor der Präsentation der ersten Sachverhalte ein Übungsdurchgang durchgeführt. Hatten die Probanden nach Durchlaufen des Übungsbeispiels keinerlei Fragen mehr, gelangten sie zum ersten Sachverhalt der zweiten Sitzung. Dieser Teil des Experimentes dauerte im Durchschnitt 25 Minuten pro Proband.

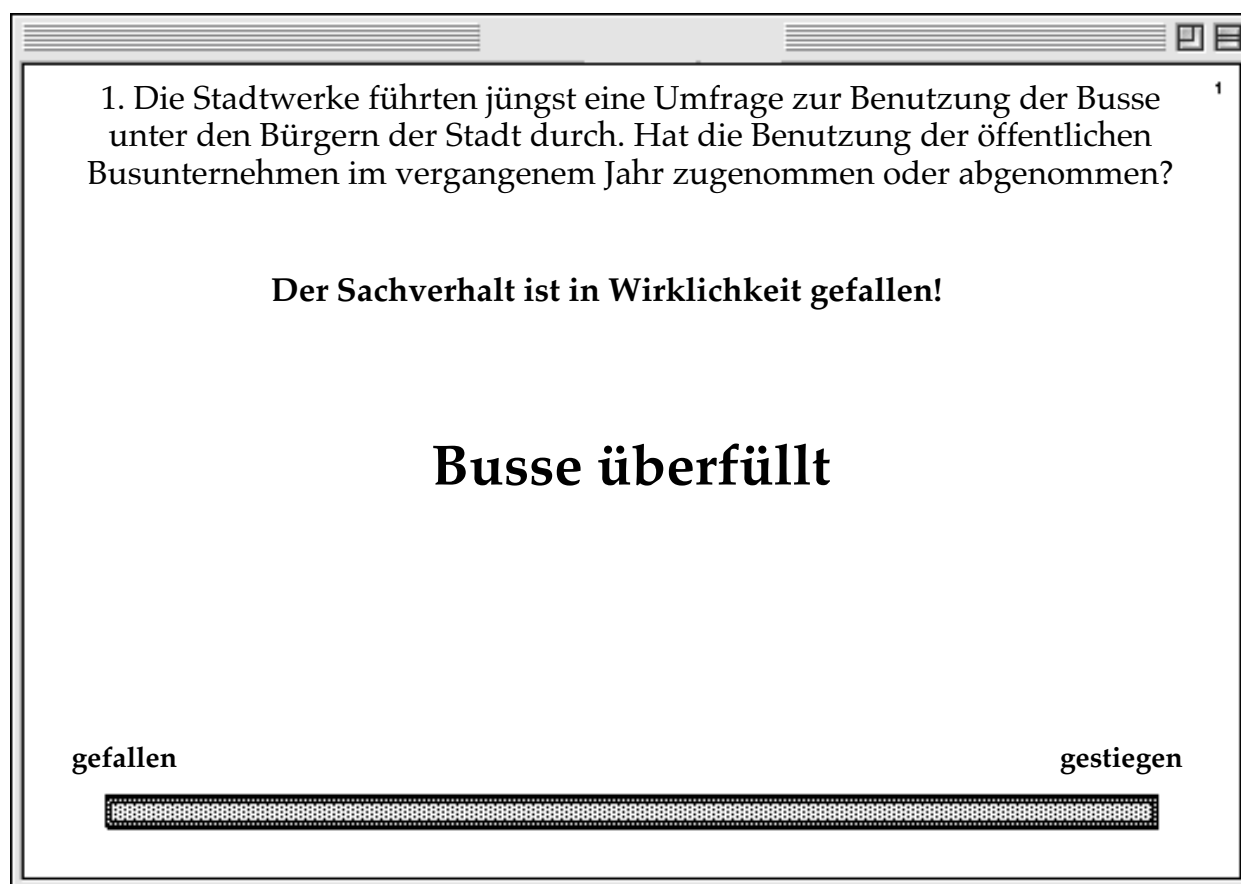


Abbildung 5.5: Bildschirm in Experiment 1 mit der Rekognitionsaufgabe: Die Probanden sollten entscheiden, ob es sich jeweils um ein neues oder altes Stichwort handelte.

Auswertung Um Extremwerte aus der Analyse auszuschließen, wurden einheitlich für alle Experimente alle Reaktionszeiten unter 200 msec aus der Auswertung ausgeschlossen. Bei Reaktionszeiten unter 200 msec ist davon auszugehen, daß die präsentierten Stichwörter nicht wahrgenommen werden konnten und daher keine sinnvolle Antwort vorlag (siehe z.B. Ratcliff & McKoon, 1981). Werte oberhalb der dreifachen Standardabweichung wurden ebenfalls von der Auswertung ausgeschlossen. Insgesamt wurden dadurch 659 von ursprünglich 13824 Werten ausgeschlossen. Das entspricht 4.8% der Daten.

ERGEBNISSE

Verteilung der Daten Betrachtet man zunächst die Verteilung von Schätzungen und Erinnerungen über die Sicherheit in Prozent, daß der Sachverhalt zugenommen habe, über alle Sachverhalte (siehe Abb. 5.6), sieht man, daß sowohl für die Rückmeldung „gestiegen“ und „gefallen“ als auch „ohne Rückmeldung“ die Mittelwerte der Verteilungen sehr eng beieinander lagen (Rückmeldung „gestiegen“: 55.82% für die Schätzung zu 55.21% für die Erinnerung; Rückmeldung „gefallen“: 56.67% zu 54.21% und „ohne“ Rückmeldung: 56.11% zu 53.58%). Die Verteilung der Schätzungen zeigte auch eine breite Streuung an, die sich dann bei

der Verteilung der Erinnerungen verminderte (Rückmeldung „gestiegen“: 23.30 zu 18.53; Rückmeldung „gefallen“: 22.89 zu 18.16 und „ohne“ Rückmeldung: 23.18 zu 18.74). Bezogen auf den systematisch variierten Faktor „Rückmeldung“ zeigten sich keinerlei augenfällige Unterschiede.

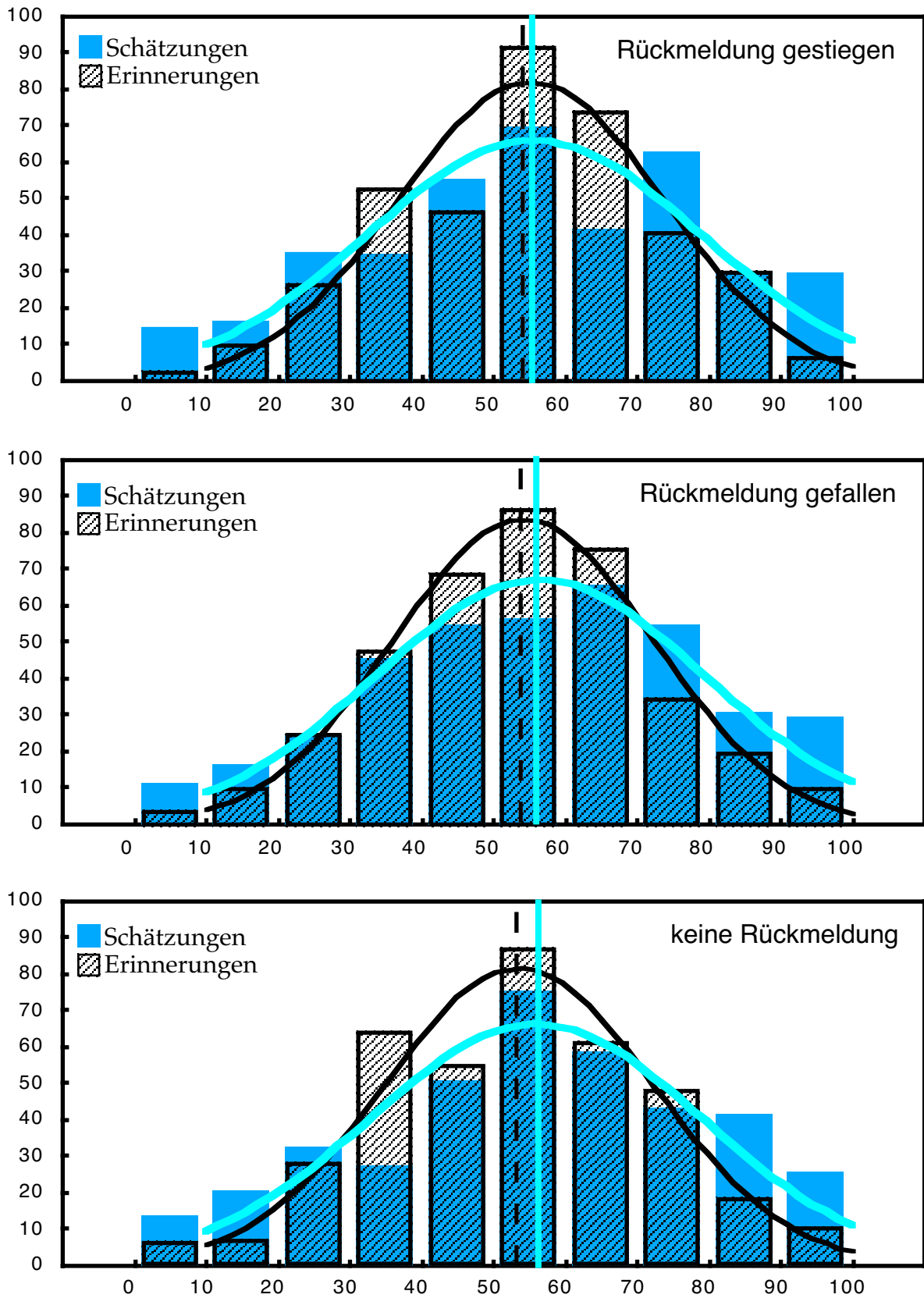


Abbildung 5.6: Relative Häufigkeiten für Schätzungen und Erinnerungen pro Zehner-Kategorie gliedert nach der Rückmeldung.

% **korrekt** Die Analyse der Anzahl korrekter Erinnerungen in der Kontroll- und den beiden Experimentalbedingungen erlaubt eine Beurteilung der Gedächtnisbeeinträchtigung durch die experimentelle Manipulation. Die Anzahl korrekter Erinnerungen ergab keinerlei signifikante Unterschiede ($\chi^2_{(2)}=1.64$). In der Kontrollbedingung traten 3.65% Fälle korrekter Erinnerung und in den Experimentalbedingungen 5.47% für die Rückmeldung „gestiegen“ und 5.21% für die Rückmeldung „gefallen“ auf.

Konfidenzurteile Die Analyse der Unterschiede in den Konfidenzurteilen zu den beiden Meßzeitpunkten erlaubt die Beurteilung, ob eine Veränderung der Konfidenz vorliegt. Herangezogen wurde dazu die Differenz ($\Delta\%$) in der subjektiven Sicherheit zum Zeitpunkt der Erinnerung minus der Sicherheit zum Zeitpunkt der ersten Einschätzung, daß der Sachverhalt „gestiegen“ ist. Ein positiver Wert von $\Delta\%$ bedeutet, daß zum zweiten Zeitpunkt die Sicherheit für „steigen“ zugenommen hat. Ein negativer Wert von $\Delta\%$ bedeutet, daß die Sicherheit für „steigen“ abgenommen hat.

In Abhängigkeit von der Rückmeldung bedeutet ein positiver Wert von $\Delta\%$ für die Rückmeldung „gestiegen“, daß das Konfidenzurteil sich in Richtung der Rückmeldung verändert hat, während er für die Rückmeldung „gefallen“ bedeutet, daß das zweite Konfidenzurteil entgegen der Richtung der Rückmeldung ausgefallen ist. Entsprechend umgekehrt verhält es sich, wenn die Differenz $\Delta\%$ negativ ausfällt. Für die Rückmeldung „gefallen“ bedeutet das eine Annäherung und für die Rückmeldung „gestiegen“ ein Entfernen von der Rückmeldung.

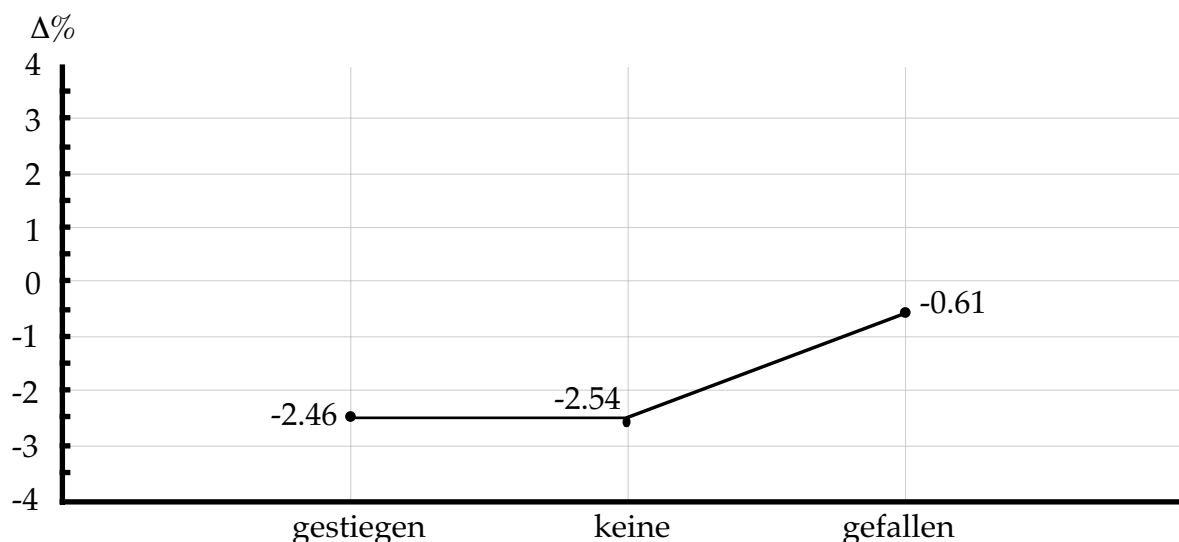


Abbildung 5.7: Differenz im Konfidenzurteil. Ein positiver Wert von $\Delta\%$ für die Rückmeldung „gestiegen“ sowie ein negativer Wert für die Rückmeldung „gefallen“ bedeuteten jeweils eine Annäherung der Erinnerung an die vorgegebene Lösung.⁶

Eine einfaktorielle Varianzanalyse mit dem Faktor „Rückmeldung“ (gestiegen/gefallen/keine) über die Differenz $\Delta\%$ zeigte keinen signifikanten Effekt der Rückmeldung ($F_{(2,126)}=1.52$): Die Rückmeldung „gefallen“ respektive „gestiegen“ und „keine“ zeigten keinerlei Unterschiede bezüglich $\Delta\%$. Es zeigten sich keinerlei Unterschiede bezüglich $\Delta\%$.

Wiedererkennensleistung Für die Beurteilung der Güte der Entscheidungen im Alt-Neu-Paradigma wurden Hits, False Alarms, die Sensitivität und der Bias erhoben (siehe: Macmillian, 1992; McNicol, 1972; Snodgrass, & Corwin, 1988).

Hits Eine zweifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (gestiegen/gefallen/keine) und „Argument“ (fallen/steigen) über die Treffer zeigte weder einen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,126)}=0.03$) noch der „Argumente“ ($F_{(1,63)}=2.06$). Die Interaktion zwischen der „Bedingung“ und der „Rückmeldung“ wurde ebenfalls nicht signifikant ($F < 1$).

False Alarms Eine einfaktorielle Varianzanalyse mit dem Faktor „Rückmeldung“ (gestiegen/gefallen/keine) über die false Alarms zeigte keinen signifikanten Effekt ($F_{(2,126)}=1.00$) (alle $t_{(63)} < 1.4$).

Diskriminationsleistung: Eine zweifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (gestiegen/gefallen/keine) und „Argument“ (fallen/steigen) über

⁶ Die Linien in dieser Grafik und auch in den folgenden Grafiken der Ergebnisdarstellung sollen keinerlei kontinuierliche Variablen implizieren. Sie sollen lediglich helfen die Zusammenhänge leichter zu erfassen.

d' zeigte weder einen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,126)}=0.53$) noch der „Argumente“ ($F_{(1,63)}=1.54$). Die Interaktion zwischen der „Bedingung“ und der „Rückmeldung“ wurde ebenfalls nicht signifikant ($F<1$).

Antwort-Bias Eine zweifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (gestiegen/gefallen/keine) und „Argument“ (fallen/steigen) über C zeigte weder einen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,126)}=0.77$) noch der „Argumente“ ($F_{(1,63)}=1.59$). Die Interaktion zwischen der „Bedingung“ und der „Rückmeldung“ wurde ebenfalls nicht signifikant ($F<1$).

Reaktionszeiten Eine dreifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (gestiegen/gefallen/keine), „Bekanntheit“ (alt/neu) und „Argument“ (fallen/steigen) über die Reaktionszeiten in Millisekunden auf die Argumente im Alt-Neu-Paradigma ergab weder einen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,126)}=.12$), der „Bekanntheit“ ($F_{(1,63)}=1.34$) noch der „Argumente“ ($F_{(1,63)}=2.72$). Keine Interaktion wurde signifikant (alle $F<1$).

DISKUSSION

Die Verteilung der Daten zeigte kaum Unterschiede zwischen den experimentellen Bedingungen. Die Mittelwerte zwischen Schätzungen und Erinnerungen unterschieden sich nur marginal. Zu erwarten gewesen wäre eine Verschiebung der Verteilung in Richtung der Rückmeldungen. Dies war bei der Rückmeldung „gestiegen“ nicht der Fall. Die Sicherheit, daß der Sachverhalt gestiegen wäre, hätte mit der gleichlautenden Rückmeldung zunehmen müssen und sich in einem höheren mittleren Konfidenzwert zeigen sollen. Die Konfidenz nahm hingegen ab. Sie fiel von 55.82% Sicherheit dafür, daß der Sachverhalt gestiegen wäre, auf nur noch 55.21%. Bei der Rückmeldung „gefallen“ zeigte sich ebenfalls eine Abnahme der Sicherheit dafür, daß der Sachverhalt gestiegen wäre, die diesmal konsistent zu der Rückmeldung erfolgte: Die Sicherheit dafür, daß der Sachverhalt gestiegen wäre, fiel von 56.67% bei der Schätzung auf nur noch 54.21% bei der Erinnerung. Allerdings zeigte sich das gleiche Muster bei Sachverhalten ohne Rückmeldung. Hier fiel die Sicherheit dafür, daß der Sachverhalt gestiegen wäre, von 56.11% auf nur noch 53.58% bei der Erinnerung. Zusammen mit der Tatsache, daß die Streuung ebenfalls insgesamt, ohne Unterschiede in dem variiertem Faktor „Rückmeldung“ enger wurde, ergibt sich das Bild einer klaren Regression zur Mitte. Die Mittelwerte näherten sich allesamt klar der 50%-Marke und die Streuung wurde ebenfalls geringer. Diese Befunde sprechen schon klar gegen einen systematischen Einfluß der Rückmeldung und deuten daraufhin, daß hier kein Rückschaufehler aufgetreten ist.

Die Analyse des Anteils korrekter Erinnerungen zeigte ebenfalls keinerlei Unterschiede zwischen Sachverhalten mit und ohne Rückmeldung. Die geringen Anteile korrekter Erinnerungen von 5% waren zu erwarten, da die Probanden, sowohl Schätzung und Erinnerung auf einer stufenlosen Skala am Bildschirm eintrugen, ohne Rückmeldung konkreter Zahlen. Die Aufgabe, sich genau die Stelle auf der Skala zu merken, auf der man vor einer Woche den Regler auf der Konfidenzskala gestellt hatte, war zu schwer zu lösen. Daher waren die Aussichten, im Anteil korrekter Erinnerungen einen Unterschied zwischen den systematisch variierten Bedingungen zu erhalten, eher gering.

Die Ergebnisse der einfaktoriellen Varianzanalyse über den Faktor „Rückmeldung“ mit der abhängigen Variablen $\Delta\%$ standen im Einklang mit der Betrachtung der Verteilungen. Eine systematische Annäherung der erinnerten Konfidenzurteile an die Rückmeldung entspräche einem Rückschaufehler. In der vorliegenden Untersuchung zeigten sich keinerlei signifikante Unterschiede im systematisch variierten Faktor „Rückmeldung“ und mithin zeigte sich kein Rückschaufehler. Das Ausbleiben des erwarteten Effekts einer systematischen Beeinflussung durch die Rückmeldung ist zunächst einmal überraschend. In der Metaanalyse von Christensen-Szalanski & Willham (1991) fanden sich unter 128 Experimenten lediglich sechs Experimente, die keinen oder einen umgekehrten Effekt zeigten.

Mögliche Gründe für das Ausbleiben des Effekts lagen eventuell in dem verwendeten Versuchsmaterial. Bei der Wahl der Sachverhalte wurde sehr viel Wert darauf gelegt, möglichst neutrale Sachverhalte zu präsentieren, bei denen wenig präexperimentelle Präferenzen zu erwarten waren. Dies sollte gewährleisten, daß die präsentierten Argumente zur Veränderung der Sachverhalte in die jeweilige Richtung, „gestiegen“ respektive „gefallen“, maximal wirken sollten. Das könnte dazu geführt haben, daß eine Entscheidung für einen Anstieg bzw. eine Abnahme des zu beurteilenden Sachverhalts schwer getroffen werden konnte und daher die Mehrzahl der Entscheidungen in der Nähe der 50% Konfidenz lagen. Schon auf deskriptiver Ebene war eine Regression zur Mitte zu beobachten. Die Mittelwerte der Erinnerung an die Schätzung tendierten daher alle zur 50%-Marke, unabhängig davon, ob eine Rückmeldung erfolgt war oder nicht und ob diese „gestiegen“ oder „gefallen“ gelautet hatte. Hinzu kommt, daß für die Abgabe der Schätzung, der Enkodierung der Lösung und die Erinnerung an die eigenen Schätzungen keinerlei Zeitbeschränkung vorlag. Das kann dazu geführt haben, daß eine Entscheidung erst nach reiflicher Überlegung und sorgfältigem Abwägen der Argumente für und wider dem Anstieg des zu beurteilenden Sachverhalts getroffen wurde. Tatsächlich lag die Zeit pro Sachverhalt im Mittel bei ca. 2 Minuten und damit relativ hoch. Die

Auswertung der Bedeutsamkeitseinschätzung der Argumente ergab zwar, daß über alle Sachverhalte die Argumente für eine Zunahme des Sachverhaltes (66.15%) wichtiger eingeschätzt wurden als Argumente, die für eine Abnahme des Sachverhaltes (60.69%) sprachen ($t_{67}=-7.02$). Da diese Einschätzung aber über alle Fragen gleich war, führte dies lediglich zu einer leichten Bevorzugung des Ausgangs, daß in Wirklichkeit der Sachverhalt gestiegen wäre. Insgesamt zeigte sich, daß alle Argumente genügend wichtig genommen wurden. Es gab kein einziges Argument, dessen Wichtigkeit unter 55% eingestuft wurde. Diese Analyse spricht daher eher für die Annahme, daß eine sehr ausgewogene Entscheidung getroffen wurde. Diese lag dann sehr nahe der 50%-Marke und war dann auch in der Erinnerungsaufgabe sehr leicht zu erinnern. Da dies aufgrund der verwendeten Konfidenzskala nicht zu einem hohen Prozentsatz korrekter Erinnerungen führen konnte, kam es zu einer Regression zur Mitte.

Das Ausbleiben des Rückschaufehlers in diesem Experiment läßt dann selbstverständlich auch keine systematischen Effekte in den anderen abhängigen Variablen erwarten, da das Hauptaugenmerk bei der Analyse der Wiedererkennensleistung und Reaktionszeiten natürlich auf einer systematischen Veränderung aufgrund der Rückmeldung lag. Wie zu erwarten war, zeigten sich in keiner Analyse signifikante Effekte.

Der ausgebliebene Rückschaufehler forderte Konsequenzen für das folgende Experiment. Typischerweise läßt sich der Rückschaufehler nur bei unsicherem Wissen nachweisen. Die experimentellen Paradigmen, etwa die Beantwortung von Almanachfragen, erfordern typischerweise eine zügige Entscheidung, schließlich gilt es, zwischen 40 und 80 Fragen zu beantworten. Der Effekt minimiert sich, je besser die Erinnerung an die Abgabe der Schätzung ist und je besser diese elaboriert wurde (siehe Kap. 2..2). Die Ergebnisse von Experiment 1 sprachen deutlich dafür, daß diese wichtige Randbedingung nicht eingehalten wurde. Insbesondere hatten sich die Probanden zu lange mit den Sachverhalten auseinandergesetzt und hatten daher die Gelegenheit, die Sachverhalte besser zu enkodieren. Hinzu kommt, daß Schätzungen in der Nähe der 50% auf der Konfidenzskala ähnlich salient wie die Endpunkte der Skala sind. Der Anteil korrekter Erinnerungen in diesem Experiment wäre vermutlich sehr hoch gewesen, hätten die Probanden eine numerische Schätzung abgegeben. Die Rückmeldung konnte daher auch nicht die gewünschte Wirkung auf die erinnerte Konfidenz zeigen.

Im folgenden zweiten Experiment wurden die Bearbeitungszeiten verkürzt. Es sollte damit erreicht werden, daß die Zeit zur gründlichen Auseinandersetzung mit allen

Argumenten für und wider nicht ausreichte und damit eine deutlich gefärbtere Entscheidung getroffen werden würde. Die Sachverhalte sollten dadurch weniger elaboriert werden und es sollte auch der Einfluß der Rückmeldung wieder verstärkt werden.

5.2 Experiment 2: Wiedererkennen mit komprimierter Schätzphase

EINLEITUNG

Das nachfolgende zweite Experiment behielt das grundsätzliche experimentelle Vorgehen des ersten Experimentes bei. Es wurden auch in diesem Experiment die Image-Sets experimentell vermittelt. Die Aufgabe bestand wieder darin, in 18 Verifikationsaufgaben zu beurteilen, ob die dort beschriebenen Sachverhalte in den letzten Jahren gestiegen oder gefallen sind, und sich zum Testzeitpunkt, genau eine Woche später, wieder an das eigene Urteil zu erinnern. Zu jedem Sachverhalt gab es, wie auch schon im ersten Experiment, zwölf Argumente, wobei jeweils sechs Argumente dafür sprachen, daß der Sachverhalt gestiegen, und sechs dafür, daß er gefallen ist. Dargeboten wurden jeweils nur vier der sechs Argumente. Die nicht dargebotenen Argumente dienten als neue Argumente für den Wiedererkennenstest. Im Unterschied zu Experiment 1 wurde in diesem zweiten Experiment besonderen Wert auf eine zügige Bearbeitung gelegt. Eine zu gründliche Auseinandersetzung mit den Sachverhalten und ihren Argumenten sollte verhindert werden. Gleichzeitig sollte dennoch gewährleistet sein, daß die Argumente nicht einfach ignoriert werden würden. Daher wurde die Durchführung gegenüber dem ersten Experiment leicht modifiziert.

Die Aufgabe zum zweiten Termin, genau eine Woche nach dem ersten, war wiederum, sich möglichst genau an das Urteil aus der ersten Sitzung zu erinnern. Zusätzlich wurde das Wiedererkennen der Argumente geprüft. In einem Alt-Neu-Paradigma wurden die Reaktionszeiten für die acht alten gegenüber vier neuen Argumenten erhoben.

Die Erinnerung an die Konfidenzurteile sollte sich in den Experimentalbedingungen systematisch von der Kontrollbedingung unterscheiden: Die Rückmeldung, ein Sachverhalt wäre gestiegen, sollte die Sicherheit für „steigen“ erhöhen bzw. für „fallen“ vermindern, während umgekehrt die Rückmeldung, daß ein Sachverhalt gefallen wäre, die Sicherheit für „fallen“ erhöhen bzw. für „steigen“ mindern sollte. Ein solches Ergebnis entspräche einem typischen Rückschlußfehler. Im Alt-Neu-Paradigma sollte sich ebenfalls ein Zusammenhang mit der Rückmeldung zeigen lassen. Argumente, die für „fallen“ sprechen, sollten häufiger und schneller erkannt werden als solche, die für „steigen“ sprechen, wenn die Rückmeldung „fallen“ lautet, und entsprechend umgekehrt. Dieses Ergebnis spräche für eine selektive Aktivierung des Image-Sets.

METHODE

Design Der Versuchsplan bestand, wie im ersten Experiment, aus einem 3×2 Design mit den Faktoren „Rückmeldung“ (gestiegen/gefallen/keine) und „Argument“ (fallen/steigen). Beide Faktoren wurden als Meßwiederholungsfaktoren realisiert. Die Zuordnung zu dem Faktor „Rückmeldung“ wurde durch systematische Permutation einer einmaligen Zufallsreihenfolge realisiert, so daß alle Fragen über alle Bedingungen gleich häufig verwirklicht wurden. Eine mit „gestiegen“ bezeichnete Rückmeldung zeigte an, daß der zu beurteilende Sachverhalt in Wirklichkeit gestiegen war. Entsprechend zeigte eine mit „gefallen“ bezeichnete Rückmeldung an, daß der zu beurteilende Sachverhalt in Wirklichkeit gefallen war. Die Bedingung „keine“ Rückmeldung diente als Kontrollbedingung. Der Faktor „Argument“ gruppiert die Wiedererkennensleistung der Argumente in Argumente, die dafür sprachen, daß der entsprechende Sachverhalt gestiegen („steigen“) und Argumente, die dafür sprachen, daß er gefallen („fallen“) war. Die Argumente sprachen jeweils zur Hälfte dafür, daß der entsprechende Sachverhalt gestiegen respektive gefallen war. Die Darbietung erfolgte blockweise, wobei die Reihenfolge der Argumente als auch der Blöcke selbst einer Zufallsreihenfolge folgte, so daß alle Argumente gleich häufig an jeder Position präsentiert wurden.

Als abhängige Maße wurden neben der Differenz der Konfidenzurteile zum Zeitpunkt der ersten Einschätzung und dem Zeitpunkt der Erinnerung ($\Delta\%$) die Antworten und Reaktionszeiten in einer Alt-Neu-Entscheidung erhoben.

Probanden An der Untersuchung nahmen 50 Personen, vorwiegend Studierende aus verschiedenen Fachbereichen der Universität Trier, teil. Davon waren 30 weiblichen und 20 männlichen Geschlechts. Das Altersspektrum reichte von 19 bis 34 Jahren. Der Altersdurchschnitt lag bei 24.6 Jahren. Die Probanden erhielten 12 DM für die Teilnahme am Experiment.

Material Die gleichen 18 Verifikationsaufgaben aus dem ersten Experiment, in denen die Probanden wiederum ihre Sicherheit angeben sollten, mit der sie glaubten, daß der entsprechende Sachverhalt gestiegen oder gefallen war. Die Beurteilung der Sachverhalte erfolgte jeweils durch Verschiebung einer Markierung auf einer analogen Skala auf dem Computerbildschirm, die von „gefallen“ zu „gestiegen“ reichte. Die Probanden bekamen keine Rückmeldung des konkreten numerischen Skalenwertes.

Wie auch schon im vorangegangenen Experiment, wurde den Probanden zufällig jeweils vier von den sechs Argumenten dargeboten, die dafür sprachen, daß der

Sachverhalt gefallen war und vier von den sechs Argumenten, die dafür sprachen, daß der Sachverhalt gestiegen war. Um die Präsentation übersichtlicher zu gestalten und weil das Platzangebot auf dem Bildschirm begrenzt war, wurden die Argumente auf dem Bildschirm in Stichworten zusammengefaßt. Zunächst wurde den Probanden jedes Argument vollständig ausformuliert zusammen mit einem Stichwort dargeboten. Im Unterschied zum ersten Experiment sollten die Probanden, nachdem sie sich das Argument durchgelesen hatten, direkt auf einer analogen Skala dessen Wichtigkeit einschätzen und anschließend die Maustaste betätigen, damit das nächste Argument auf dem Bildschirm erscheinen konnte. Nachdem alle Argumente präsentiert worden waren, blieben auf dem Bildschirm, in der Reihenfolge ihrer Darbietung, nur die Stichwörter zu den Argumenten sichtbar. Die Probanden konnten aber jederzeit durch Klicken auf ein Stichwort das entsprechende Argument nochmals lesen.

Ablauf Die Durchführung folgte dem Gedächtnisdesign und fand in zwei Sitzungen mit einer Woche Abstand statt. Der gesamte Versuch erfolgte wieder am Rechner. Es bestand jeweils für vier Probanden gleichzeitig die Möglichkeit teilzunehmen. Der Ablauf erfolgte analog zum ersten Experiment mit dem Unterschied, daß sowohl von Seiten der Versuchsleiter als auch in den Instruktionen besonderer Wert auf eine zügige Beantwortung der Sachverhalte gelegt wurde. Die Einschätzung der Wichtigkeit der Argumente wurde vorgezogen. Damit sollte ebenfalls eine zügigere Bearbeitung der ersten Aufgabe ermöglicht werden.

Wie im ersten Experiment wurden die Probanden in der ersten Sitzung zunächst in einer ausführlichen Instruktion am Bildschirm aufgeklärt, daß sie im folgenden 18 Sachverhalte danach beurteilen sollten, ob der betreffende Sachverhalt gestiegen oder gefallen wäre. Sie wurden informiert, daß sie ihre Antwort in einer analogen Skala am unteren Bildschirmrand eintragen sollten, wobei der jeweilige Abstand zu den Extrempunkten der Skala, die mit „gestiegen“ bzw. „gefallen“ bezeichnet waren, die subjektive Sicherheit ihres Urteils anzeigen sollte. Eine Antwort nahe einem Extrempunkt drücke hohe Sicherheit, eine Antwort in der Nähe der Mitte der Skala Unsicherheit aus. Sie wurden informiert, daß sie zu jedem Sachverhalt acht Argumente präsentiert bekämen. Sie wurden angehalten, diese zu beachten, da sie für den weiteren Verlauf des Experimentes wichtig seien und ihnen ihre Entscheidung erleichtern sollten. Die Probanden wurden darüber informiert, daß aus Platzgründen auf dem Bildschirm jeweils nur noch Stichwörter zu den Argumenten erschienen, sie aber jeweils durch Klicken auf ein entsprechendes Stichwort die Gelegenheit hätten, das zugehörige Argument nochmals vollständig zu sehen. Weiterhin wurden die Probanden instruiert, daß es zur Evaluation der Argumente

notwendig wäre, jedes Argument in seiner Wichtigkeit einzuschätzen. Dazu erschiene nach jedem Stichwort eine analoge Skala, auf der sie durch Mausklick einen Regler so einstellen könnten, daß er ihre Einschätzung der Wichtigkeit dieses Arguments wiedergäbe. Die Instruktion wies die Probanden an, nicht zu viel Zeit dem Studium der Argumente zu widmen und die Entscheidung, ob der Sachverhalt gestiegen oder gefallen wäre, möglichst spontan zu treffen. Das Hauptinteresse der Untersuchung läge auf schnellen spontanen Entscheidungen, wie sie auch in der Wirtschaft üblich wären, verbunden mit einer sehr kurzen Abwägung der Argumente für und wider eines Sachverhaltes. Vor der Präsentation der ersten Frage wurde ein Übungsdurchgang (mit nicht ausgewerteten Sachverhalten) durchgeführt.

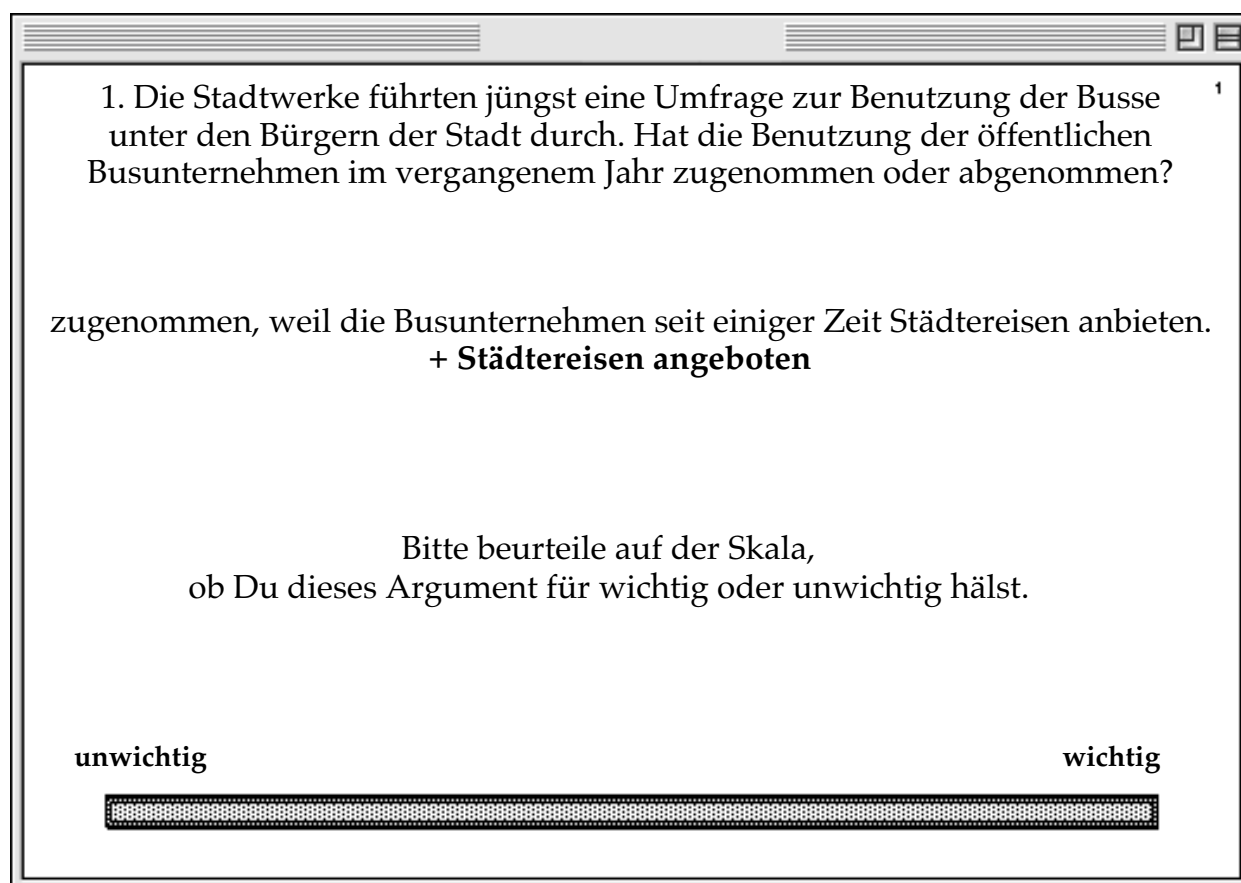


Abbildung 5.8: Bildschirm in Experiment 2 bei der Präsentation eines Arguments mit zugehörigem Stichwort (unterhalb des Arguments durch fetten Schriftschnitt hervorgehoben). Die untere Skala diente gleichzeitig zur Einschätzung der Wichtigkeit der Argumente.

Zu Beginn eines Versuchsdurchgangs erschien am oberen Teil des Bildschirms jeweils ein Sachverhalt. Anschließend wurden nacheinander in zufälliger Reihenfolge, jeweils einzeln, die acht Argumente dargeboten. Es erschien am Bildschirm jeweils das Argument, gefolgt von einem textlich abgehobenen Stichwort

(siehe Abb. 5.8). Die Probanden sollten die Wichtigkeit des Arguments einschätzen. Dazu mußten Sie mit der Maus am unteren Bildschirmrand auf eine analoge Skala klicken, deren jeweilige Endpunkte mit „unwichtig“ bzw. „wichtig“ gekennzeichnet waren. Durch Klicken auf die Skala wurde an dieser Stelle ein Schieberegler sichtbar, der mit der Maus verschoben werden konnte. Gleichzeitig erschien auf dem Bildschirm eine Taste mit der Aufschrift „weiter“, mit der die Probanden zum nächsten Argument gelangen konnten. Nach dem letzten Argument erschien am unteren Bildschirmrand eine analoge Skala, deren jeweilige Endpunkte mit „gestiegen“ und „gefallen“ gekennzeichnet waren. Auf dem Bildschirm war zu diesem Zeitpunkt der Sachverhalt und in der Reihenfolge ihres Erscheinens die Stichworte zu den Argumenten und die Skalen zu sehen. Den Probanden war es freigestellt, durch Klicken auf ein Stichwort nochmals das ausformulierte Argument zu sehen.

Durch Klicken auf die Skala wurde an dieser Stelle wieder ein Schieberegler sichtbar, der mit der Maus verschoben werden konnte. Gleichzeitig erschien auf dem Bildschirm eine Taste mit der Aufschrift „weiter“, mit der die Probanden zum nächsten Sachverhalt gelangen konnten. Erfolgte die Entscheidung der Probanden nicht schnell genug, erschien 20 Sekunden nach der Einschätzung des letzten Arguments eine Aufforderung, sich jetzt zu entscheiden. Der nächste Sachverhalt wurde dennoch erst eingeblendet, wenn die Probanden sich entschieden hatten. Dieser Teil des Experimentes dauerte im Durchschnitt 25 Minuten pro Proband.

Die zweite Sitzung, genau eine Woche nach der ersten, erfolgte vollkommen analog zum ersten Experiment. Mit der Ausnahme, daß in den Instruktionen nochmals Wert darauf gelegt wurde, sich zügig zu entscheiden. Die Probanden wurden instruiert, daß nun die gleichen Sachverhalte wie in Sitzung 1 nochmals dargeboten würden und es darum ginge, sich jetzt möglichst genau an die eigenen Schätzungen aus der vorangegangenen Sitzung zu erinnern. Sie wurden informiert, daß ihnen zu einigen Fragen das tatsächliche Ergebnis, ob der Sachverhalt gestiegen oder gefallen wäre, präsentiert werde.

Wie in Experiment 1 erschien nach der Erinnerung an das Konfidenzurteil eine zweite Instruktion, die die Probanden darüber informierte, daß ihnen im folgenden jeweils nach Abgabe ihrer Antworten alte und neue Stichworte zu den Sachverhalten präsentiert werden würden. Ihre Aufgabe wäre es, möglichst schnell und korrekt zu entscheiden, ob es sich um ein altes, im Zusammenhang mit diesem Sachverhalt schon genanntes Stichwort, oder um ein neues Stichwort handelte. Hatten die

Probanden nach Durchlaufen des Übungsbeispiels keinerlei Fragen mehr, gelangten sie zum ersten Sachverhalt der zweiten Sitzung.

Die Reihenfolge der Sachverhalte und der Bildschirmaufbau wurde beibehalten, außer daß zum Zeitpunkt der Erinnerung an die Antwort aus der ersten Sitzung die Stichworte und Argumente nicht präsentiert wurden. Nachdem die Probanden zum jeweiligen Sachverhalt ihr Urteil von der ersten Sitzung zu replizieren versucht hatten, wurden ihnen in zufälliger Reihenfolge die acht alten und vier neue Argumente präsentiert. Anschließend gelangten sie zum nächsten Sachverhalt. Dieser Teil des Experimentes dauerte im Durchschnitt 25 Minuten pro Proband.

Auswertung Um Extremwerte aus der Analyse auszuschließen, wurden wiederum alle Reaktionszeiten unter 200 msec aus der Auswertung ausgeschlossen. Werte oberhalb der dreifachen Standardabweichung wurden ebenfalls von der Auswertung ausgeschlossen. Insgesamt wurden dadurch 410 von ursprünglich 10800 Werten ausgeschlossen. Das entspricht 3.8% der Daten.

ERGEBNISSE

Verteilung der Daten Es zeigte sich bei der Verteilung von Schätzungen und Erinnerungen über die Sicherheit in Prozent, daß der Sachverhalt zugenommen habe, über alle Sachverhalte (siehe Abb. 5.9), daß sowohl für die Rückmeldung „gestiegen“ und „gefallen“ als auch „ohne Rückmeldung“ die Mittelwerte der Verteilungen sehr eng beieinander lagen (Rückmeldung „gestiegen“: 55.86% für die Schätzung zu 56.45% für die Erinnerung; Rückmeldung „gefallen“: 52.01% zu 50.05% und „ohne“ Rückmeldung: 55.99% zu 52.34%). Die Verteilung der Schätzungen zeigten auch eine breite Streuung an, die sich dann bei der Verteilung der Erinnerungen verminderte (Rückmeldung „gestiegen“: 24.43 zu 22.16; Rückmeldung „gefallen“: 26.13 zu 22.89 und „ohne“ Rückmeldung: 26.39 zu 22.69). Bezogen auf den systematisch variierten Faktor „Rückmeldung“ zeigten sich keinerlei augenfällige Unterschiede.

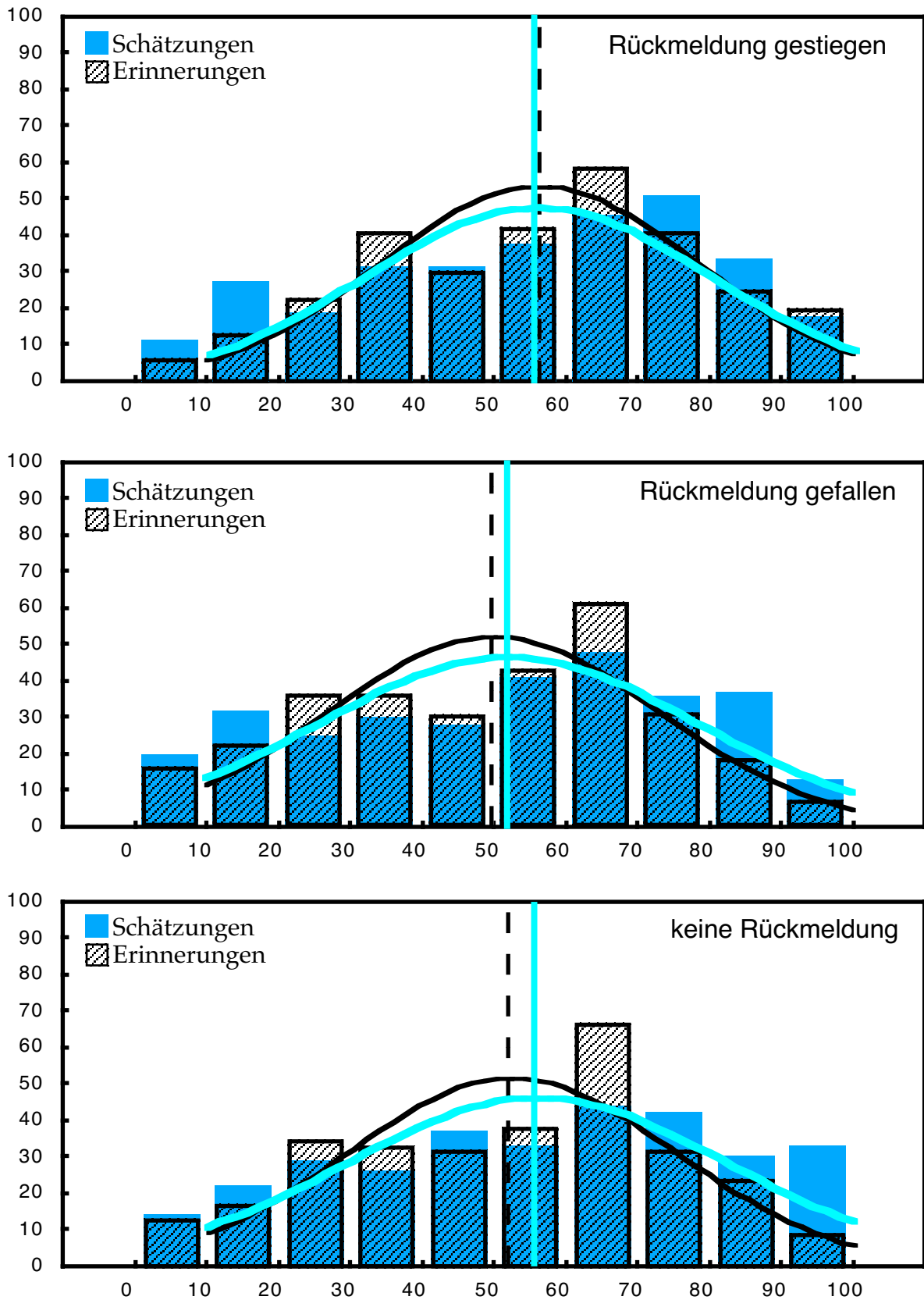


Abbildung 5.9: Relative Häufigkeiten für Schätzungen und Erinnerungen pro Zehner-Kategorie gegliedert nach der Rückmeldung.

% korrekt Die Analyse der Anzahl korrekter Erinnerungen in der Kontroll- und den beiden Experimentalbedingungen erlaubt eine Beurteilung der Gedächtnisbeeinträchtigung durch die experimentelle Manipulation. Die Anzahl korrekter Erinnerungen ergab signifikante Unterschiede ($\chi^2_{(2)}=6.85$; Power (1-beta)=0.46). In der Kontrollbedingung zeigte sich in 4.00% der Fälle eine korrekter Erinnerung an das eigene Konfidenzurteil. In den Experimentalbedingungen zeigte sich für die Rückmeldung „gestiegen“ in nur 1.66% und für die Rückmeldung „gefallen“ in lediglich 1.00% der Fälle eine korrekte Erinnerung an das eigene Konfidenzurteil aus der ersten Sitzung.

Konfidenzurteile Eine einfaktorielle Varianzanalyse mit dem Faktor „Rückmeldung“ (gestiegen/gefallen/keine) über die Differenz $\Delta\%$ zeigte keinen signifikanten Effekt der Rückmeldung ($F_{(2,98)}=1.58$): Die Rückmeldung „gefallen“ respektive „gestiegen“ und „keine“ zeigten keinerlei Unterschiede bezüglich $\Delta\%$ (siehe Abb. 5.10).

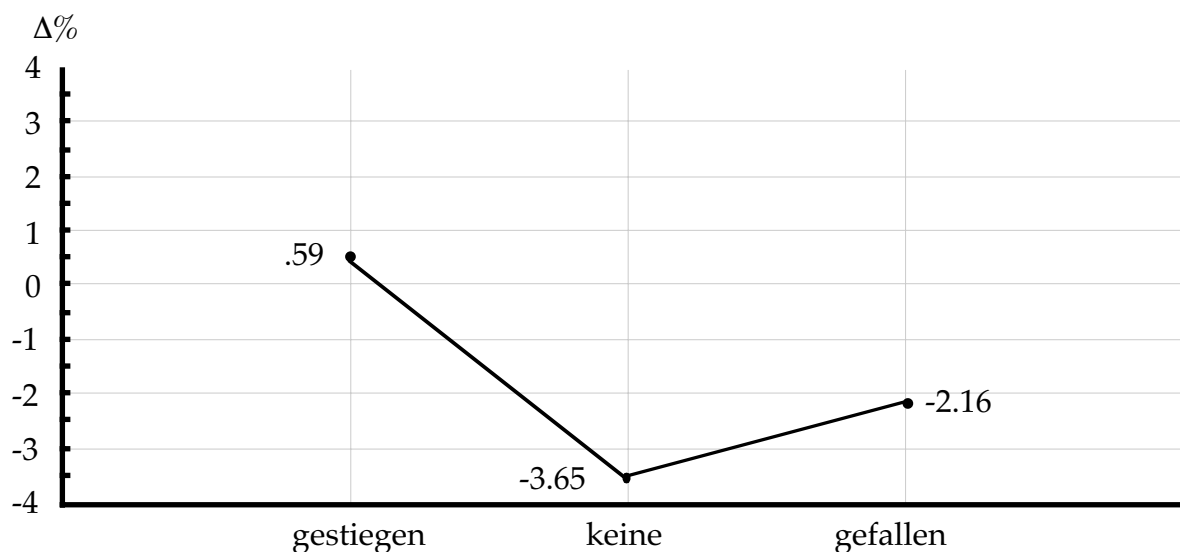


Abbildung 5.10: Differenz im Konfidenzurteil. Ein positiver Wert von $\Delta\%$ für die Rückmeldung „gestiegen“ sowie ein negativer Wert für die Rückmeldung „gefallen“ bedeuteten jeweils eine Annäherung der Erinnerung an die vorgegebene Lösung.

Wiedererkennensleistung Wie schon im ersten Experiment wurden für die Beurteilung der Güte der Entscheidungen im Alt-Neu-Paradigma Hits, False Alarms, die Sensitivität und der Bias erhoben.

Hits Eine zweifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (gestiegen/gefallen/keine) und „Argument“ (fallen/steigen) über die Treffer zeigte weder einen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,98)}=0.26$) noch der „Argumente“ ($F_{(1,49)}=3.60$). Die Interaktion zwischen der „Bedingung“ und der

„Rückmeldung“ wurde allerdings signifikant ($F_{(2,98)}=3.22$, Power (1-beta)=0.94): Es gab mehr Treffer in der Bedingung „keine“ Rückmeldung für Argumente, die für „steigen“ (.66) als für solche, die für „fallen“ sprachen (.61) ($t_{(49)}=3.52$; $p=.0009$). Die anderen Einzelvergleiche erreichten keine Signifikanz (alle $t<1.24$) (siehe Abb. 5.11).

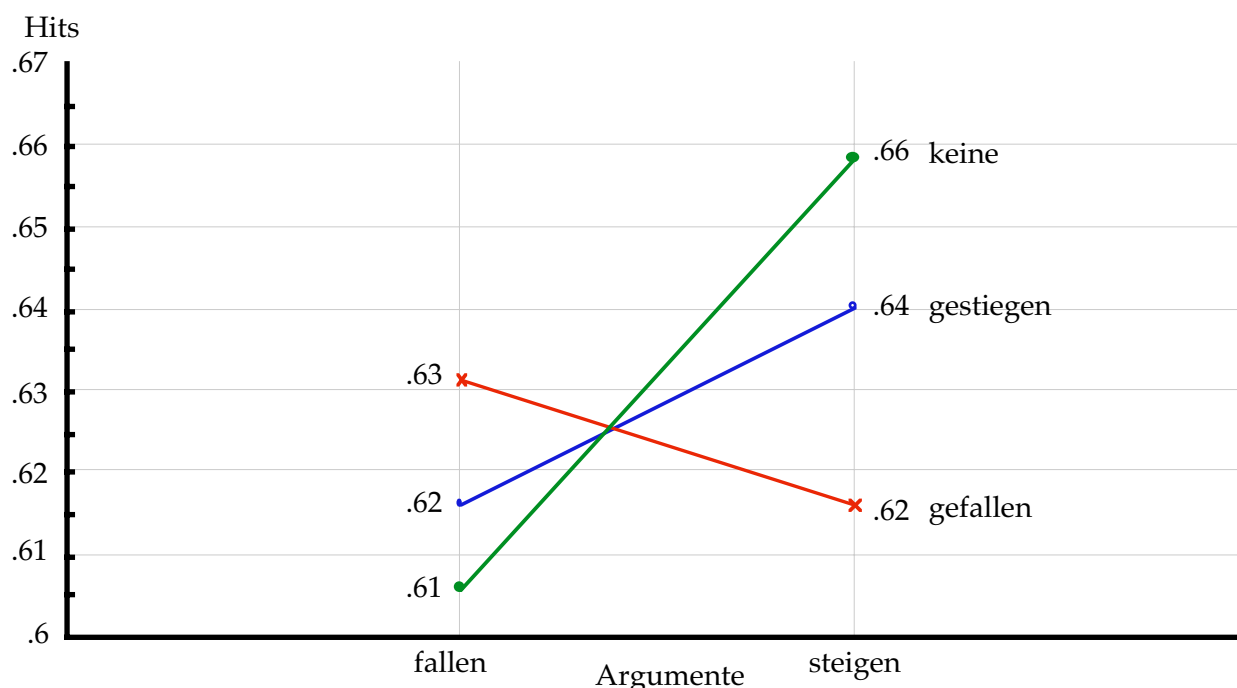


Abbildung 5.11: Interaktion zwischen Rückmeldung (gestiegen/gefallen/keine) und Argumenten (fallen/steigen) über die Treffer.

False Alarms Eine einfaktorielle Varianzanalyse mit dem Faktor „Rückmeldung“ (gestiegen/gefallen/keine) über die false Alarms zeigte keinen signifikanten Effekt ($F_{(2,98)}=.08$) (alle $t_{(49)}<.09$).

Diskriminationsleistung Eine zweifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (gestiegen/gefallen/keine) und „Argument“ (fallen/steigen) über d' zeigte keinen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,98)}=0.16$). Die „Argumente“ unterschieden sich jedoch signifikant ($F_{(1,49)}=4.25$, Power (1-beta)=0.97): Probanden reagierten auf Argumente für „steigen“ (-0.02) sensibler als auf solche für „fallen“ (-0.07). Die Interaktion zwischen der „Rückmeldung“ und den Faktor „Argument“ wurde ebenfalls signifikant ($F_{(2,98)}=3.42$, Power (1-beta)=0.94): Auf Argumente, die für „steigen“ sprachen (.04), wurde in der Bedingung „keine“ Rückmeldung sensibler reagiert als auf solche, die für „fallen“ sprachen (-0.09) ($t_{(49)}=3.65$; $p=.0006$). Alle anderen Einzelvergleiche erreichten keine Signifikanz (alle $F<1.36$; $t<1.43$) (siehe Abb. 5.12).

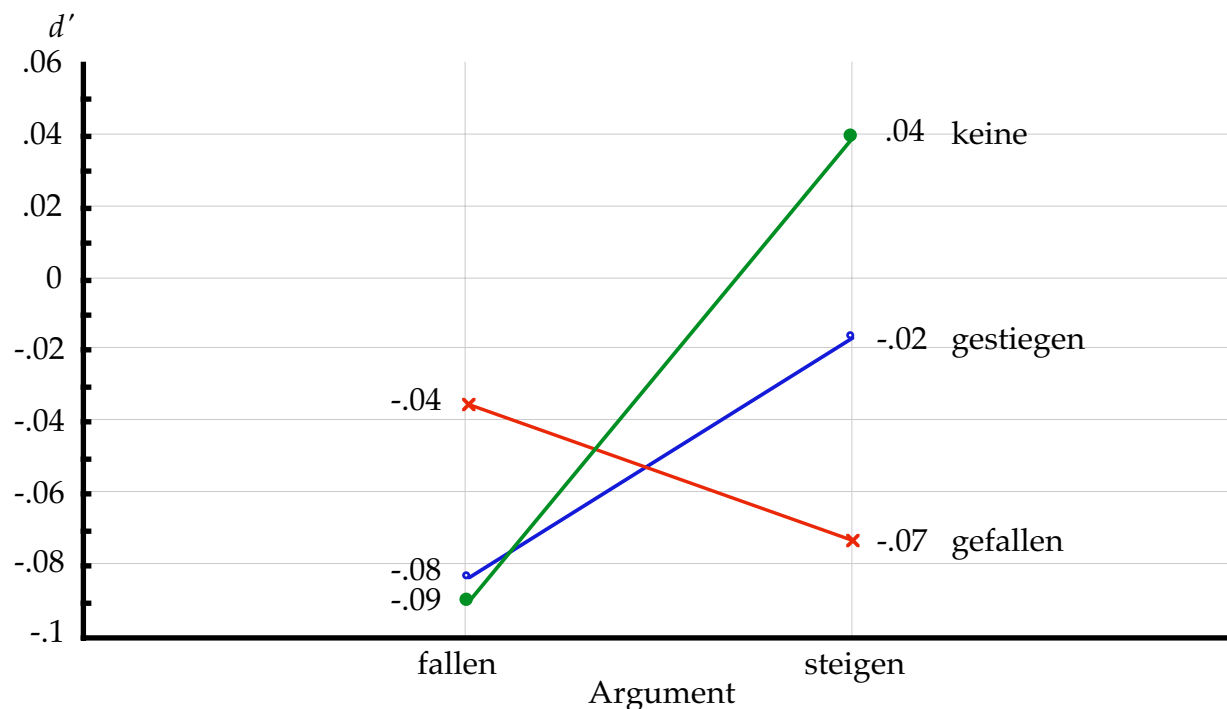


Abbildung 5.12: Interaktion Rückmeldung (gestiegen/gefallen/keine) mit Argumente (fallen/steigen) über die Diskriminationsleistung d' .

Antwort-Bias Eine zweifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (gestiegen/gefallen/keine) und „Argument“ (fallen/steigen) über C zeigte keinen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,98)}=0.08$). Die „Argumente“ unterschieden sich jedoch signifikant ($F_{(1,49)}=4.45$, Power (1-beta)=0.97): auf Argumente für „steigen“ (-.34) wurde liberaler reagiert als auf solche für „fallen“ (-.31). Die Interaktion zwischen der „Rückmeldung“ und dem Faktor „Argumente“ wurde ebenfalls signifikant ($F_{(2,98)}=3.40$, Power (1-beta)=0.94): Auf Argumente, die für „steigen“ sprachen (-.36), wurde in der Bedingung „keine“ Rückmeldung liberaler reagiert als auf solche, die für „fallen“ sprachen (-.29) ($t_{(49)}=3.62$; $p=.0007$). Alle anderen Einzelvergleiche erreichten keine Signifikanz (alle $F < 1.83$; $t < 1.50$) (siehe Abb. 5.13).

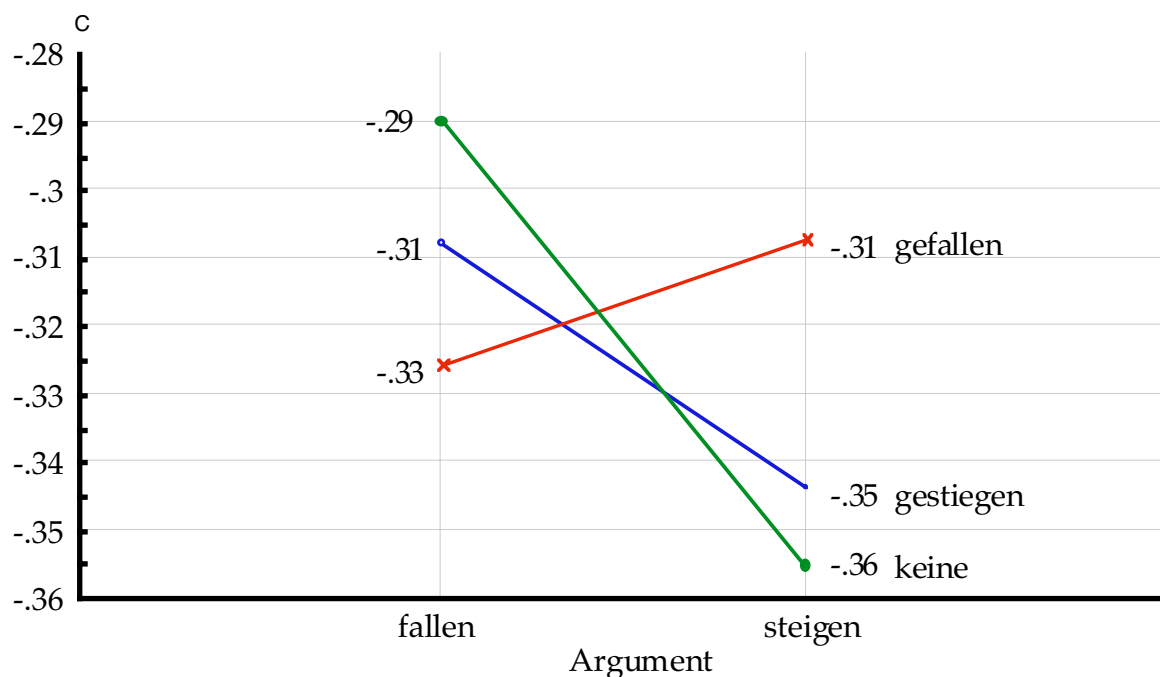


Abbildung 5.13: Interaktion Rückmeldung (gestiegen/ gefallen/ keine) mit Argumente (fallen/ steigen) über den Antwort-Bias C.

Reaktionszeiten Eine dreifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (gestiegen/ gefallen/ keine), „Bekanntheit“ (alt/ neu) und „Argument“ (fallen/ steigen) über die Reaktionszeiten in Millisekunden auf die Argumente im Alt-Neu-Paradigma ergab keinen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,126)}=0.14$) noch der „Argumente“ ($F_{(1,63)}=0.02$). Es zeigte sich jedoch ein signifikanter Haupteffekt über die „Bekanntheit“ ($F_{(1,63)}=4.46$, Power (1-beta)=0.97): Auf „neue“ Argumente (922 ms) wurde schneller reagiert als auf alte Argumente (1043 ms). Keine Interaktion wurde signifikant (alle $F < .79$).

DISKUSSION

Die Verteilung der Daten zeigte kaum Unterschiede zwischen den experimentellen Bedingungen. Die Mittelwerte zwischen Schätzungen und Erinnerungen unterschieden sich wie in Experiment 1 nur geringfügig. Eine Verschiebung der Verteilung in Richtung der Rückmeldungen lag zwar vor, aber das Ausmaß betrug nur wenige Prozentpunkte. So nahm die Sicherheit, daß der Sachverhalt gestiegen wäre, mit der gleichlautenden Rückmeldung zu: Sie stieg aber von 55.86% Sicherheit dafür, daß der Sachverhalt gestiegen wäre, auf lediglich 56.45%. Bei der Rückmeldung „gefallen“ zeigte sich eine Abnahme der Sicherheit dafür, daß der Sachverhalt gestiegen sei, die ebenfalls konsistent zu der Rückmeldung erfolgte: Die Sicherheit dafür, daß der Sachverhalt gestiegen wäre, fiel von 52.01% bei der Schätzung auf nur noch 50.05% bei der Erinnerung. Allerdings zeigte sich das gleiche

Muster auch bei Sachverhalten ohne Rückmeldung. Hier fiel die Sicherheit dafür, daß der Sachverhalt gestiegen wäre, von 55.99% auf 52.34% bei der Erinnerung. Zusammen mit der Tatsache, daß die Streuung ebenfalls insgesamt ohne Unterschiede in dem variierten Faktor „Rückmeldung“ enger wurde, ergab sich auch in Experiment 2 das Bild einer Regression zur Mitte. Die Mittelwerte bei den Sachverhalten ohne Rückmeldung näherten sich allesamt klar der 50%-Marke und die Streuung wurde ebenfalls geringer. Die Konfidenzen blieben auch in Experiment 2 sehr nahe der 50%-Marke. Die zügigere Bearbeitung der Sachverhalte führte nicht zu extremeren Konfidenzen. Die Befunde ließen schon auf dieser Ebene der Analyse erkennen, daß ein deutlicher systematischer Einfluß der Rückmeldung nicht zu erwarten war und somit auch in diesem Experiment nur ein kleiner oder kein Rückschaufehler zu erwarten war.

Die Analyse des Anteils korrekter Erinnerungen zeigte in Experiment 2 deutliche Unterschiede zwischen Sachverhalten mit und ohne Rückmeldung. Mit 4% korrekten Erinnerungen an das eigene Konfidenzurteil aus der ersten Sitzung bei Sachverhalten ohne Rückmeldung verglichen mit 1.66% bei der Rückmeldung „gestiegen“ und 1% bei der Rückmeldung „gefallen“ lag der Anteil hier deutlich höher. Das Ergebnis deutet auf eine leichte Verschlechterung der Erinnerungsleistung bei Sachverhalten mit Rückmeldung hin. Dennoch waren auch hier wie im ersten Experiment die Anteile korrekter Erinnerungen sehr gering.

Wie schon im ersten Experiment standen die Ergebnisse der einfaktoriellen Varianzanalyse über den Faktor „Rückmeldung“ mit der abhängigen Variablen $\Delta\%$ im Einklang mit der Betrachtung der Verteilungen. Es zeigten sich keinerlei signifikanten Unterschiede im systematisch variierten Faktor „Rückmeldung“ und mithin zeigte sich auch kein Rückschaufehler. Damit bestätigte das zweite Experiment den Befund aus dem ersten Experiment. Im zweiten Experiment wurde durch die Veränderung der Instruktionen und der Evaluation der Argumente eine deutlich schnellere Bearbeitung der Aufgaben als in der ersten Sitzung erreicht. Im Schnitt dauerte die Beurteilung eines Sachverhalts und die Evaluation der Argumente nur noch 1.3 Minuten. Damit waren aber, wie schon bei der deskriptiven Analyse zu erkennen war, keine extremeren Konfidenzurteile und keine stärkere Wirkung der Rückmeldung verbunden. Die Befunde dieses zweiten Experiments legen nahe, daß auch hier wieder eine Entscheidung für einen Anstieg bzw. eine Abnahme des zu beurteilenden Sachverhalts den Probanden schwer fiel und daher die Mehrzahl der Entscheidungen in der Nähe 50% Konfidenz lagen. Eventuell war die Bearbeitungszeit für die einzelnen Sachverhalte immer noch zu lang und führte auch in diesem zweiten Experiment zu sehr ausgewogenen Urteilen. Die Auswertung der

Bedeutsamkeitseinschätzung der Argumente ergab auch im zweiten Experiment, daß über alle Sachverhalte die Argumente für eine Zunahme des Sachverhaltes (64.84%) wichtiger eingeschätzt wurden als Argumente, die für eine Abnahme des Sachverhaltes (60.42%) sprachen ($t_{49} = -3.94$). Erneut zeigte sich, daß alle Argumente genügend wichtig genommen wurden. Es gab kein einziges Argument, dessen Wichtigkeit unter 52% eingestuft wurde. Daher ist auch in diesem Experiment davon auszugehen, daß eine sehr ausgewogene Entscheidung getroffen wurde. Diese lag wieder sehr nahe der 50%-Marke und war daher leicht zu erinnern. Wie in Experiment 1 führte dies zu einer starken Regression zur Mitte. Es bleibt festzustellen, daß die Veränderungen des Versuchsablaufs nicht die gewünschte Wirkung gezeigt haben. Die Erklärung für das Ausbleiben des Effekts liegt daher eher in der Auswahl des Materials. Das Bemühen, möglichst neutrale Sachverhalte zu präsentieren, bei denen wenig präexperimentelle Präferenzen zu erwarten waren, führte daher, wie auch schon in der Diskussion von Experiment 1 angedeutet, zu indifferenten Urteilen. Die Wahl der Sachverhalte und deren Argumente für und wider eines Anstiegs ermöglichten keine klare Entscheidung. Die Wahl möglichst neutralen Materials in diesem zweiten und auch dem ersten Experiment mag zusätzlich dazu geführt haben, daß die Motivation, eine klare Entscheidung zu treffen, fehlte. Die Zu- oder Abnahme des öffentlichen Nahverkehrs einer fiktiven Stadt mag nicht das gleiche Interesse hervorgerufen haben, wie es beispielsweise die gleiche Frage bezogen auf die eigene Stadt getan hätte. So könnte auch ein allgemeines Desinteresse gegenüber dem verwendeten Material zu Konfidenzen in der Nähe der 50%-Marke geführt haben und damit für das Ausbleiben des Effekts verantwortlich gewesen sein.

Das Ausbleiben des Rückschaufehlers in diesem Experiment läßt somit wiederum keine systematischen Effekte in den anderen abhängigen Variablen erwarten, da das Hauptaugenmerk bei der Analyse der Wiedererkennensleistung und Reaktionszeiten natürlich auf einer systematischen Veränderung aufgrund der Rückmeldung lag.

Bei der Analyse der Wiedererkennensleistungen zeigte sich sowohl bei den Treffern als auch in der Diskriminationsleistung und im Antwort-Bias ein einheitliches Bild. Sachverhalte ohne Rückmeldung zeigten eine höhere Trefferrate für Argumente, die für „steigen“ sprachen, wobei auch die Diskriminationsleistung hier besser war und der Antwort-Bias eine liberalere Antwortstrategie andeutete. Diese Befunde sind konsistent zur Bedeutsamkeitseinschätzung der Argumente. Über alle Sachverhalte wurden Argumente, die für eine Zunahme des Sachverhaltes sprachen, wichtiger eingeschätzt als Argumente, die für eine Abnahme des Sachverhaltes sprachen.

Die Analyse der Reaktionszeiten zeigte lediglich einen deutlichen Vorteil neuer Argumente gegenüber alten. Dieser Befund entsprach ebenfalls nicht der Erwartung. Die Reaktion auf bekannte Argumente hätte in einem Wiedererkennenstest schneller erfolgen sollen als auf ein neues Argument. Möglicherweise zeigte sich hier jedoch ein Effekt der Präsentation der Argumente. Alle alten Stichworte zu den Argumenten waren mit bekannten Argumenten verknüpft, wohingegen die neuen Stichworte für die Probanden nicht mit ausformulierten Argumenten verknüpft waren. Dies könnte dazu geführt haben, daß bei alten Stichworten zunächst die dazugehörigen Argumente mit abgerufen wurden, wohingegen bei neuen Argumenten nichts abgerufen werden mußte, da diese nicht präsentiert worden waren. Die Folge wäre dann eine schnellere Antwort auf neue Stichworte.

Die Konsequenzen für das folgende dritte Experiment lagen auf der Hand. Offenbar sollten die Sachverhalte nicht zu neutral sein. Das Setting einer fiktiven Stadt sollte verlassen werden. Die Themen sollten so gewählt werden, daß sie auf ein gewisses Interesse bei den Probanden stoßen. Die Motivation, eine klare dezidierte Entscheidung zu treffen, sollte erhöht werden. Damit sollte auch vermieden werden können, daß leicht zu erinnernde Konfidenzen um die 50%-Marke getroffen werden würden.

5.3 Experiment 3: Wiedererkennen mit relevanten Sachverhalten

EINLEITUNG

Der Auswahl der Sachverhalte galt das besondere Augenmerk dieses Experimentes. Es wurde darauf geachtet, Material mit einer größeren Alltagsrelevanz zu verwenden. Außerdem wurde die Art der Rückmeldung variiert. Das grundsätzliche experimentelle Vorgehen wurde aber beibehalten. Auch in diesem dritten Experiment wurden die Image-Sets experimentell vermittelt und das Experiment folgte dem Gedächtnisdesign. Die Aufgabe bestand weiterhin darin, in 24 Verifikationsaufgaben zu beurteilen, ob die dort beschriebenen Sachverhalte in den letzten Jahren eher gestiegen oder gefallen sind, und sich zum Testzeitpunkt, genau eine Woche später, wieder an das eigene Urteil zu erinnern. Zu jedem Sachverhalt wurden wieder acht von insgesamt zwölf Argumenten dargeboten, wobei jeweils vier von sechs Argumenten dafür sprachen, daß der Sachverhalt gestiegen, und vier von sechs dafür, daß er gefallen war. Im Gegensatz zu den vorangegangenen Experimenten wurde in diesem Experiment besonderer Wert darauf gelegt, daß die verwendeten Sachverhalte einen Bezug zum Erfahrungsschatz der Probanden hatten. Es wurde besonders darauf geachtet, daß die beschriebenen Sachverhalte Themen ansprachen, die nach Möglichkeit ein gewisses Interesse bei den Probanden auslöste. Es sollte zu möglichst vielen Sachverhalten präexperimentelle Ansichten vorhanden sein. Ein Beispiel soll das verdeutlichen (kursiv sind jeweils zusammenfassende Stichworte zu den Argumenten angegeben):

1. 1988 glaubten 88% der amerikanischen Erwachsenen, das Recht zu haben, ihre Kinder zu schlagen. Dieser Prozentsatz hat sich bis zum jetzigen Zeitpunkt verändert.

Er hat abgenommen,

weil der autoritäre Erziehungsstil zunehmend kritisiert wird. Kritisierte autoritäre Erziehung

weil Werbekampagnen im Fernsehen über die Folgen von Gewalterziehung aufklären. Werbekampagnen gegen Gewalt

weil die zunehmende Berichterstattung in den Medien über Mißhandlungen in der Familie wirkt. Berichterstattung über Mißhandlungen

weil die Gewalt in der Familie zu einem wichtigen und vieldiskutierten Thema des Weltkindertages wurde. Weltkindertag

Er hat zugenommen,

weil die antiautoritäre Erziehung auch ihre Mängel zeigt. *Mängel antiautoritärer Erziehung*

weil autoritär ausgerichtete Glaubensgemeinschaften in den USA wieder mehr Zulauf bekommen. *Autoritäre Glaubensgemeinschaften*

weil strenge und konservative Colleges in den USA an Popularität gewinnen. *Strenge Colleges*

weil die Zusammenhänge zwischen Jugendgewalt und den fehlenden aber notwendigen Grenzen in der „Laissez-faire“-Erziehung immer häufiger kritisch diskutiert werden. *Jugendgewalt*

Die so gewählten Sachverhalte sollten zu einer größeren Motivation der Probanden führen und damit indifferente Angaben um die 50% in der Sicherheit für die Veränderung eines Sachverhaltes vermeiden. Die Probanden wurden wie in den beiden vorangegangenen Experimenten gebeten, durch eine Prozentangabe (0-100%) ihre Sicherheit anzugeben, ob der entsprechende Sachverhalt gestiegen oder gefallen wäre. Eine Prozentangabe von 0% sollte dabei „ganz sicher gefallen“ und eine Prozentangabe von 100% sollte entsprechend „ganz sicher gestiegen“ bedeuten. Eine Prozentangabe um 50% sollte hingegen anzeigen, daß keine Entscheidung getroffen werden konnte.

Zur zweiten Sitzung, die genau eine Woche nach der ersten stattfand, bekamen die Probanden erneut alle 24 Sachverhalte (allerdings ohne Argumente) präsentiert, wobei ihnen für jeweils acht der Sachverhalte eine bestätigende, synonyme, für weitere acht eine widersprechende, antonyme und für acht keine Rückmeldung gegeben wurde. Eine synonyme Rückmeldung bestand darin, das Urteil der Probanden aus dem ersten Fragebogen zu bestätigen, also rückzumelden, der Sachverhalt wäre „gestiegen“, wenn die Probanden im ersten Fragebogen selbst so geantwortet hatten, bzw. „gefallen“, falls das erste Urteil so lautete. Eine antonyme Rückmeldung behauptete das Gegenteil. Die Rückmeldung lautete also „gestiegen“, wenn die Probanden im ersten Fragebogen „gefallen“ geantwortet hatten, und „gefallen“, falls sie sich für „gestiegen“ entschieden hatten. Die acht Sachverhalte ohne Rückmeldung dienten als Kontrollitems. Die Aufgabe bestand darin, sich möglichst genau an das Urteil aus der ersten Sitzung zu erinnern.

Wie in den vorangegangenen Experimenten wurde zusätzlich das Wiedererkennen der Argumente geprüft. In einem Alt-Neu-Paradigma wurden die Reaktionszeiten für die acht alten gegenüber vier neuen Argumenten erhoben.

Die Erinnerung an die Konfidenzurteile sollte sich in den Experimentalbedingungen systematisch von der Kontrollbedingung unterscheiden: Die Rückmeldung, ein Sachverhalt wäre gestiegen, sollte die Sicherheit für „steigen“ erhöhen und für „fallen“ vermindern, während umgekehrt die Rückmeldung, daß ein Sachverhalt gefallen wäre, die Sicherheit für „fallen“ erhöhen und für „steigen“ mindern sollte. Ein solches Ergebnis entspräche einem typischen Rückschaufehler. Die Bedingungsvariation, daß ein Konfidenzurteil durch die Rückmeldung bestätigt oder widerlegt wurde, sollte im Falle der Bestätigung eine stärkere Sicherheit und im Falle der Widerlegung eine verminderte Sicherheit für eine Erinnerung im Sinne der Rückmeldung ergeben: Eine Entscheidung für „gestiegen“ im ersten Durchgang und eine synonyme Rückmeldung „gestiegen“ zum Testzeitpunkt sollte eine absolut gesehen stärkere Sicherheit im erinnerten Konfidenzurteil für „steigen“ ergeben als die antonyme Rückmeldung „gefallen“.

Im Alt-Neu-Paradigma sollte sich ebenfalls ein Zusammenhang mit der Rückmeldung zeigen lassen. Argumente, die für „fallen“ sprechen, sollten häufiger und schneller erkannt werden als solche, die für „steigen“ sprechen, wenn die Rückmeldung „fallen“ lautet und entsprechend umgekehrt. Dieses Ergebnis spräche für eine selektive Aktivierung des Image-Sets. Die Reaktionszeiten sollten mit synonyme Rückmeldung einen stärkeren Unterschied zeigen als mit antonymer Rückmeldung.

METHODE

Design Der Versuchsplan bestand in einem 3x2x2 Design mit den Faktoren „Rückmeldung“ (synonym/antonym/keine), „Schätzung“ (gestiegen/gefallen) und „Argument“ (fallen/steigen) (siehe Abb. 5.14). Alle Faktoren wurden als Meßwiederholungsfaktoren realisiert. Die Zuordnung zu dem Faktor „Rückmeldung“ wurde durch systematische Permutation einer einmaligen Zufallsreihenfolge realisiert, so daß alle Fragen über alle Bedingungen gleich häufig verwirklicht wurden. Eine „synonyme“ Rückmeldung bestand darin, das Urteil der Probanden aus dem ersten Fragebogen zu bestätigen, also rückzumelden, der Sachverhalt wäre „gestiegen“, wenn die Probanden im ersten Fragebogen selbst so geantwortet hatten, bzw. „gefallen“, falls das erste Urteil so lautete. Eine „antonyme“ Rückmeldung behauptete das Gegenteil. Die Rückmeldung lautete also „gestiegen“, wenn die Probanden im ersten Fragebogen „gefallen“ geantwortet hatten, und „gefallen“, falls sie sich für „gestiegen“ entschieden hatten. Die acht Sachverhalte ohne Rückmeldung dienten als Kontrollitems. Der Faktor „Argument“ gruppiert die Wiedererkennensleistung der Argumente in Argumente, die dafür sprachen, daß der entsprechende Sachverhalt gestiegen („steigen“) und Argumente, die dafür

sprachen, daß er gefallen („fallen“) war. Die Argumente sprachen jeweils zur Hälfte dafür, daß der entsprechende Sachverhalt gestiegen respektive gefallen war. Die Darbietung erfolgte blockweise, wobei die Reihenfolge der Argumente als auch der Blöcke selbst einer Zufallsreihenfolge folgte, so daß alle Argumente gleich häufig an jeder Position präsentiert wurden.

Als abhängige Maße wurden neben der Differenz der Konfidenzurteile zum Zeitpunkt der ersten Einschätzung und dem Zeitpunkt der Erinnerung ($\Delta\%$) die Antworten und Reaktionszeiten in einer Alt-Neu-Entscheidung erhoben.

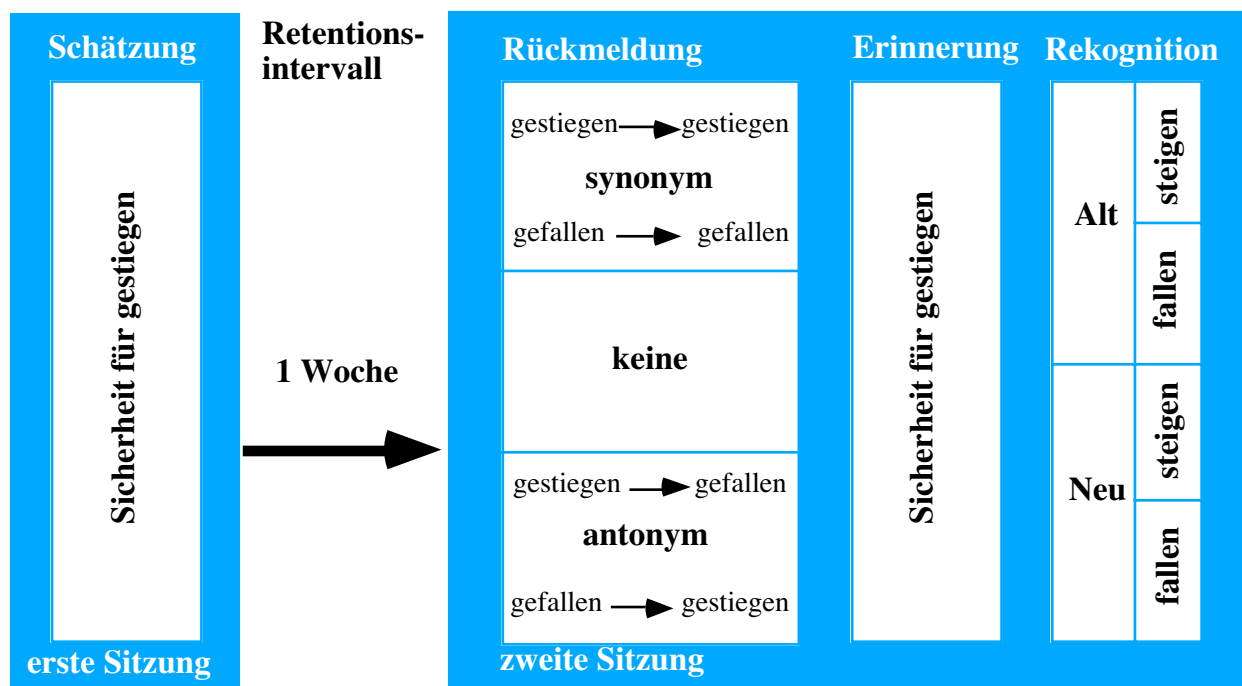


Abbildung 5.14: Versuchsplan von Experiment 3. Gedächtnisdesign mit einer Woche Retentionsintervall und Meßwiederholung auf den Faktoren „Rückmeldung“ und „Argumente“. „Synonym“ bedeutete eine gleichlautende Rückmeldung der Entscheidung der Probanden, also „gefallen“, falls diese sich in der ersten Sitzung für „gefallen“ entschieden hatten. „Antonym“ hingegen bedeutete ein Widerspruch zur ersten Sitzung: hatten die Probanden sich dort für „gestiegen“ entschieden, wurde jetzt „gefallen“ rückgemeldet. Entsprechendes gilt für die anderen Entscheidungen.

Probanden An der Untersuchung nahmen 64 Personen, vorwiegend Studierende aus verschiedenen Fachbereichen der Universität Trier, teil. Davon waren 39 weiblichen und 25 männlichen Geschlechts. Das Altersspektrum reichte von 19 bis 34 Jahren. Der Altersdurchschnitt lag bei 23.6 Jahren. Die Probanden erhielten 12 DM für die Teilnahme am Experiment.

Material Wie schon in den beiden vorangegangenen Experimenten wurden auch in diesem dritten Experiment 24 Verifikationsaufgaben vorgegeben, die danach beurteilt werden sollten, wie sicher sich die Probanden jeweils waren, ob ein betreffender Sachverhalt gestiegen oder gefallen war. Die Beurteilung der Sachverhalte erfolgte jeweils durch Verschiebung einer Markierung auf einer analogen Skala auf dem Computerbildschirm, die von „gefallen“ zu „gestiegen“ reichte. Die Probanden bekamen keine Rückmeldung des konkreten numerischen Skalenwertes. Im Gegensatz zu den beiden vorausgegangenen Experimenten wurde bei diesem Experiment jedoch Wert darauf gelegt, daß die Themen gerade so gewählt waren, daß sie das allgemeine Interesse der Probanden ansprachen. Ziel war es, indifferente Entscheidungen nahe der 50%-Marke zu vermeiden und damit die Voraussetzungen für das Auftreten eines Rückschaufehlers zu schaffen.

Wie auch schon in den vorangegangenen Experimenten wurden die Argumente auf dem Bildschirm in Stichworten zusammengefaßt. Zunächst wurde den Probanden jedes Argument vollständig ausformuliert zusammen mit einem Stichwort dargeboten. Die Probanden sollten sich das Argument durchlesen und anschließend die Maustaste betätigen, damit das nächste Argument auf dem Bildschirm erscheinen konnte. Nachdem alle Argumente präsentiert worden waren, blieben auf dem Bildschirm, in der Reihenfolge ihrer Darbietung, nur die Stichwörter zu den Argumenten sichtbar. Die Probanden konnten aber jederzeit durch Klicken auf ein Stichwort das entsprechende Argument nochmals lesen. Auf die Einschätzung der Wichtigkeit der Argumente wurde in diesem Experiment verzichtet, um die Aufmerksamkeit der Probanden nicht zu stark an die Argumente zu binden. Die Argumente sollten nicht zu stark enkodiert werden zum Einen, um präexperimentellen Präferenzen nicht durch zu lange Beschäftigung mit Gegenargumenten zu verwässern und eine spontane dezidierte Entscheidung bei den Sachverhalten zu ermöglichen, und zum Anderen, um eine spätere systematische Beeinflussung durch die Rückmeldung erst möglich zu machen (siehe Abb. 5.14).

Um im zweiten Durchgang einen Wiedererkennenstest realisieren zu können, existierten zu jedem Sachverhalt noch vier nicht präsentierte Stichwörter, die im letzten Teil des Experiments als Foils dienten.

Ablauf Analog zu den beiden vorangegangenen Experimenten folgte auch dieses Experiment dem Gedächtnisdesign und fand in zwei Sitzungen mit einer Woche Abstand statt. Der gesamte Versuch erfolgte am Rechner. Es bestand jeweils für vier Probanden gleichzeitig die Möglichkeit teilzunehmen. Der Ablauf deckte sich in

sämtlichen Details mit den vorangegangenen Experimenten, so daß auf eine ausführliche Beschreibung an dieser Stelle verzichtet wird. Die erste Sitzung dauerte im Durchschnitt 30 Minuten pro Proband, während die zweite Sitzung im Durchschnitt 25 Minuten pro Proband dauerte.

Auswertung Um Extremwerte aus der Analyse auszuschließen, wurden wieder alle Reaktionszeiten unter 200 msec aus der Auswertung ausgeschlossen. Werte oberhalb der dreifachen Standardabweichung wurden ebenfalls von der Auswertung ausgeschlossen. Insgesamt wurden dadurch 659 von ursprünglich 13824 Werten ausgeschlossen. Das entspricht 4.8% der Daten.

Die Verteilung der Schätzungen für „gefallen“ respektive „gestiegen“ zeigten eine Bevorzugung des Ausgangs „gestiegen“. Die Probanden entschieden sich in 59.16% für „gestiegen“. Bei „synonym“ rückgemeldeten Sachverhalten wurde in 58.18%, bei „antonym“ rückgemeldeten Sachverhalten in 63.35% und bei Sachverhalten ohne Rückmeldung in 55.98% „gestiegen“ gewählt. Die Datenmenge für die Mittelwerte des Faktors „Schätzung“ sind daher noch ausreichend groß und vergleichbar.

ERGEBNISSE

Verteilung der Daten Betrachtet man zunächst die Verteilung von Schätzungen und Erinnerungen über die Sicherheit in Prozent, daß der Sachverhalt zugenommen habe, getrennt nach der Rückmeldung (synonym/antonym/keine), sieht man, daß in der Bedingung mit „synonymer“ Rückmeldung (siehe Abb. 5.15) sowohl für die Rückmeldung „gestiegen“ als auch „gefallen“ die Mittelwerte der Verteilungen von Schätzungen und Erinnerungen in Richtung der Skalenmitte wanderten (Schätzung „gestiegen“: 71.82% zu 65.85%; Schätzung „gefallen“: 33.06% zu 40.67%). Die Streuung nahm in der Verteilung der Erinnerungen zu (Schätzung „gestiegen“: 14.40 zu 15.10; Schätzung „gefallen“: 11.85 zu 15.40).

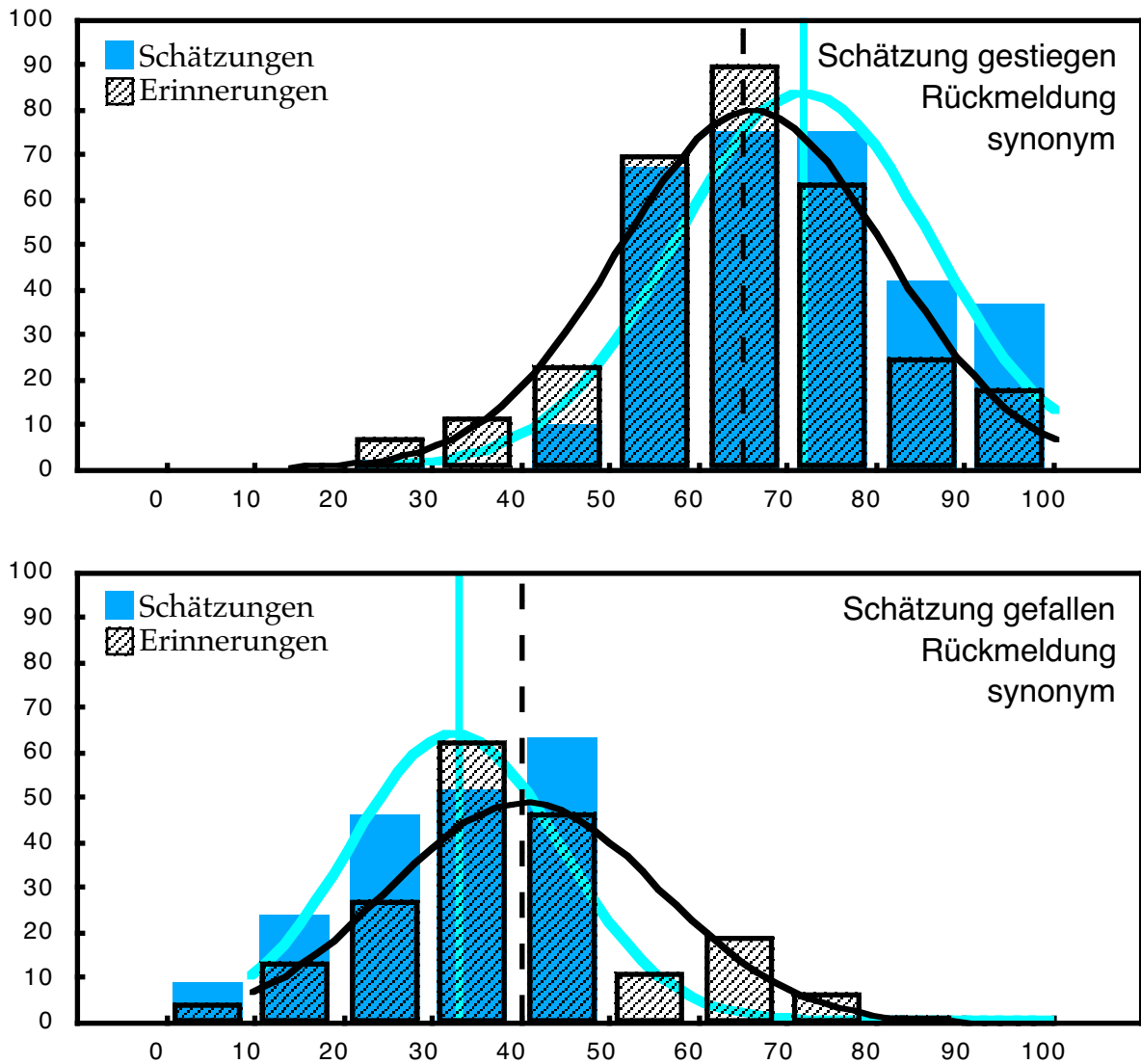


Abbildung 5.15: Relative Häufigkeiten für Schätzungen und Erinnerungen pro Zehner-Kategorie gegliedert nach der Rückmeldung bei synonyme Rückmeldung.

In der Bedingung mit „antonymer“ Rückmeldung (siehe Abb. 5.16) wanderten ebenfalls sowohl für die Rückmeldung „gestiegen“ als auch „gefallen“ die Mittelwerte der Verteilungen von Schätzungen und Erinnerungen in Richtung der Skalenmitte (Schätzung „gestiegen“: 71.93% zu 58.95%; Schätzung „gefallen“: 34.19% zu 46.34%). Die Streuung nahm in der Verteilung der Erinnerungen zu (Schätzung „gestiegen“: 13.50% zu 16.01%; Schätzung „gefallen“: 12.91% zu 14.60%).

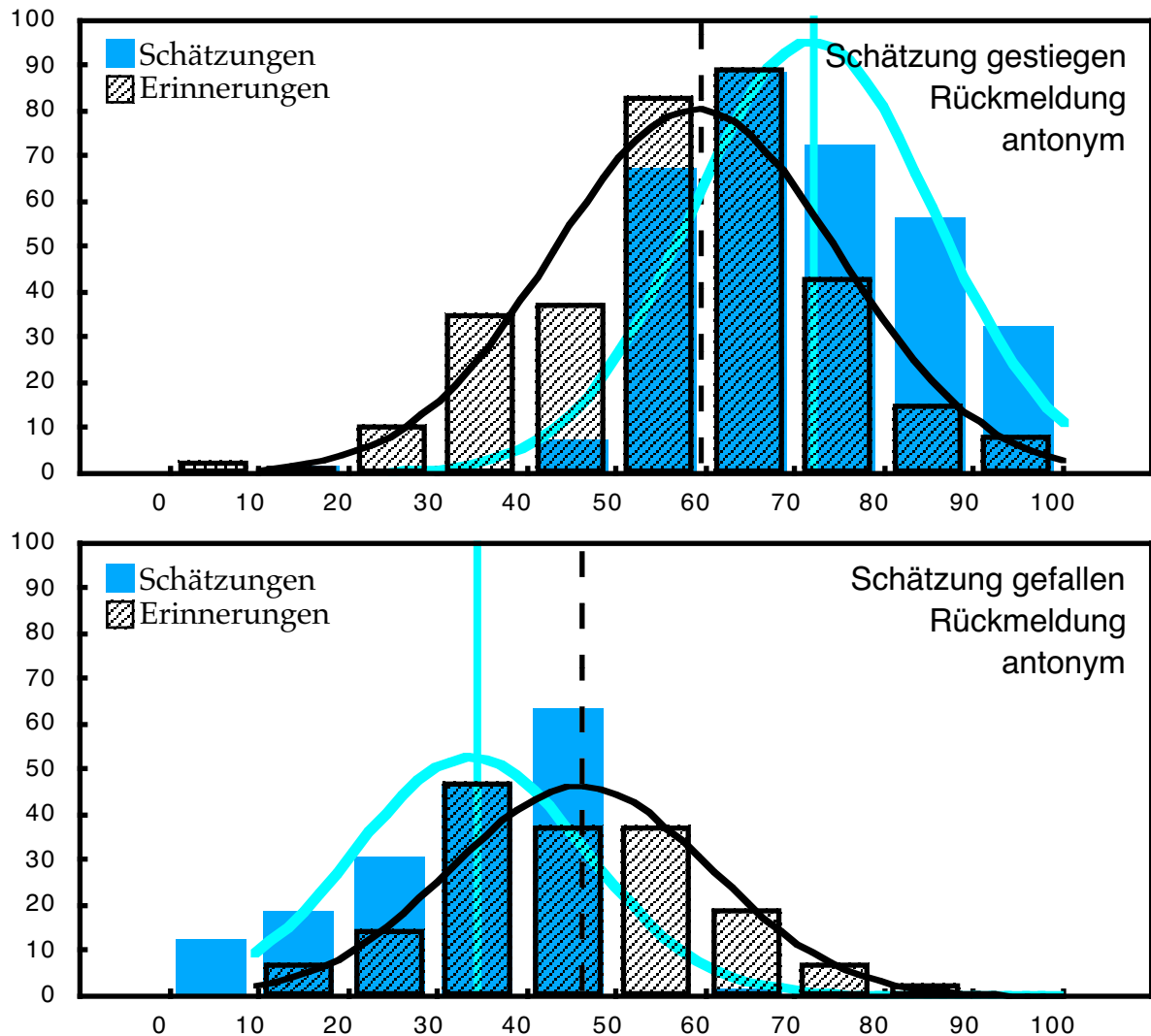


Abbildung 5.16: Relative Häufigkeiten für Schätzungen und Erinnerungen pro Zehner-Kategorie gegliedert nach der Rückmeldung bei antonymer Rückmeldung.

Teilt man in der Bedingung „keine“ Rückmeldung (siehe Abb. 5.17) die Schätzungen und Erinnerungen nach der Lage der Schätzungen in solche die „steigen“ und solche, die „fallen“ geschätzt haben, ein, wanderten ebenfalls sowohl für die Schätzung „gestiegen“ als auch „gefallen“ die Mittelwerte der Verteilungen von Schätzungen und Erinnerungen in Richtung der Skalenmitte (Schätzung „gestiegen“: 70.44% zu 64.54%; Schätzung „gefallen“: 30.67% zu 40.08%). Die Streuung nahm in der Verteilung der Erinnerungen zu (Schätzung „gestiegen“: 13.62 zu 15.75; Schätzung „gefallen“: 13.94 zu 17.29).

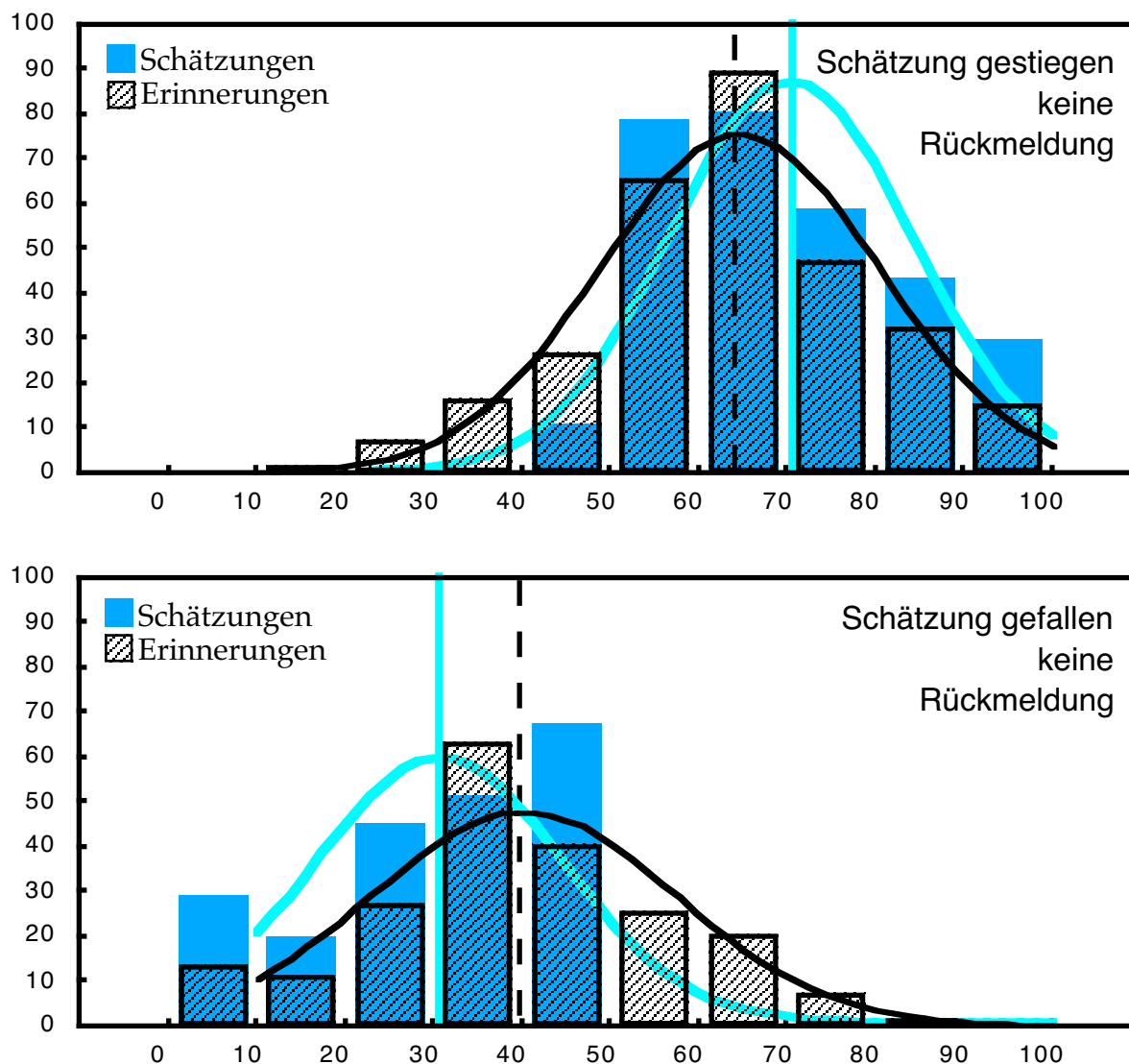


Abbildung 5.17: Relative Häufigkeiten für Schätzungen und Erinnerungen pro Zehner-Kategorie gegliedert nach der Schätzung in der Kontrollbedingung.

% **korrekt** Die Analyse der Anzahl korrekter Erinnerungen in der Kontroll- und den beiden Experimentalbedingungen erlaubt eine Beurteilung der Gedächtnisbeeinträchtigung durch die experimentelle Manipulation. Die Anzahl korrekter Erinnerungen ergab keine signifikanten Unterschiede zwischen synonymen (6.94%), antonymen (5.48%) und keiner Rückmeldung (7.13%) ($\chi^2_{(2)}=1.33$). Auch die Trennung nach der Präferenz der Schätzung in „gestiegen“ (6.77%) versus „gefallen“ (6.26%) zeigte keinerlei signifikante Unterschiede ($\chi^2_{(1)}=.15$). Weder innerhalb synonymen ($\chi^2_{(1)}=1.29$), antonymen ($\chi^2_{(1)}=.08$) noch ohne Rückmeldung ($\chi^2_{(1)}=1.30$) zeigte sich ein Unterschied zwischen der Präferenz in der Schätzung (siehe Tabelle 5.1). Die Rückmeldung unterschied sich in der Anzahl korrekter Erinnerung weder, wenn die Entscheidung für „gestiegen“ ($\chi^2_{(2)}=1.16$), noch wenn sie für „gefallen“ ($\chi^2_{(2)}=2.94$) ausgefallen war (siehe Tabelle 5.1).

Tabelle 5.1: % korrekt getrennt nach Schätzung (gestiegen/ gefallen) und Rückmeldung (synonym/ antonym/ keine).

| Schätzung | Rückmeldung | Anzahl | Prozent |
|-----------|-------------|--------|---------|
| gestiegen | synonym | 24 | 7.97 |
| gefallen | synonym | 10 | 5.29 |
| gestiegen | antonym | 10 | 5.88 |
| gefallen | antonym | 17 | 5.26 |
| gestiegen | keine | 18 | 6.04 |
| gefallen | keine | 18 | 8.70 |

Konfidenzurteile Bei einer zweifaktoriellen Varianzanalyse mit den Faktoren „Rückmeldung“ (synonym/antonym/keine) und „Schätzung“ (gestiegen/gefallen) über die Differenz $\Delta\%$ zeigte sich kein signifikanter Haupteffekt des Faktors „Rückmeldung“ ($F_{(2,104)}=2.64$). Der Faktor „Schätzung“ ($F_{(1,52)}=188.90$, Power (1-beta)=0.85) wurde jedoch signifikant: Es zeigten sich bei der Schätzung „gestiegen“ ein $\Delta\%$ von -7.84 und bei der Schätzung „gefallen“ ein $\Delta\%$ von 9.41. Die Interaktion (siehe Abb. 5.18) zwischen „Rückmeldung“ und „Schätzung“ wurde ebenfalls signifikant ($F_{(2,104)}=6.12$, Power (1-beta)=0.79): Die Schätzung „gestiegen“ zeigte Unterschiede in den Rückmeldungen ($F_{(2/104)}=13.69$; $p=.000004$). Ein Tukey Test zeigte, daß sich jeweils nur die antonyme Rückmeldung (-12.38 $\Delta\%$) von den anderen unterschied (synonym/antonym $p=.00002$; kontroll/antonym ebenfalls $p=.00002$). Die Schätzung „gefallen“ zeigte keine signifikanten Unterschiede in den Rückmeldungen ($F_{(2/108)}=1.51$). Der Unterschied im Wert von $\Delta\%$ für die Schätzung „gestiegen“ (-5.21 $\Delta\%$) zu „gefallen“ (7.40 $\Delta\%$) mit der Rückmeldung synonym wurde signifikant ($t_{(55)}=5.40$; $p=.0000015$). Auch mit der Rückmeldung antonym unterschied sich die Schätzung „gefallen“ (11.04 $\Delta\%$) von „gestiegen“ (-12.38 $\Delta\%$) ($t_{(57)}=9.93$; $p=.0000000$). In der Bedingung „keine“ Rückmeldung unterschieden sich Werte von $\Delta\%$ für Schätzungen, die für „gestiegen“ (5.21 $\Delta\%$) respektive „gefallen“ (-9.79 $\Delta\%$) ausgefallen waren, gleichfalls ($t_{(57)}=9.62$; $p=.0000000$).

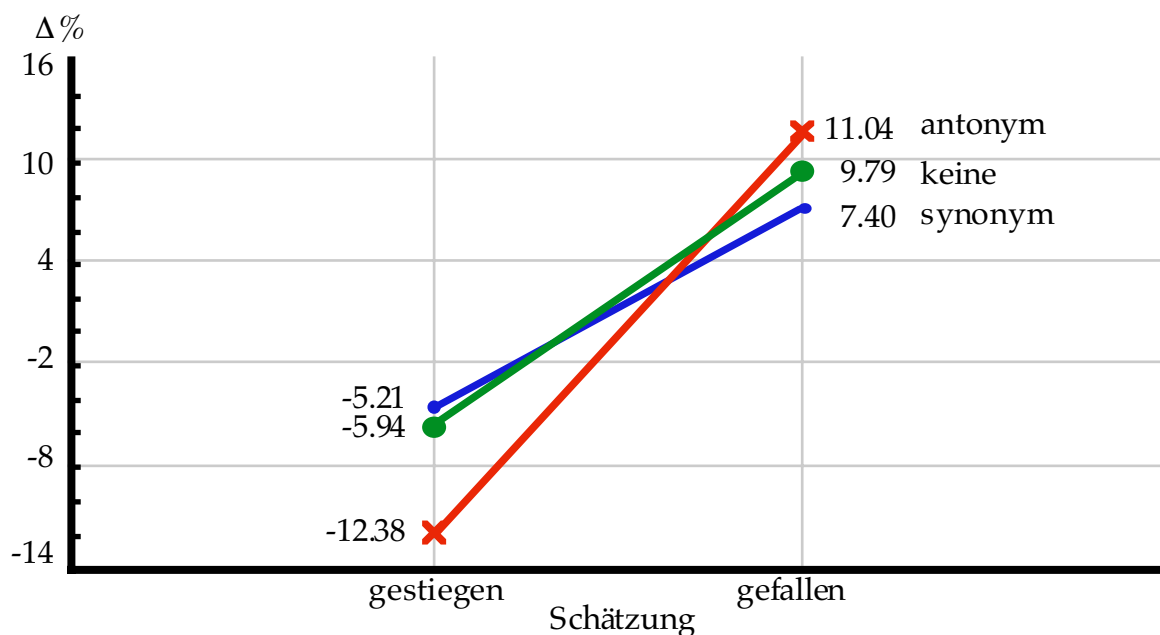


Abbildung 5.18: Interaktion zwischen „Rückmeldung“ und „Schätzung“ über die Differenz im Konfidenzurteil. Eine „antonyme“ Rückmeldung erzeugt eine Annäherung der Erinnerung an die Rückmeldung (positiver Wert für die Schätzung „gefallen“ und negativer Wert für die Schätzung „gestiegen“), während eine synonyme Rückmeldung eine Abweichung von der Rückmeldung erzeugt.

Wiedererkennensleistung Für die Beurteilung der Güte der Entscheidungen im Alt-Neu-Paradigma wurden wiederum Hits, False Alarms, die Sensitivität und der Bias erhoben.

Hits Eine dreifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (synonym/antonym/keine), „Schätzung“ (gestiegen/gefallen) und „Argument“ (fallen/steigen) über die Treffer zeigte keinen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,120)}=2.38$). Der Faktor „Schätzung“ zeigte signifikante Unterschiede in den Treffern ($F_{(1,60)}=6.98$; Power (1-beta)=0.85): Es wurden mehr Treffer bei der Schätzung „gestiegen“ (.72) als bei der Schätzung „gefallen“ (.66) erzielt. Signifikant wurde ebenfalls der Unterschied in den Trefferhäufigkeiten beim Faktor „Argument“ ($F_{(1,60)}=8.19$; Power (1-beta)=0.86): Argumente, die für „fallen“ sprechen (.71), werden häufiger richtig erkannt als solche, die für „steigen“ sprechen (.67). Die zweifachen Interaktionen wurden nicht signifikant (alle $F < 1$).

Die Interaktion zwischen der „Rückmeldung“, „Schätzung“ und „Argument“ (siehe Abb. 5.19) wurde signifikant ($F_{(2,120)}=3.74$; Power (1-beta)=0.79): Der Faktor „Rückmeldung“ zeigte signifikante Unterschiede in der Anzahl Treffer bei Argumenten, die für „steigen“ sprachen und wo die Schätzung „gefallen“ gelautet hatte ($F_{(2,120)}=4.06$). Ein Tukey Test zeigte, daß sich lediglich die „synonyme“

Rückmeldung von „keiner“ Rückmeldung unterschied (synonym/antonym $p=.14$, synonym/keine $p=.014$ und antonym/keine $p=.63$). Alle weiteren Einzelvergleiche zeigten jedoch keine Signifikanz (alle $F < 3.48$; $T < 2.95$).

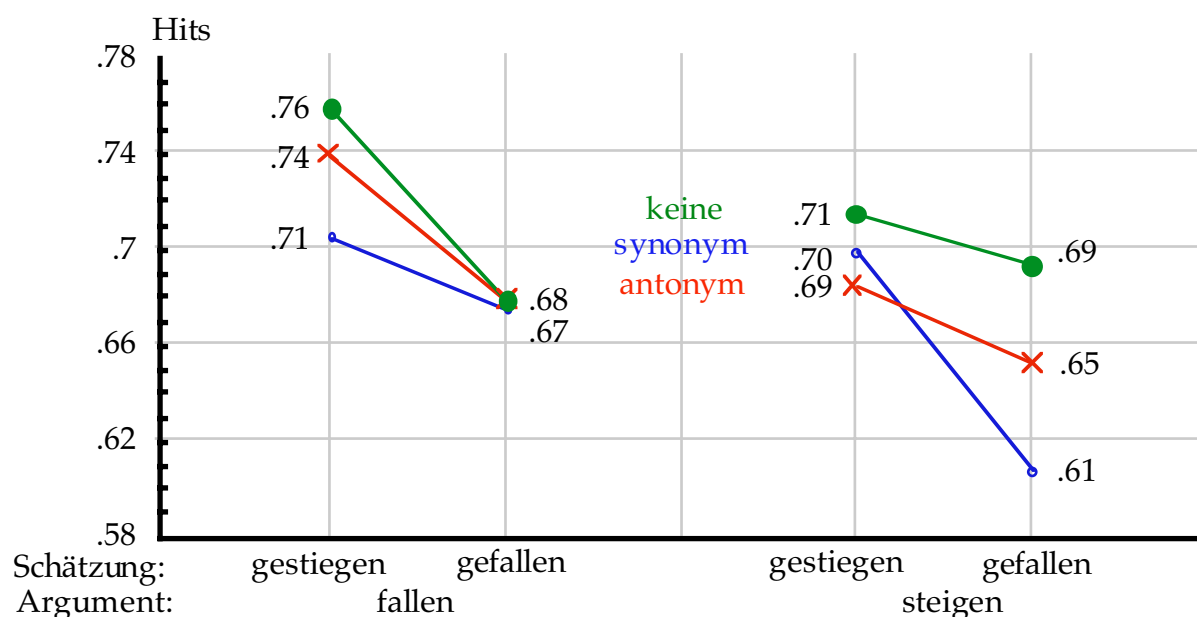


Abbildung 5.19: Interaktion Rückmeldung (synonym/antonym/keine), Schätzung (gestiegen/gefallen) und Argument (fallen/steigen) über die Treffer.

False Alarms Eine zweifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (synonym/antonym/keine), „Schätzung“ (gestiegen/gefallen) über die false Alarms zeigte weder einen signifikanten Effekt der „Rückmeldung“ ($F_{(1,61)}=0.08$) noch der „Schätzung“ ($F_{(2,122)}=0.66$). Die Interaktion wurde ebenfalls nicht signifikant ($F_{(2,122)}=0.71$).

Diskriminationsleistung Eine dreifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (synonym/antonym/keine), „Schätzung“ (gestiegen/gefallen) und „Argument“ (fallen/steigen) über d' zeigte keinen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,122)}=2.07$). Der Faktor „Schätzung“ zeigte signifikante Unterschiede in der Diskriminationsleistung ($F_{(1,61)}=7.41$, Power (1-beta)=0.86): Lautete die Schätzung „gestiegen“ wurde sensibler reagiert (1.09), als wenn die Schätzung „gefallen“ gelautet hatte (.98). Die „Argumente“ unterschieden sich ebenfalls signifikant ($F_{(1,49)}=6.19$; Power (1-beta)=0.86): Auf Argumente für „steigen“ (1.07) wurde sensibler reagiert als auf Argumente für „fallen“ (1.00). Die zweifach Interaktionen wurden nicht signifikant (alle $F < 1$).

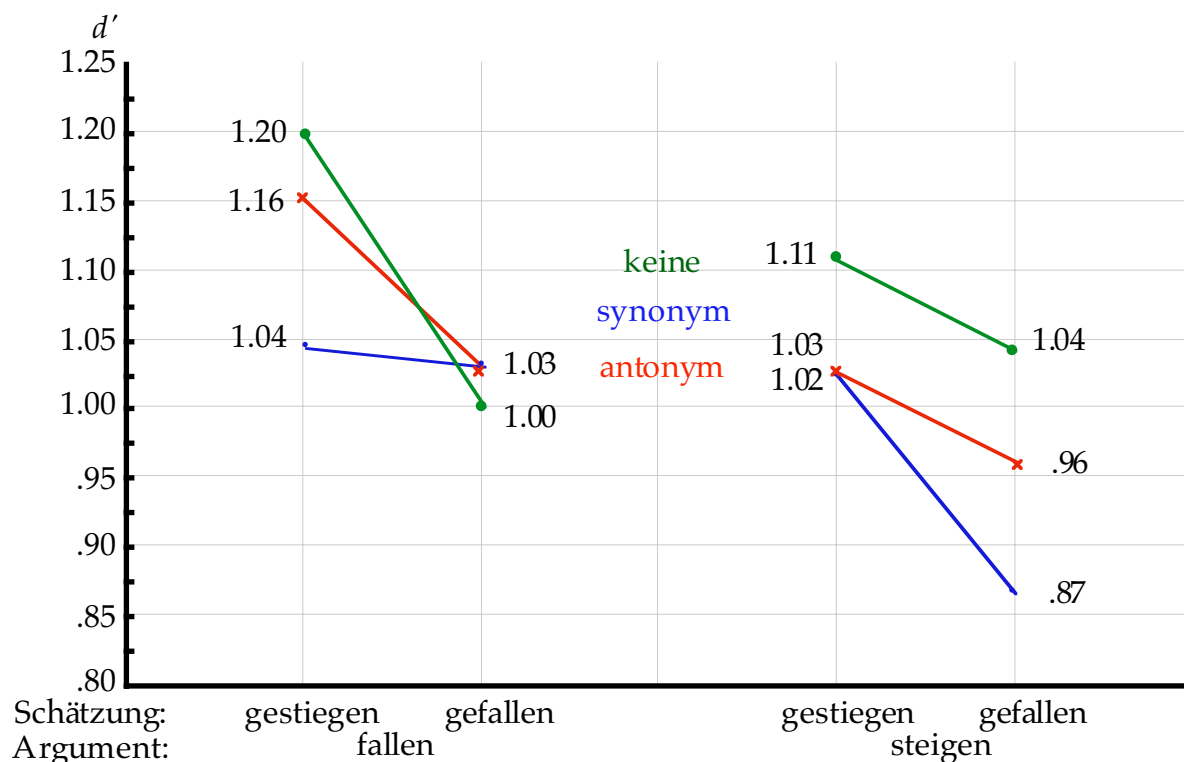


Abbildung 5.20: Interaktion Rückmeldung (synonym/antonym/keine), Schätzung (gestiegen/gefallen) und Argument (fallen/steigen) über die Diskriminationsleistung d' .

Die Interaktion zwischen der „Rückmeldung“, „Schätzung“ und „Argument“ (siehe Abb. 5.20) wurde signifikant ($F_{(2,122)}=3.44$; Power (1-beta)=0.79): Wurde „gestiegen“ geschätzt, fanden sich signifikante Unterschiede über die „Rückmeldung“ ($F_{(2,122)}=4.06$; $p=0.02$). Ein Tukey-Test zeigte, daß sich lediglich „keine“ Rückmeldung von „synonymer“ Rückmeldung unterschied (synonym/keine $p=.016$). Es zeigten sich jedoch keine signifikanten Unterschiede in den weiteren Einzelvergleichen (alle $F<3.64$; $t<2.94$).

Antwort-Bias Eine dreifaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (synonym/antonym/keine), „Schätzung“ (gestiegen/gefallen) und „Argument“ (fallen/steigen) über C zeigte keinen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,122)}=1.32$). Der Faktor „Schätzung“ zeigte ebenfalls keine signifikanten Unterschiede in dem Antwort-Bias ($F_{(1,61)}=.03$). Die „Argumente“ unterschieden sich jedoch signifikant ($F_{(1,61)}=6.33$; Power (1-beta)=.86): Argumente für „steigen“ (.04) wurden konservativer beantwortet als solche für „fallen“ (.01). Die zweifachen Interaktionen wurden nicht signifikant (alle $F<1$). Die Interaktion zwischen der „Rückmeldung“, „Schätzung“ und „Argument“ (siehe Abb. 5.21) wurde signifikant ($F_{(2,122)}=3.47$; Power (1-beta)=0.79): Wurde „gefallen“ geschätzt fanden sich signifikante Unterschiede über die „Rückmeldung“ ($F_{(2,122)}=3.04$; $p=0.05$). Ein Tukey-Test zeigte, daß sich lediglich „keine“ Rückmeldung von „synonymer“

Rückmeldung unterschied (synonym/keine $p=.05$). Es zeigten sich jedoch keine weiteren signifikanten Unterschiede in den Einzelvergleichen (alle $F<2.98$; $T<2.96$).

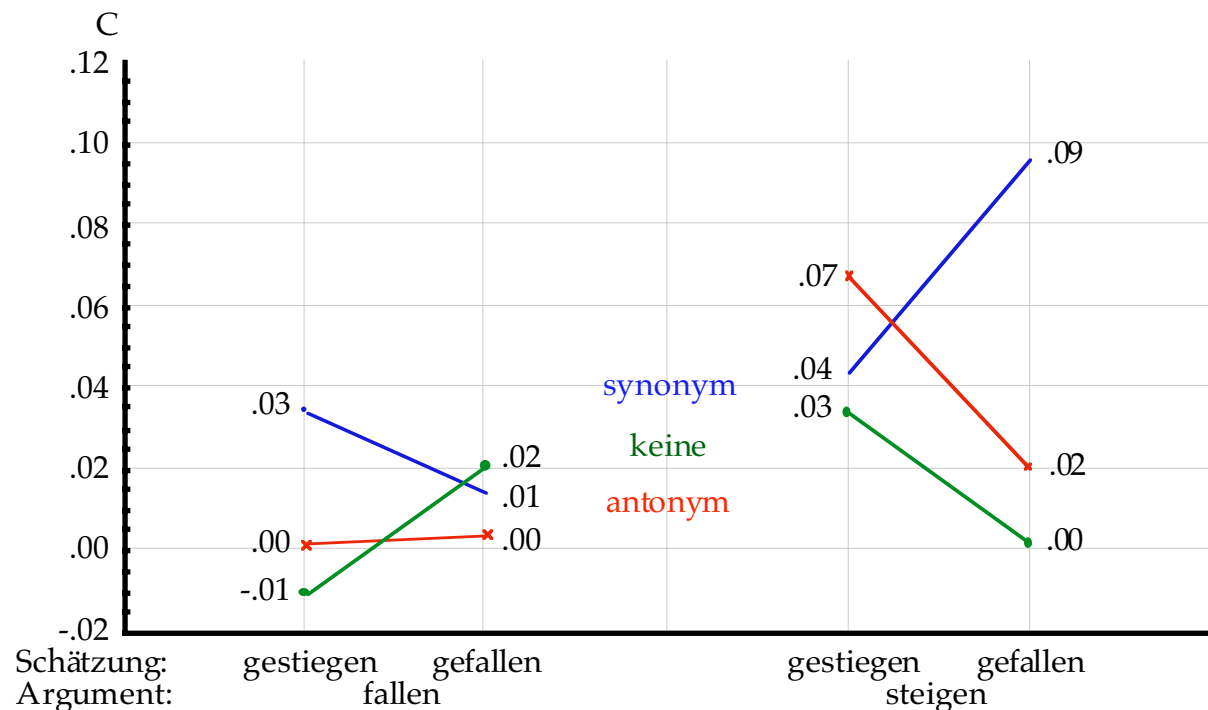


Abbildung 5.21: Interaktion Rückmeldung (synonym/antonym/keine), Schätzung (gestiegen/gefallen) und Argument (fallen/steigen) über den Antwort-Bias C.

Reaktionszeiten Eine vierfaktorielle Varianzanalyse mit den Faktoren „Rückmeldung“ (synonym/antonym/keine), „Schätzung“ (gestiegen/gefallen), „Bekanntheit“ (alt/neu) und „Argument“ (fallen/steigen/neue) über die Reaktionszeiten auf die Argumente im Alt-Neu-Paradigma ergab weder einen signifikanten Haupteffekt der „Rückmeldung“ ($F_{(2,104)}=.38$) der „Schätzung“ ($F_{(1,52)}=1.26$), der „Bekanntheit“ ($F_{(1,52)}=.44$) noch des Faktors „Argument“ ($F_{(2,104)}=.01$). Keinerlei Interaktionen wurden signifikant (alle $F<2.59$).

DISKUSSION

Die Verteilung der Daten zeigte schon in der Bedingung mit synonyme Rückmeldung eine Erinnerung, die nicht in Richtung der vorgegebenen Rückmeldung, sondern in Richtung der 50%-Marke verschoben war. In der Bedingung mit antonymer Rückmeldung zeigte sich das gleiche Bild, wobei hier die Rückmeldung auch eine Verschiebung in die Richtung erwarten ließen. Nimmt man jetzt auch noch die Sachverhalte ohne Rückmeldung hinzu und unterteilt die Schätzungen in solche, die für „gefallen“ und solche, die für „gestiegen“ sprachen, rundet sich das Bild ab. Auch hier zeigte sich eine starke Verschiebung zur 50%-Marke. Daher zeigte sich in Experiment 3 ganz eindeutig eine starke Tendenz zur

Mitte. Die Streuung nahm in allen Fällen zu. Der Grund für die Zunahme der Streuung lag in der Einteilung der Daten durch den Faktor „Rückmeldung“ in Sachverhalte, die bei der Schätzung durch die Probanden „gestiegen“ respektive „gefallen“ eingestuft wurden. Die Schätzungen der Probanden dienten der gezielten „synonymen“ respektive „antonymen“ Rückmeldung der Sachverhalte. Das Kriterium für die erfolgte Rückmeldung lag bei der 50%-Marke. Lag eine Schätzung unter diesem Wert wurde im Falle „synonymer“ Rückmeldung zurückgemeldet, der Sachverhalt sei in Wirklichkeit „gefallen“, während bei „antonymer“ Rückmeldung, rückgemeldet wurde, er sei „gestiegen“. Dadurch wurden die Verteilungen der Schätzung systematisch an der 50%-Marke gekürzt. Es ergibt sich also bei der Betrachtung der Daten ein unvollständiges Bild. Die Streuung muß bei dieser Darstellung der Daten zunehmen, da die wahre Streuung der Schätzung nicht berücksichtigt wird. Auch wenn Probanden im Mittel eher die Einschätzung bevorzugten, der Sachverhalt sei „gestiegen“ lagen doch eine erhebliche Menge an Daten unter der 50%-Marke. Betrachtet man die vollständigen Verteilungen so nahmen die Streuungen auch ab: bei „synonymer“ Rückmeldung von 23.19 auf 19.53, bei „antonymer“ Rückmeldung von 22.34 auf 16.64 und ohne Rückmeldung von 23.92 auf 20.94.

Insgesamt zeigte sich also bei der deskriptiven Betrachtung der Daten ein uneinheitliches Bild. Während in der Bedingung mit antonymer Rückmeldung eine Verschiebung der Konfidenzen in Richtung der Rückmeldung auftrat, zeigte die synonyme Rückmeldung keinerlei Wirkung. Es zeigte sich im Gegenteil eine Verschiebung der Erinnerung zur 50%-Marke, die auch bei Sachverhalten ohne Rückmeldung zu erkennen ist. Analysiert man die Daten traditionell nur nach der erfolgten Rückmeldung, deutete sich eine Verschiebung in Richtung der Rückmeldung nur bei der Rückmeldung „gefallen“ an. Schon die Schätzungen deuteten a priori auf eine Präferenz für die Rückmeldung „gestiegen“. Die Mittelwerte für die Schätzungen liegen alle über der 50%-Marke (56.64% über alle Schätzungen). Betrachtet man daher die Rückmeldung „gestiegen“ als eine Bestätigung und „gefallen“ als einen subjektiv erlebten Widerspruch, zeigte auch die traditionelle Rückmeldung ein konsistentes Ergebnis. Es deutet sich ein Rückschaufehler nur im Falle antonymer, widersprechender Rückmeldung an.

Die Befunde ließen schon auf dieser Ebene der Analyse erkennen, daß ein deutlicher systematischer Einfluß der Rückmeldung nicht zu erwarten sei und somit auch in diesem Experiment nur ein kleiner oder kein Rückschaufehler zu erwarten war.

Die Analyse des Anteils korrekter Erinnerungen zeigte in Experiment 3 keine Unterschiede zwischen Sachverhalten mit synonymem, antonymem und ohne Rückmeldung. Dennoch waren auch hier, wie in den vorangegangenen Experimenten, die Anteile korrekter Erinnerungen sehr gering.

Die Analyse der Differenz $\Delta\%$ erfordert eine genaue Betrachtung des Faktors „Rückmeldung“. So ist zu beachten, daß bei einer „antonymen“ Rückmeldung und einer Schätzung, die „gestiegen“ gelautet hatte, die tatsächlich erfolgte Rückmeldung „gefallen“ gelautet hatte. Entsprechend sind die Ergebnisse der Varianzanalyse zu werten. Der signifikant gewordene Haupteffekt der „Schätzung“ deutete auf eine allgemeine Distanzierung der Erinnerungen von den Schätzungen. Die Erinnerungen lagen in allen Fällen entgegen der getroffenen Entscheidung bei der Abgabe der Schätzung. Die signifikant gewordene Interaktion zwischen der „Rückmeldung“ und der „Schätzung“ bestätigte zunächst den Haupteffekt der „Schätzung“: Auch hier zeigte sich in allen Fällen eine Erinnerung entgegen der Schätzung. Allerdings zeigten die Einzelvergleiche, daß dieser Effekt am stärksten bei „antonymer“ Rückmeldung ausgeprägt war. Und nur im Falle einer „antonymen“ Rückmeldung bedeuten die Erinnerungen auch eine Annäherung an die tatsächlich erfolgte Rückmeldung. Es zeigte sich daher ein Rückschafehler bei „antonymer“ Rückmeldung. Die Effekte der Interaktion stehen damit im Einklang mit der Betrachtung der Verteilungen. Ein Rückschafehler nur bei antonymer, widersprechender Rückmeldung könnte darauf hindeuten, daß eventuell eine Diskrepanz zwischen eigenem Urteil und der Rückmeldung wichtig für das Auftreten eines Rückschafehlers sein könnte. Da in diesem wie auch in den beiden vorangegangenen Experimenten nicht die Nennung eines konkreten numerischen Wertes erforderlich war, konnte eventuell erst eine widersprechende Rückmeldung einen Effekt auslösen. Die Probanden mußten in diesem Experiment lediglich den Schieberegler verschieben, um ihre Sicherheit anzugeben, daß der Sachverhalt gestiegen wäre. Die Variation, die mit der analogen Skala gewählt wurde, hatte eventuell den unerwünschten Nebeneffekt, daß die Rückmeldung „gestiegen“ respektive „gefallen“ nicht so wirkte wie intendiert. Die Probanden hatten diese Rückmeldung nicht so verstanden, daß sie 100% Sicherheit ausdrückte, der Sachverhalt wäre in Wirklichkeit gestiegen respektive gefallen. Eine solche Rückmeldung, wenn die eigene Entscheidung ebenso gelautet hatte, löste dann keinerlei weiteren Verarbeitungsprozeß aus. Die Rückmeldung wurde einfach gar nicht enkodiert. Erst die Nennung eines konkreten Wertes zur ersten Sitzung, beispielsweise 60% Sicherheit, daß der Sachverhalt gestiegen sei, und die numerische Rückmeldung 100% Sicherheit, daß der Sachverhalt gestiegen sei in der zweiten Sitzung, würde zu einem solchen Enkodierungsprozeß führen. Eine antonyme

Rückmeldung wurde hingegen verarbeitet und enkodiert, da offensichtlich eine Diskrepanz zwischen eigenem Urteil und der Rückmeldung bestand. Allerdings ist bei dieser Interpretation Vorsicht geboten, da hier eine Konfidenzskala verwendet wurde, bei der sich auch schon in den beiden vorangegangenen Experimenten eine starke Tendenz zur Mitte zeigte. Die Einzelvergleiche zeigten auch hier signifikante Unterschiede zwischen Schätzung und Erinnerung, die jedoch bei synonyme Rückmeldung entgegen der Erwartung von der Rückmeldung weg zur Mitte erfolgte. Auch bei Sachverhalten ohne Rückmeldung konnte man diese starke Tendenz zur Mitte beobachten. Es könnte daher sein, daß diese starke Tendenz zur Mitte alle intendierten systematischen Effekte der Rückmeldung überlagert hat. Bei Sachverhalten mit antonymer Rückmeldung wiesen die Tendenz zur Mitte und der Effekt der Rückmeldung in die gleiche Richtung, während bei Sachverhalten mit synonyme Rückmeldung, diese entgegen der Tendenz zur Mitte wies.

Das erlaubt natürlich eine nachträgliche Einschätzung der Größe der Effekte. Dadurch daß der Faktor „Rückmeldung“ systematisch mit und entgegen der Entscheidung der Probanden bei der Abgabe der eigenen Schätzung die Rückmeldung variierte, ist eine Abschätzung des Rückschaufehlers und des Regressionseffekts möglich. Im Falle ohne Rückmeldung erwarten wir, daß der Differenzwert $\Delta\%$ nur die Regression zur Mitte anzeigt. Im Falle „synonyme“ Rückmeldung läuft der Effekt des Rückschaufehlers entgegen der Tendenz zur Mitte, während bei „antonymer“ Rückmeldung der Rückschaufehler in die gleiche Richtung, wie die Regression, weist.

Tabelle 5.2: $\Delta\%$ getrennt nach Schätzung (gestiegen/ gefallen) und Rückmeldung (synonym/ antonym/ keine).

| Schätzung | Rückmeldung | beteiligte Effekte | $\Delta\%$ |
|-----------|-------------|--------------------|------------|
| gestiegen | synonym | RegrE - RF | -5.21 |
| gefallen | synonym | RegrE - RF | 7.40 |
| gestiegen | antonym | RegrE + RF | -12.38 |
| gefallen | antonym | RegrE + RF | 11.04 |
| gestiegen | keine | RegrE | -5.94 |
| gefallen | keine | RegrE | 9.79 |

RegrE= Regressionseffekt; RF= Rückschaufehler

Daraus ergibt sich ein Regressionseffekt aus dem Mittel der Sachverhalte ohne Rückmeldung von $7.87 \Delta\%$ und ein Rückschaufehler von $2.70 \Delta\%$ ⁷. Der Regressionseffekt ist damit mehr als drei Mal so groß wie der Rückschaufehler. Es zeigte sich damit eine starke Tendenz zur Mitte, die den Effekt „synonymer“ Rückmeldung überlagert hat und auch für einen starken Effekt bei Sachverhalten ohne Rückmeldung sorgte.

Die Auswahl der Sachverhalte mit interessanteren Themen bewirkte wie beabsichtigt etwas extremere Schätzungen, aber leider zeigte sich dennoch kein genereller Rückschaufehler. Möglicherweise hätten die Rückmeldungen stärker wirken können, wenn die Rückmeldung extremer erfolgt wäre. Eine Rückmeldung, der Sachverhalt sei in Wirklichkeit gestiegen respektive gefallen, wurde bei Sachverhalten mit synonyme Rückmeldung nicht von der eigenen Entscheidung unterschieden. Eine Rückmeldung, der Sachverhalt sei sehr stark gestiegen respektive gefallen, hingegen würde höchstwahrscheinlich von der eigenen Entscheidung zu unterscheiden sein und dann auch in diesem Fall zu einer Verarbeitung der Rückmeldung führen.

Das Ausbleiben eines generellen Rückschaufehlers in diesem Experiment läßt dann wiederum keine generellen systematischen Effekte in den anderen abhängigen

⁷ Es ergibt sich der Regressionseffekt (RegrE) aus dem Mittel für $\Delta\%$ für die Schätzung (S) „gestiegen“ und „gefallen“: $(5.94 + 9.79)/2=7.87$). Der Rückschaufehler (RF) bei synonyme Rückmeldung (RM) ergibt sich aus der Differenz des RegrE minus dem Mittelwert von $\Delta\%$ für die S „gestiegen“ und „gefallen“: $(7.87-(5.21+7.40))/2=1.56$). Der RF bei antonymer RM ergibt sich aus der Differenz des Mittelwerts von $\Delta\%$ für die S „gestiegen“ und „gefallen“ minus dem RegrE: $((12.38+11.04)/2-7.87=3.84)$. Der RF ergibt sich dann aus dem Mittel von „synonymer“ und „antonymer“ RM: $((1.56+3.84)/2=2.70)$.

Variablen erwarten. Dennoch legte die signifikante Interaktion zwischen Rückmeldung und Schätzung nahe, daß sich in der Bedingung mit antonymer Rückmeldung Effekte zeigen könnten.

Bei der Analyse der Wiedererkennensleistungen zeigte sich sowohl bei den Treffern als auch in der Diskriminationsleistung und im Antwort-Bias kein einheitliches Bild. Die Schätzung „gestiegen“ bewirkte mehr Treffer, und es wurde sensitiver reagiert. Dieser Befund entspräche auch der Interpretation, daß eine Antworttendenz zugunsten des Ausgangs „gestiegen“ vorhanden war, die sich ja bereits schon in der Verteilung der Daten gezeigt hatte. Sachverhalte ohne Rückmeldung zeigten eine höhere Trefferrate für Argumente, die für „fallen“ sprachen, wobei auch die Diskriminationsleistung hier besser war und der Antwort-Bias eine konservativere Antwortstrategie andeutete. Dieser Befund lässt sich nicht eindeutig interpretieren.

Die Analyse der Reaktionszeiten zeigte keinerlei signifikante Effekte.

Die Konsequenzen für das folgende vierte Experiment lagen auf der Hand. Das grundsätzliche Vorgehen mußte weiter verändert werden. Ohne einen Rückschaufehler konnten keine weitergehenden Effekte erwartet werden. Aus diesem Grunde sollte das vierte Experiment auf die analoge Konfidenzskala verzichten. Die Abgabe eines konkreten numerischen Wertes sollte, wie bereits erläutert, eher dazu angetan sein, einen Rückschaufehler zu verursachen. Die Methode, selektive Aktivierung über Reaktionszeitunterschiede in einem Wiedererkennensparadigma nachweisen zu wollen, war eventuell auch nicht geeignet, so daß im folgenden Experiment eine einfache freie Wiedergabe der Argumente verlangt werden sollte.

6 Experimente mit Erinnerungsmaßen

6.1 Experiment 4: Freie Wiedergabe der Argumente im Gedächtnisdesign

EINLEITUNG

Dieses vierte Experiment, wie auch das folgende fünfte Experiment, folgt einer anderen Methode zum Nachweis *selektiver Aktivierung*: Es sollte jeweils in einem nachfolgenden Test die Erinnerung an die präsentierten Argumente in einer freien Wiedergabe getestet werden. Des weiteren wurde in diesem Experiment ein Fragebogen gewählt. Es wurden die gleichen 24 Sachverhalte aus dem dritten Experiment verwendet. Außerdem wurde die Betonung der Rückmeldung variiert. Die Image-Sets wurden weiterhin experimentell vermittelt.

Die Verwendung eines Fragebogens statt einer Versuchsdurchführung am Computer folgte praktischen Erwägungen. So sollten im Anschluß an Vorlesungen leicht Probanden rekrutiert werden können, ohne daß es für diese erforderlich sein würde, sich in einen Versuchsraum zu begeben. Zumal in diesem Experiment, im Gegensatz zu den vorangegangenen, auch eine explizite numerische Rückmeldung gefordert war, die sehr gut schriftlich zu erledigen war. Besonders auch für die freie Wiedergabe der Argumente in der zweiten Sitzung bot sich die Wahl eines Fragebogens an, da davon auszugehen war, daß handschriftliche Aufzeichnungen auf einem extra Bogen für die meisten Probanden angenehmer und schneller zu anzufertigen waren als diese am Computer eintippen zu müssen. Die Fragebögen enthielten exakt die gleichen 24 Sachverhalte aus Experiment 3. Der eigentliche Unterschied zum dritten Experiment, abgesehen von der Wahl eines Fragebogens, fand in der zweiten Sitzung statt. Es gab zweierlei Rückmeldungen. Eine Gruppe von Probanden erhielt eine extreme Rückmeldung derart, daß rückgemeldet wurde, der Sachverhalt wäre „stark gestiegen“ respektive „stark gefallen“. Die andere Gruppe von Probanden erhielt eine normale Rückmeldung, in der lediglich rückgemeldet wurde, der Sachverhalt wäre „gestiegen“ respektive „gefallen“. Diese Variation der Rückmeldung sollte die verfälschende Wirkung auf die Erinnerung der Probanden verstärken. Der dritte Unterschied betraf, wie bereits erwähnt, den Test: Probanden sollten sich in freier Wiedergabe an Argumente zu jedem Sachverhalt erinnern.

Die Probanden wurden im Gegensatz zu den vorangegangenen Experimenten gebeten, durch eine konkrete Prozentangabe (0-100%) ihre Sicherheit anzugeben, ob der entsprechende Sachverhalt gestiegen oder gefallen war. Eine Prozentangabe von 0% bedeutete „ganz sicher gefallen“ und 100% bedeutete „ganz sicher gestiegen“. Ein Wert um 50% bedeutete, daß keine Entscheidung getroffen werden konnte.

Analog zum dritten Experiment bekamen die Probanden auch in diesem vierten Experiment zum zweiten Termin, genau eine Woche nach dem ersten, erneut alle 24 Sachverhalte (allerdings ohne Argumente) vorgelegt, wobei für jeweils acht der Sachverhalte eine bestätigende (synonyme) für weitere acht eine widersprechende (antonyme) und für acht keine Rückmeldung gegeben wurde.

Die Erinnerung an die Konfidenzurteile sollte sich in den Experimentalbedingungen systematisch von der Kontrollbedingung unterscheiden: Die Rückmeldung, ein Sachverhalt wäre gestiegen, sollte die Sicherheit für „steigen“ erhöhen und für „fallen“ vermindern, während umgekehrt die Rückmeldung, daß ein Sachverhalt gefallen wäre, die Sicherheit für „fallen“ erhöhen und für „steigen“ mindern sollte. Ein solches Ergebnis entspräche einem typischen Rückschaufehler. Die Bedingungsvariation, daß ein Konfidenzurteil durch die Rückmeldung bestätigt (synonyme Rückmeldung) oder widerlegt (antonyme Rückmeldung) wurde, sollte im Falle der Bestätigung eine stärkere Sicherheit und im Falle der Widerlegung eine verminderte Sicherheit für eine Erinnerung im Sinne der Rückmeldung ergeben: Eine Entscheidung für „gestiegen“ im ersten Durchgang und eine synonyme Rückmeldung „gestiegen“ zum Testzeitpunkt sollte eine absolut gesehen stärkere Sicherheit im erinnerten Konfidenzurteil für „steigen“ ergeben als die antonyme Rückmeldung „gefallen“.

In den Experimenten wurde zusätzlich die Erinnerung an die Argumente geprüft. Hier sollte sich ebenfalls ein Zusammenhang mit der Rückmeldung zeigen lassen. Argumente, die für „fallen“ sprechen, sollten häufiger erinnert werden als solche, die für „steigen“ sprechen, wenn die Rückmeldung „fallen“ lautet und entsprechend umgekehrt. Dieses Ergebnis spräche für eine selektive Aktivierung des Image-Sets. Die Anzahl erinnelter Argumente sollte mit synonyme Rückmeldung einen stärkeren Unterschied zeigen als mit antonymer Rückmeldung.

METHODE

Design Der Versuchsplan bestand in einem 2x3x2 Design mit den Faktoren „Betonung“ (extrem/normal), „Rückmeldung“ (synonym/antonym/keine) und „Argument“ (fallen/steigen) (siehe Abb. 6.1). Der Faktor „Betonung“ wurde als Gruppenfaktor realisiert, wobei darauf geachtet wurde, daß Probanden per Zufall den beiden Gruppen zugeordnet wurden. Die Faktoren „Rückmeldung“ und „Argument“ wurden als Meßwiederholungsfaktoren realisiert. Die Gruppen des Faktors „Betonung“ unterschieden sich darin, daß Probanden in der Gruppe mit „extremer“ Rückmeldung, im Gegensatz zu der Gruppe mit „normaler“ Rückmeldung, durch die Hinzufügung des Adjektivs „stark“, also „Der Sachverhalt ist stark gestiegen“, respektive „stark gefallen“, eine betontere, zwingendere Rückmeldung erhielten. Die Zuordnung zu dem Faktor „Rückmeldung“ wurde durch systematische Permutation einer einmaligen Zufallsreihenfolge realisiert, so daß alle Fragen über alle Bedingungen gleich häufig verwirklicht wurden. Eine mit „gestiegen“ bezeichnete Rückmeldung zeigte an, daß der zu beurteilende Sachverhalt in Wirklichkeit gestiegen sei. Entsprechend zeigte eine mit „gefallen“ bezeichnete Rückmeldung an, daß der zu beurteilende Sachverhalt in Wirklichkeit gefallen sei. Die Bedingung „keine“ Rückmeldung diente als Kontrollbedingung. Der Faktor „Argument“ gruppiert die Wiedererkennensleistung der Argumente in Argumente, die dafür sprachen, daß der entsprechende Sachverhalt gestiegen („steigen“) und Argumente, die dafür sprachen, daß er gefallen („fallen“) war. Die Argumente sprachen jeweils zur Hälfte dafür, daß der entsprechende Sachverhalt gestiegen respektive gefallen war. Die Darbietung erfolgte blockweise, wobei die Reihenfolge der Argumente als auch der Blöcke selbst einer Zufallsreihenfolge folgte, so daß alle Argumente gleich häufig an jeder Position präsentiert wurden.

Als abhängige Maße wurden neben der Differenz der Konfidenzurteile zum Zeitpunkt der ersten Einschätzung und dem Zeitpunkt der Erinnerung ($\Delta\%$) die Anzahl erinnerter Argumente in einer freien Wiedergabe erhoben.

Probanden An der Untersuchung nahmen 104 Personen, vorwiegend Studierende aus verschiedenen Fachbereichen der Universität Trier, teil. Davon waren 73 weiblichen und 31 männlichen Geschlechts. Das Altersspektrum reichte von 18 bis 56 Jahren. Der Altersdurchschnitt lag bei 22.5 Jahren. Die Probanden erhielten 12 DM für die Teilnahme am Experiment.

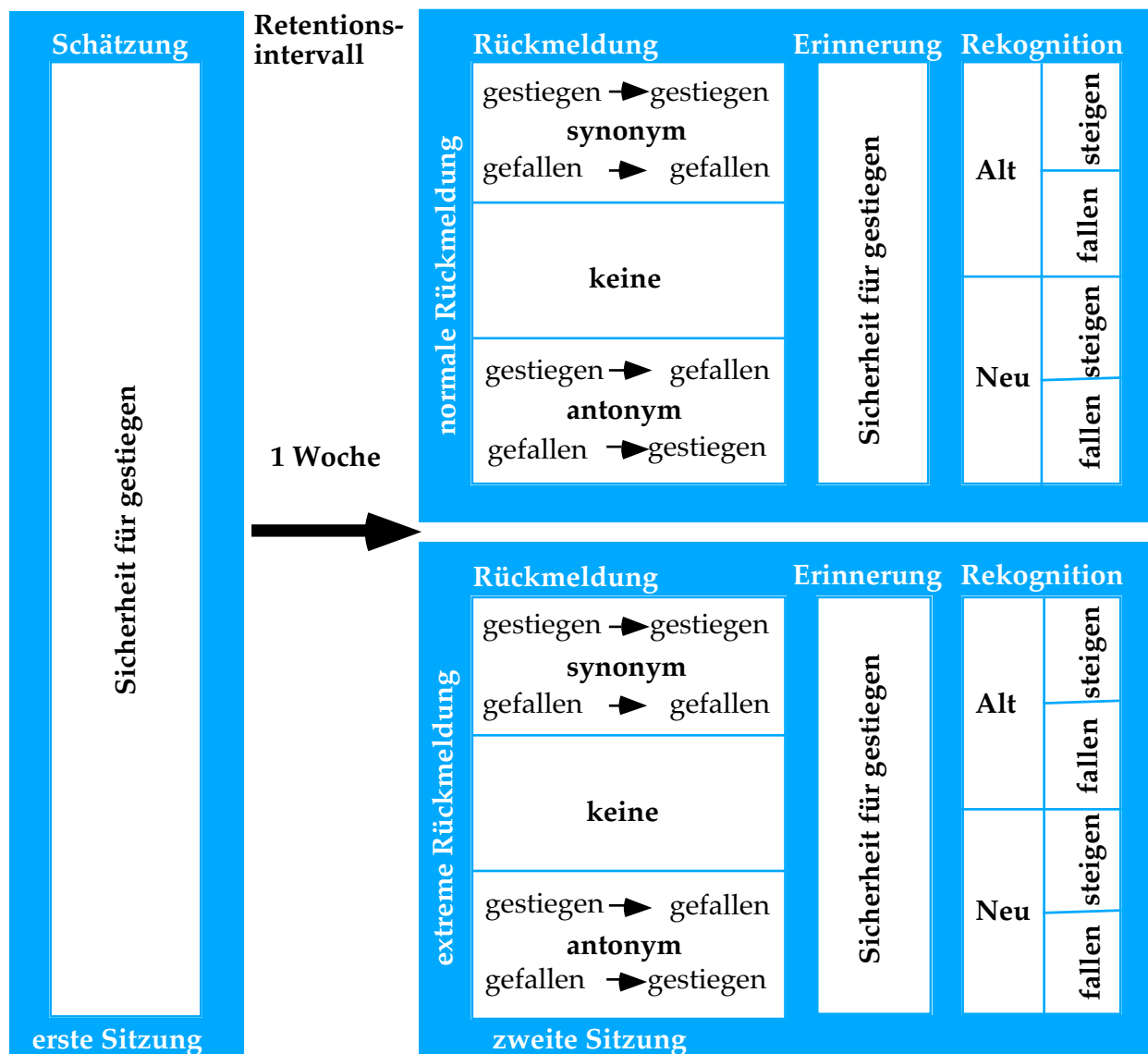


Abbildung 6.1: Versuchsplan von Experiment 4. Gedächtnisdesign mit einer Woche Retentionsintervall mit dem Gruppenfaktor „Betonung“ und Meßwiederholung auf den Faktoren „Rückmeldung“, „Schätzung“ und „Argumente“.

Material 24 Verifikationsaufgaben, die danach beurteilt werden sollten, wie sicher die Probanden jeweils waren, ob ein betreffender Sachverhalt gestiegen oder gefallen war. Wie im vorausgegangenen Experiment wurde bei diesem Experiment Wert darauf gelegt, daß die Themen so gewählt waren, daß sie das allgemeine Interesse der Probanden ansprachen. Indifferente Entscheidungen nahe der 50%-Marke sollten vermieden werden. Im Gegensatz zu den vorangegangenen Experimenten erfolgte dieses Experiment nicht am Computer. Wie bereits erwähnt wurde in diesem Experiment ein Fragebogen verwendet, zum einem, um Flexibilität bei der Rekrutierung von Probanden zu gewinnen, zum anderen, da es leichter schien, die freie Wiedergabe von Argumenten handschriftlich durchzuführen. Die Konfidenzangaben zu den Sachverhalten erfolgte in diesem Experiment, im

Gegensatz zu den vorangegangenen, durch Angabe eines konkreten numerischen Werts.

Für den zweiten Termin wurde individuell für jeden Probanden ein zweiter Fragebogen zusammengestellt, der jeweils acht synonyme, acht antonyme und acht Sachverhalte ohne Rückmeldungen enthielt.

Ablauf Die Durchführung folgte dem Gedächtnisdesign und fand in zwei Sitzungen mit einer Woche Abstand statt. Es konnten immer mehrere Personen gleichzeitig teilnehmen. Die Probanden wurden gebeten, nach dem Ende von Lehrveranstaltungen im Raum zu verweilen und einen Fragebogen auszufüllen. Die Gruppengröße schwankte von 2 bis 8 Personen, die gleichzeitig den ersten Teil des Experiments bearbeiteten. Zum ersten Termin erhielten die Probanden den ersten Fragebogen mit der Instruktion, jeweils den Sachverhalt und die Argumente sorgfältig durchzulesen und durch Prozentangaben ihre Sicherheit für eine Veränderung in Richtung gestiegen bzw. gefallen anzugeben. Die Probanden wurden ausdrücklich darauf hingewiesen, daß die Argumente zur Bearbeitung des zweiten Teils wichtig wären. Zusätzlich enthielt die Instruktion auch ein Beispiel, wie die Aufgabe zu bearbeiten sein werde. Die Probanden gaben dann für alle 24 Sachverhalte ihr Konfidenzurteil ab. Sie hatten beliebig viel Zeit für diese Aufgabe. Im Durchschnitt benötigten sie ca. 30 Minuten zur Bearbeitung dieses ersten Teils des Experimentes und wurden gebeten, pünktlich in einer Woche zum zweiten Termin zu erscheinen.

In der Zwischenzeit wurde individuell für jeden Probanden ein zweiter Fragebogen mit den entsprechenden Sachverhalten aber ohne die Argumente zusammengestellt. Zu Beginn des zweiten Termins nach einer Woche wurde den Probanden dieser Fragebogen präsentiert, mit der Instruktion, daß ihnen nun die gleichen Sachverhalte wie zum ersten Termin vorgelegt würden und es darum ginge, sich jetzt möglichst genau an das eigene Konfidenzurteil aus der vorangegangenen Sitzung zu erinnern. Sie wurden darauf hingewiesen, daß bei mehreren der Sachverhalte der tatsächliche Verlauf als Hilfestellung, ob gestiegen oder gefallen, mitgeteilt werden würde. Die Probanden trugen für alle 24 Sachverhalte ihre Erinnerung an ihr Konfidenzurteil ein. Es bestand wiederum keine Zeitbegrenzung für diese Aufgabe. Nach Beendigung dieser Aufgabe wurden die Probanden instruiert, sich noch einmal alle Sachverhalte des Fragebogens vorzunehmen und zu versuchen, sich an möglichst viele der Argumente zu dem entsprechenden Sachverhalt zu erinnern und diese jeweils in Stichworten zu notieren. Auch für diese Aufgabe bestand keine Zeitbegrenzung. Die Probanden benötigten ca. 60 Minuten

zur Bearbeitung des gesamten zweiten Teils des Experiments und wurden auf Wunsch abschließend über die Hintergründe des Experiments aufgeklärt.

Auswertung Die Verteilung der Schätzungen für „gefallen“ respektive „gestiegen“ zeigten wieder eine Bevorzugung des Ausgangs „gestiegen“. Die Probanden entschieden sich in 61.35% für „gestiegen“. Bei „synonym“ rückgemeldeten Sachverhalten wurde in 59.22%, bei „antonym“ rückgemeldeten Sachverhalten in 60.75% und bei Sachverhalten ohne Rückmeldung in 63.65% „gestiegen“ gewählt. Die Datenmenge für die Mittelwerte des Faktors „Schätzung“ sind daher noch ausreichend groß und vergleichbar.

Die Zuordnung der Argumente in die entsprechenden Kategorien in Argumente, die für „gestiegen“ respektive für „gefallen“ sprachen, geschah durch zwei Rater. Die Übereinstimmung der zwei Rater bezüglich der Zuordnung zu den Kategorien war sehr hoch ($K=.89$; siehe Cohen; 1960). Die Argumente der Probanden ließen sich außer in drei Fällen eindeutig den Kategorien „steigen“ und „fallen“ des Faktors „Argument“ zuordnen. Neben alten, mit den Sachverhalten präsentierten, traten auch neue, erfundene Argumente auf, die ebenfalls in die Kategorien „steigen“ und „fallen“ eingeordnet wurden. Der Anteil neuer Argumente lag bei 39.4% („steigen“: 56.4% alte und 43.6% neue Argumente; „fallen“: 65.9% alte und 34.1% neue Argumente). Es wurden jeweils getrennte Analysen mit alten und neuen Argumenten durchgeführt. Es ergaben sich keine Unterschiede in den Ergebnissen. Auf die getrennte Darstellung der Ergebnisse für neue und alte Argumente wurde daher im Ergebnisteil verzichtet. Es wird nur die Analyse aller (alte und neue) Argumente dargestellt.

ERGEBNISSE

Verteilung der Daten Betrachtet man zunächst die Verteilung von Schätzungen und Erinnerungen, über die Sicherheit in Prozent, daß der Sachverhalt zugenommen habe getrennt nach der Rückmeldung (synonym/antonym/keine), sieht man, daß in der Bedingung mit „synonymer“ Rückmeldung (siehe Abb. 6.2) sowohl für die Rückmeldung „gestiegen“ als auch „gefallen“ die Mittelwerte der Verteilungen von Schätzungen und Erinnerungen in Richtung der Skalenmitte wanderten (Schätzung „gestiegen“: 68.98% zu 64.87%; Schätzung „gefallen“: 31.60% zu 41.15%). Die Streuung nahm in der Verteilung der Erinnerungen zu (Schätzung „gestiegen“: 13.20 zu 14.43; Schätzung „gefallen“: 10.67 zu 17.27).

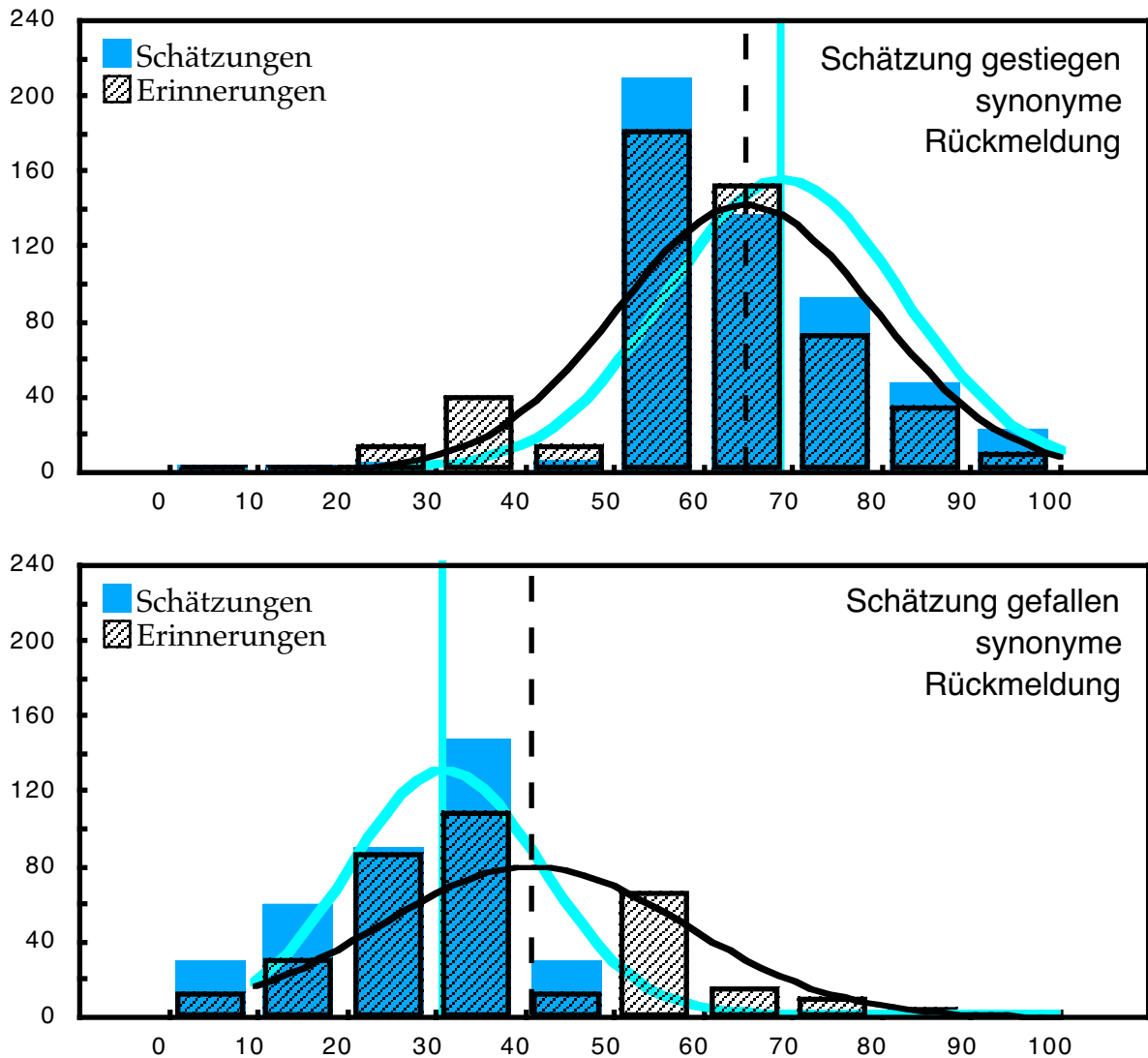


Abbildung 6.2: Relative Häufigkeiten für Schätzungen und Erinnerungen pro Zehner-Kategorie gegliedert nach der Rückmeldung bei synonyme Rückmeldung.

In der Bedingung mit „antonymer“ Rückmeldung (siehe Abb. 6.3) wanderten ebenfalls sowohl für die Rückmeldung „gestiegen“ als auch „gefallen“ die Mittelwerte der Verteilungen von Schätzungen und Erinnerungen in Richtung der Skalenmitte (Schätzung „gestiegen“: 68.92% zu 59.65%; Schätzung „gefallen“: 37.04% zu 47.98%). Die Streuung nahm in der Verteilung der Erinnerungen zu (Schätzung „gestiegen“: 10.50 zu 14.80; Schätzung „gefallen“: 17.15 zu 17.93).

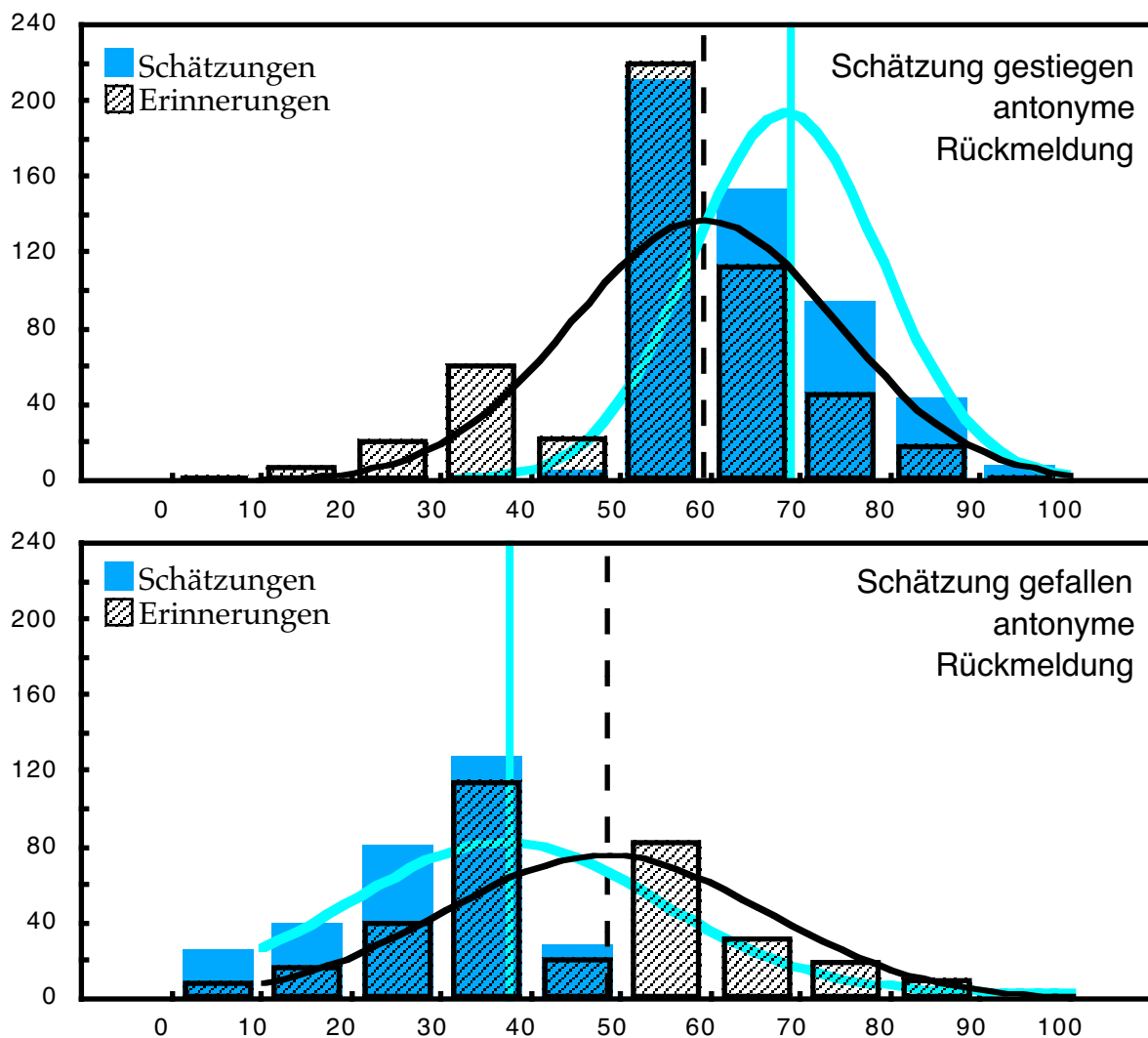


Abbildung 6.3: Relative Häufigkeiten für Schätzungen und Erinnerungen pro Zehner-Kategorie gegliedert nach der Rückmeldung bei antonymer Rückmeldung.

Teilt man in der Bedingung „keine“ Rückmeldung (siehe Abb. 6.4) die Schätzungen und Erinnerungen nach der Lage der Schätzungen in solche, die „steigen“ und solche, die „fallen“ geschätzt haben, wanderten ebenfalls sowohl für die Rückmeldung „gestiegen“ als auch „gefallen“ die Mittelwerte der Verteilungen von Schätzungen und Erinnerungen in Richtung der Skalenmitte (Schätzung „gestiegen“: 65.94% zu 61.77%; Schätzung „gefallen“: 32.12% zu 43.15%). Die Streuung nahm in der Verteilung der Erinnerungen zu (Schätzung „gestiegen“: 13.71 zu 15.07; Schätzung „gefallen“: 9.73 zu 16.44).

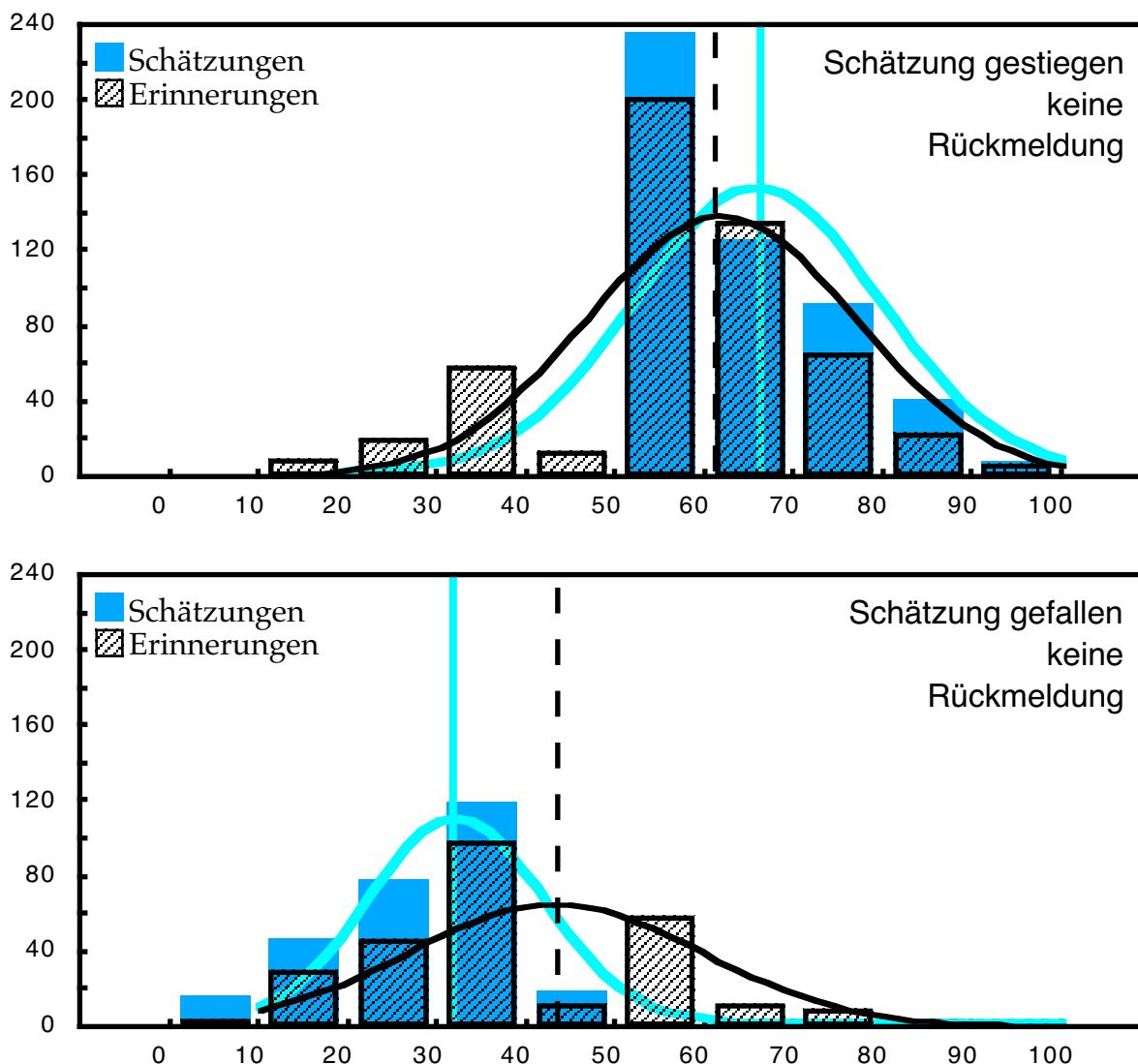


Abbildung 6.4: Relative Häufigkeiten für Schätzungen und Erinnerungen pro Zehner-Kategorie gegliedert nach der Schätzung in der Kontrollbedingung.

% korrekt Die Analyse der Anzahl korrekter Erinnerungen in der Kontroll- und den beiden Experimentalbedingungen erlaubt eine Beurteilung der Gedächtnisbeeinträchtigung durch die experimentelle Manipulation. Die Anzahl korrekter Erinnerungen ergab signifikante Unterschiede (siehe Tabelle 6.1). Die Anzahl korrekter Erinnerungen ergab signifikante Unterschiede zwischen synonymem (35.75%), antonymem (27.33%) und keiner Rückmeldung (37.01%) ($\chi^2_{(2)}=20.84$; Power (1-beta)=0.79). Auch die Trennung nach der Präferenz der Schätzung in „gestiegen“ (35.45%) versus „gefallen“ (30.63%) zeigte signifikante Unterschiede ($\chi^2_{(1)}=6.45$; Power (1-beta)=0.79). Weder innerhalb „synonymer“ ($\chi^2_{(1)}=2.57$) noch „antonymer“ Rückmeldung ($\chi^2_{(1)}=1.18$) zeigte sich ein Unterschied zwischen der Präferenz in der Schätzung. Jedoch bei „keiner“ Rückmeldung ($\chi^2_{(1)}=4.63$; Power (1-beta)=0.79) zeigte sich ein signifikanter Unterschied (siehe Tabelle 6.1): Es fanden sich mehr korrekte Erinnerungen, wenn die Schätzung „gestiegen“ gelautet

hatte. Die Anzahl korrekter Erinnerungen in der Rückmeldung unterschieden sich bei der Schätzung „gestiegen“ ($\chi^2_{(2)}=20.71$; Power (1-beta)=0.79), jedoch nicht für „gefallen“ ($\chi^2_{(2)}=1.70$): Bei einer Entscheidung für „gestiegen“ fanden sich weniger korrekte Erinnerungen, wenn es sich um eine „antonyme“ Rückmeldung handelte.

Tabelle 6.1: % korrekt getrennt nach Schätzung (gestiegen/gefallen) und Rückmeldung (synonym/antonym/keine).

| Schätzung | Rückmeldung | Anzahl | Prozent |
|-----------|-------------|--------|---------|
| gestiegen | synonym | 193 | 37.92 |
| gefallen | synonym | 113 | 32.56 |
| gestiegen | antonym | 146 | 28.68 |
| gefallen | antonym | 86 | 25.29 |
| gestiegen | keine | 207 | 39.66 |
| gefallen | keine | 85 | 31.84 |

Konfidenzurteile Bei einer dreifaktoriellen Varianzanalyse mit dem Gruppenfaktor „Betonung“ (normal/extrem) und den Meßwiederholungsfaktoren „Rückmeldung“ (synonym/antonym/keine) und „Schätzung“ (gefallen, gestiegen) über die Differenz $\Delta\%$, wurden weder der Gruppenfaktor „Betonung“ ($F_{(1,101)}=.20$) noch der Meßwiederholungsfaktor „Rückmeldung“ ($F_{(2,202)}=.92$) signifikant. Es zeigte sich jedoch ein Haupteffekt der „Schätzung“ ($F_{(1,101)}=262.63$; Power (1-beta)=0.97): Die Schätzung „gefallen“ ($4.07\Delta\%$) resultierte in einem positivem Wert für $\Delta\%$, der sich deutlich von dem negativem Wert für $\Delta\%$ für die Schätzung „gestiegen“ ($-3.93\Delta\%$) unterschied. Die Interaktion (siehe Abb. 6.5) zwischen „Rückmeldung“ und „Schätzung“ wurde ebenfalls signifikant ($F_{(2,202)}=18.18$; Power (1-beta)=0.95). Bei „synonymer“ Rückmeldung unterschied sich die Schätzung „gestiegen“ ($-2.60\Delta\%$) von „gefallen“ ($2.98\Delta\%$) ($t_{(103)}=5.67$; $p=.0000001$). Bei „antonymer“ Rückmeldung unterschied sich die Schätzung „gefallen“ ($-6.05\Delta\%$) ebenfalls von „gestiegen“ ($5.39\Delta\%$) ($t_{(103)}=17.49$; $p=.0000000$). Auch in der Bedingung „keine“ Rückmeldung unterschieden sich die Differenzwerte $\Delta\%$ bezüglich der zur Schätzung getroffenen Entscheidung. Eine Entscheidung für „gestiegen“ ($-3.15\Delta\%$) führte zu einem negativem $\Delta\%$, und eine Entscheidung für „gefallen“ ($3.83\Delta\%$) zu einem positiven $\Delta\%$ ($t_{(103)}=11.64$; $p=.0000000$). Die Schätzung „gestiegen“ zeigte Unterschiede in den Rückmeldungen ($F_{(2/202)}=118.07$; $p=.0000001$). Ein Tukey Test zeigte, daß sich jeweils nur die „antonyme“ Rückmeldung von den anderen unterschied

(synonym/antonym $p=.00002$; keine/antonym ebenfalls $p=.00002$). Die Schätzung „gefallen“ zeigte gleichfalls die Unterschiede in den Rückmeldungen ($F_{(2/202)}=79.17$; $p=.0000001$). Ein Tukey Test zeigte, daß sich nur die „antonyme“ Rückmeldung von der Bedingung mit „synonymer“ und „keiner“ Rückmeldung unterschied (synonym/antonym $p=.000002$; keine/antonym ebenfalls $p=.000002$). Die anderen zweifachen Interaktionen und die dreifach Interaktion wurden nicht signifikant (alle $F<1.97$).

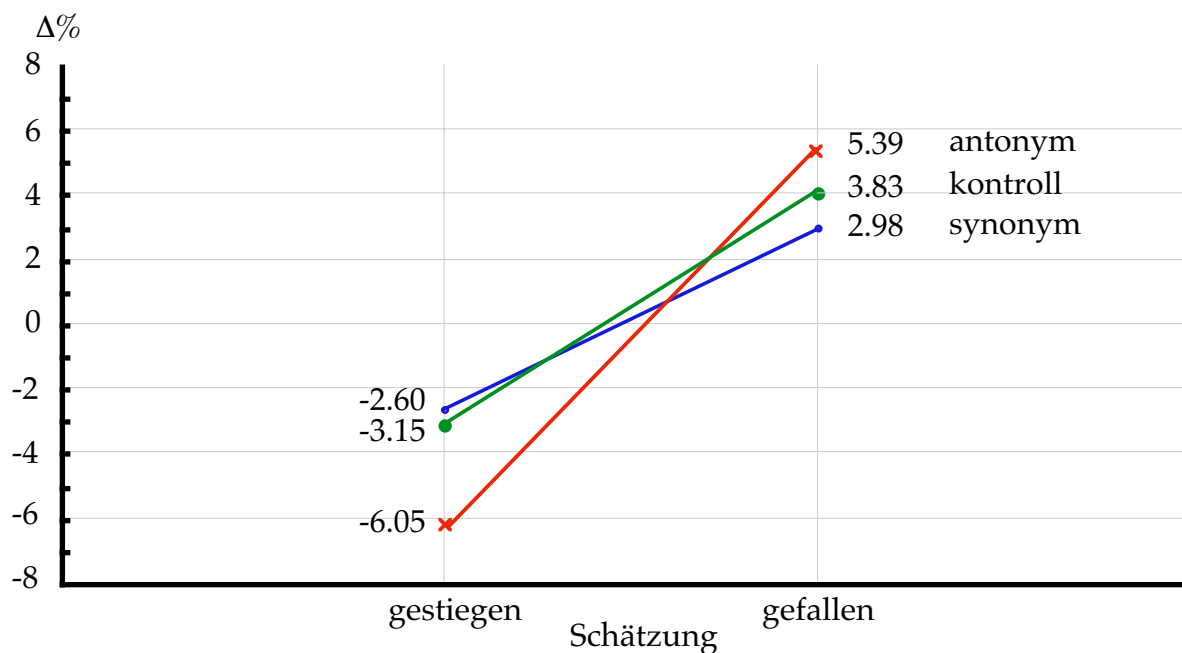


Abbildung 6.5: Interaktion zwischen „Rückmeldung“ und „Schätzung“ über die Differenz im Konfidenzurteil. Eine „antonyme“ Rückmeldung erzeugt eine Annäherung der Erinnerung an die Rückmeldung (positiver Wert für die Schätzung „gefallen“ und negativer Wert für die Schätzung „gestiegen“), während eine synonyme Rückmeldung eine Abweichung von der Rückmeldung erzeugt.

Freie Wiedergabe Die Analyse der Anzahl erinnerter Argumente in einer freien Wiedergabe erlaubt die Beurteilung eines direkten Effekts auf die Güte der Erinnerung durch die experimentelle Manipulation. Eine vierfaktorielle Varianzanalyse mit dem Gruppenfaktor „Betonung“ (normal/extrem) und den Meßwiederholungsfaktoren „Rückmeldung“ (synonym/antonym/keine), „Schätzung“ (gefallen, gestiegen) und „Argument“ (fallen/steigen) über die Anzahl erinnerter Argumente in einer freien Wiedergabe zeigte weder einen signifikanten Effekt des Gruppenfaktors „Betonung“ ($F_{(1,101)}=.07$) noch des Meßwiederholungsfaktors „Rückmeldung“ ($F_{(2,102)}=.14$).

Der Faktor „Schätzung“ zeigte jedoch signifikante Unterschiede in der Anzahl erinnerter Argumente ($F_{(1,101)}=29.45$; Power (1-beta)=0.97): Lautete die Schätzung „gestiegen“ (.59), wurden mehr Argumente erinnert, als wenn die Schätzung „gefallen“ (.47) gelautet hatte. Es zeigte sich kein Haupteffekt des Faktors „Argument“ ($F_{(1,101)}=2.93$).

Die zweifach Interaktion zwischen „Rückmeldung“ und „Schätzung“ (siehe Abb. 6.6) wurde signifikant ($F_{(2,202)}=50.86$; Power (1-beta)=0.95): Der Faktor „Rückmeldung“ zeigte signifikante Unterschiede, wenn die Schätzung „gefallen“ gelautet hatte ($F_{(2,202)}=39.35$, $p=.0000001$). Ein Tukey Test zeigte, daß mit „antonymer“ Rückmeldung mehr Argumente erinnert wurden (.63), als mit „synonymer“ (.41) und ohne Rückmeldung (.37) (synonym/antonym $p=.00002$; keine/antonym $p=.00002$). Der Faktor „Rückmeldung“ zeigte ebenfalls signifikante Unterschiede, wenn die Schätzung „gestiegen“ gelautet hatte ($F_{(2,202)}=40.30$, $p=.0000001$). Ein Tukey Test zeigte, daß mit „antonymer“ Rückmeldung weniger Argumente erinnert wurden (.42), als mit „synonymer“ (.66) und ohne Rückmeldung (.70) (synonym/antonym $p=.00002$; keine/antonym $p=.00002$). Bei „synonymer“ Rückmeldung wurden mehr Argumente erinnert, wenn die Schätzung „gestiegen“ gelautet hatte ($F_{(1,101)}=31.57$, $p=.0000001$). Bei „antonymer“ Rückmeldung wurden mehr Argumente erinnert, wenn die Schätzung „gefallen“ gelautet hatte ($F_{(1,101)}=31.59$, $p=.0000001$). Bei „keine“ Rückmeldung wurden mehr Argumente erinnert, wenn die Schätzung „gestiegen“ gelautet hatte ($F_{(1,101)}=70.65$, $p=.0000001$).

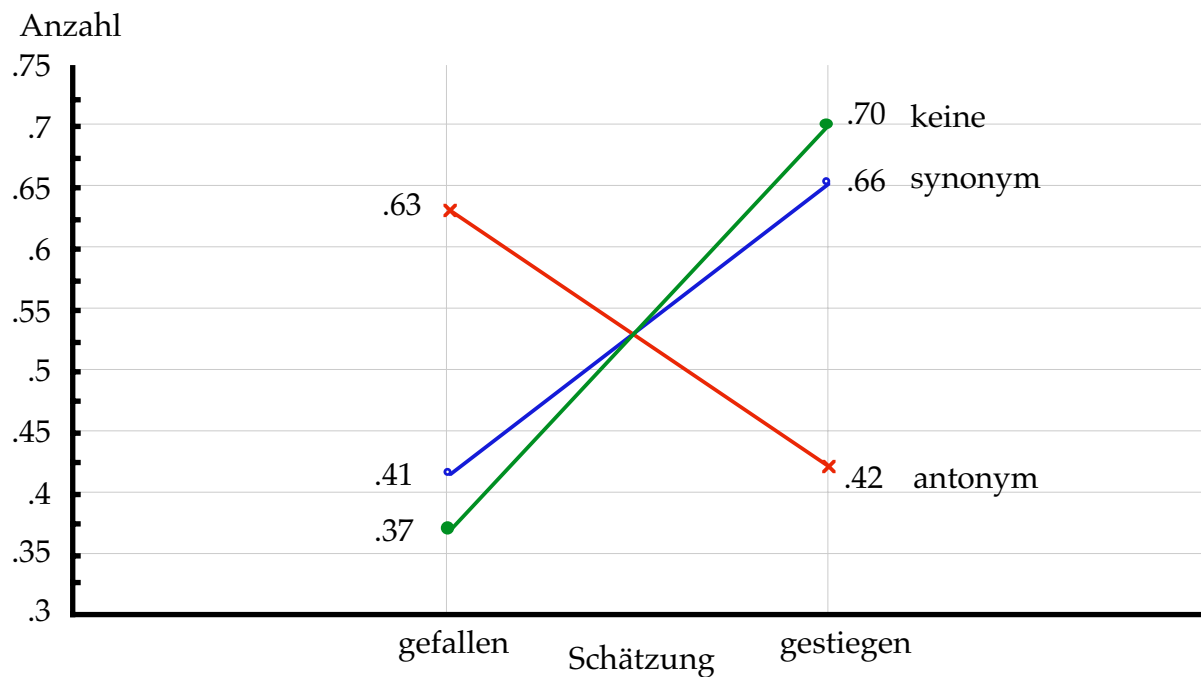


Abbildung 6.6: Interaktion Rückmeldung (synonym/antonym/keine) mit Schätzung (gestiegen, gefallen) über die Anzahl erinnerter Argumente in der freien Wiedergabe.

Die zweifach Interaktion zwischen „Rückmeldung“ und „Argument“ (siehe Abb. 6.7) wurde signifikant ($F_{(2,202)}=4.80$; Power (1-beta)=0.95): Die Einzelvergleiche über den Faktor „Schätzung“ zeigten nur in der Bedingung „keine“ Rückmeldung ($F_{(1,101)}=10.13$, $p=.0019$) signifikante Unterschiede zwischen Argumenten, die für „fallen“ (.50) bzw. „steigen“ (.57) sprachen. Der Faktor „Rückmeldung“ wirkte sich weder bei Argumenten, die für „fallen“ sprachen, noch bei solchen, die für „steigen“ sprachen, aus (alle $F < 2.73$).

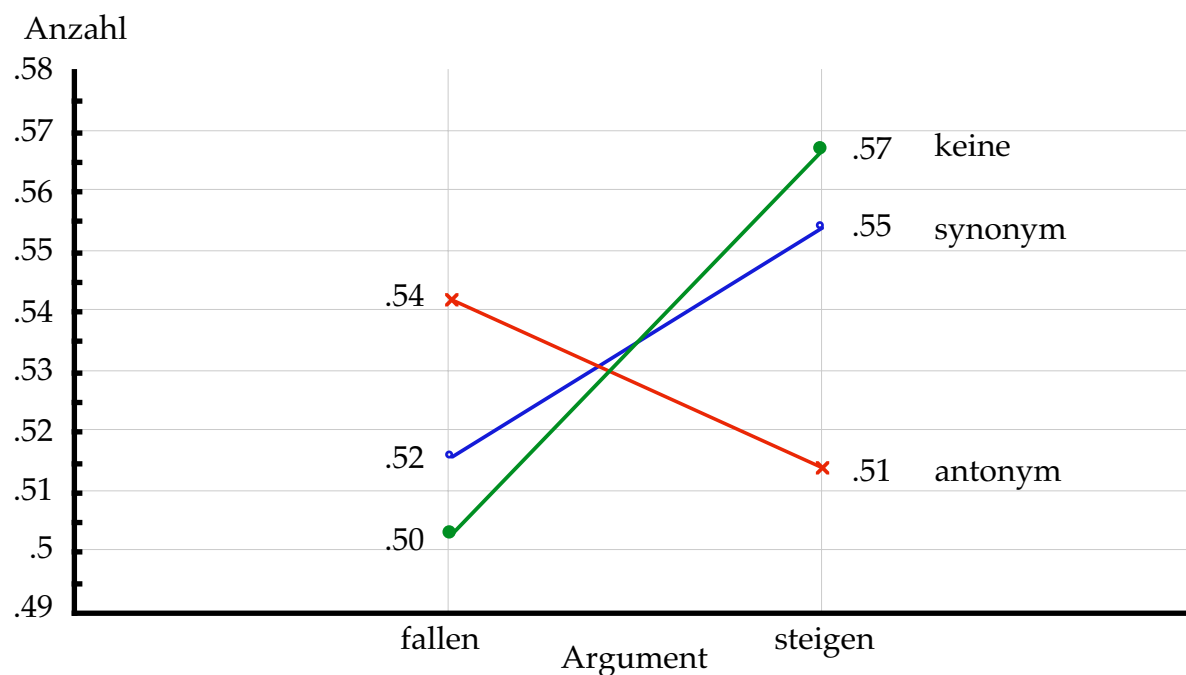


Abbildung 6.7: Interaktion Rückmeldung (synonym/antonym/keine) mit Argument (fallen, steigen) über die Anzahl erinnerter Argumente in der freien Wiedergabe.

Ebenfalls signifikant wurde die Interaktion zwischen „Schätzung“ und „Argument“ ($F_{(1,101)}=94.91$; Power (1-beta)=0.97) (siehe Abb. 6.8): Bei der Schätzung „gestiegen“ wurden signifikant mehr Argumente, die für „steigen“ (.67) sprachen, als für „fallen“ (.51) erinnert ($F_{(1,101)}=61.28$, $p=.0000001$). Bei der Schätzung „gefallen“ wurden signifikant mehr Argumente, die für „fallen“ (.53) sprachen, als für „steigen“ (.42) erinnert ($F_{(1,101)}=30.14$, $p=.0000003$). Argumente, die für „steigen“ sprachen, wurden öfter bei der Schätzung „gestiegen“ (.67) als bei der Schätzung „gefallen“ (.42) erinnert ($F_{(1,101)}=90.27$, $p=.0000001$).

Die anderen zweifachen Interaktionen wurden nicht signifikant (alle $F < 2.91$).

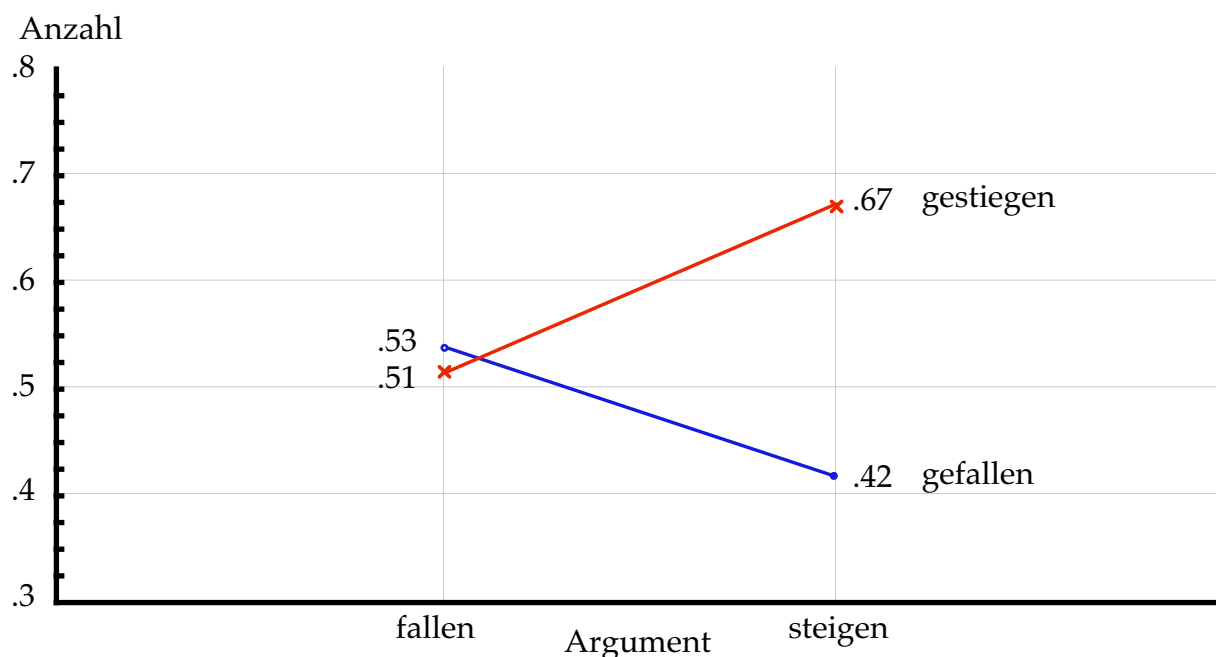


Abbildung 6.8: Interaktion Schätzung (gestiegen, gefallen) mit Argument (fallen/steigen) über die Anzahl erinnertes Argumente in der freien Wiedergabe.

Die dreifach Interaktion zwischen „Rückmeldung“, „Schätzung“ und „Argument“ (siehe Abb. 6.9) wurde ebenfalls signifikant ($F_{(2,202)}=32.92$): Die Schätzung „gestiegen“ bewirkte bei Argumenten, die für „fallen“ sprachen, signifikante Unterschiede über den Faktor „Rückmeldung“ ($F_{(2,202)}=4.57$; $p=.01$). Ein Tukey Test zeigte, daß gegenüber der Rückmeldung „antonym“ (.46) bei „keine“ Rückmeldung (.57) mehr Argumente erinnert wurden (keine/antonym $p=.007$). Die Schätzung „gefallen“ bewirkte bei Argumenten, die für „fallen“ sprachen, signifikante Unterschiede über den Faktor „Rückmeldung“ ($F_{(2,202)}=11.87$; $p=.00001$). Ein Tukey Test zeigte, daß gegenüber der Rückmeldung „synonym“ (.51) und „keine“ Rückmeldung (.44) bei der Rückmeldung „antonym“ (.63) mehr Argumente erinnert wurden (synonym/antonym $p=.01$; keine/antonym $p=.00003$). Bei Argumenten, die für „steigen“ sprachen, bewirkte die Schätzung „gestiegen“ signifikante Unterschiede über den Faktor „Rückmeldung“ ($F_{(2,202)}=51.65$; $p=.0000001$). Ein Tukey Test zeigte, daß gegenüber der Rückmeldung „synonym“ (.79) respektive „keine“ Rückmeldung (.83) bei der Rückmeldung „antonym“ (.38) weniger Argumente erinnert wurden (synonym/antonym $p=.00002$; keine/antonym $p=.00002$). Die Schätzung „gefallen“ bei Argumenten, die für „steigen“ sprachen, bewirkte ebenfalls signifikante Unterschiede über den Faktor „Rückmeldung“ ($F_{(2,202)}=54.06$; $p=.0000001$). Ein Tukey Test zeigte wiederum, daß gegenüber der Rückmeldung „synonym“ (.31) respektive „keine“ Rückmeldung (.30) bei der Rückmeldung „antonym“ (.64) mehr Argumente erinnert wurden (synonym/antonym $p=.00002$; keine/antonym $p=.00002$).

Bei der „Rückmeldung“ „synonym“ und der Schätzung „gestiegen“ ($t_{(103)}=6.72$; $p=.0000000$) wurden mehr Argumente, die für „steigen“ (.79) sprachen, als für „fallen“ (.52) erinnert. Bei der Rückmeldung „synonym“ und der Schätzung „gefallen“ ($t_{(103)}=7.39$; $p=.0000000$) wurden mehr Argumente, die für „fallen“ (.51) sprachen, als für „steigen“ (.32) erinnert. Bei der Rückmeldung „synonym“ und Argumenten, die für „steigen“ sprachen ($t_{(103)}=9.02$; $p=.0000000$), wurden mehr Argumente bei der Schätzung „gestiegen“ (.79) als bei der Schätzung „gefallen“ (.31) erinnert.

Bei der Rückmeldung „antonym“ und Argumenten, die für „fallen“ sprachen ($t_{(103)}=3.75$; $p=.00029$), wurden mehr Argumente bei der Schätzung „gefallen“ (.63) als bei der Schätzung „gestiegen“ (.46) erinnert. Bei der Rückmeldung „antonym“ und Argumenten, die für „steigen“ sprachen ($t_{(103)}=5.82$; $p=.0000001$), wurden mehr Argumente bei der Schätzung „gefallen“ (.64) als bei der Schätzung „gestiegen“ (.39) erinnert.

Bei der Bedingung „keine“ Rückmeldung und einer Schätzung, die für „gestiegen“ sprach ($t_{(103)}=6.61$; $p=.0000000$), wurden mehr Argumente, die für „steigen“ (.83) sprachen als für „fallen“ (.57), erinnert. Bei der Bedingung „keine“ Rückmeldung und einer Schätzung, die für „gefallen“ sprach ($t_{(103)}=5.39$; $p=.0000004$), wurden mehr Argumente, die für „fallen“ (.44) sprachen als für „steigen“ (.30), erinnert. Bei der Bedingung „keine“ Rückmeldung und Argumenten, die für „fallen“ sprachen ($t_{(103)}=2.99$; $p=.0035$), wurden bei Schätzungen, die für „gestiegen“ (.57) sprachen, mehr Argumente erinnert als bei einer Entscheidung für „gefallen“ (.44). Bei der Bedingung „keine“ Rückmeldung und Argumenten, die für „steigen“ sprachen ($t_{(103)}=11.02$; $p=.0000000$), wurden bei Schätzungen, die für „gestiegen“ (.83) sprachen, mehr Argumente erinnert, als bei einer Entscheidung für „gefallen“ (.30).

Alle weiteren Einzelvergleiche zeigten keine Signifikanz (alle $t < 2.57$).

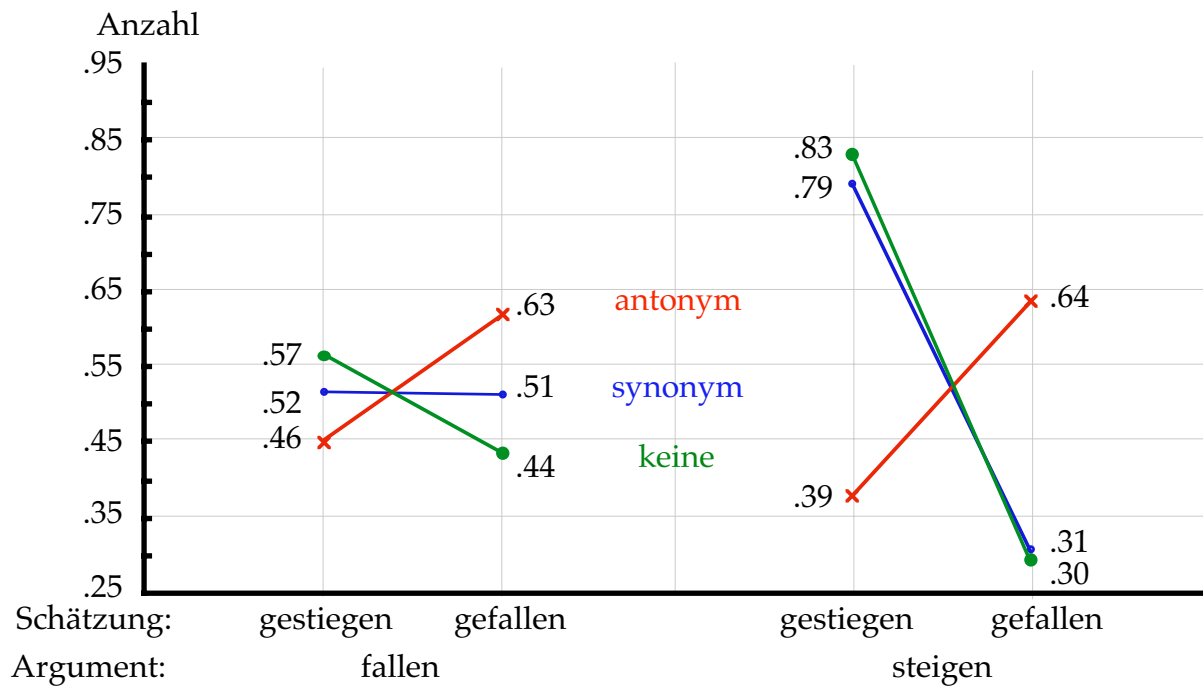


Abbildung 6.9: Interaktion Rückmeldung (synonym/antonym/keine), Schätzung (gestiegen, gefallen) und Argument (fallen/steigen) über die Anzahl erinnertes Argumente in der freien Wiedergabe.

Alle weiteren dreifach Interaktionen sowie die vierfach Interaktion wurden nicht signifikant (alle $F < .78$).

DISKUSSION

Die Verteilung der Daten zeigte schon in der Bedingung mit „synonym“ Rückmeldung eine Erinnerung, die in Richtung der 50%-Marke verschoben war. In der Bedingung mit „antonym“ Rückmeldung zeigte sich das gleiche Bild, wobei hier die Rückmeldung auch eine Verschiebung in die Richtung erwarten ließ. Auch Sachverhalte mit „keiner“ Rückmeldung unterteilt nach Schätzungen, die für „gefallen“ und für „gestiegen“ sprachen, zeigten das gleiche Bild. Auch hier zeigte sich eine starke Verschiebung zur 50%-Marke. Daher zeigte sich in Experiment 4 wieder ganz eindeutig eine Tendenz zur Mitte. Die Streuung nahm in allen Fällen zu. Der Grund für die Zunahme der Streuung ist in diesem vierten Experiment der gleiche, wie im vorangegangenen dritten Experiment: eine Kürzung der Verteilungen an der 50%-Marke, um abhängig vom Faktor „Schätzungen“ eine „synonyme“ oder „antonyme“ Rückmeldung verwirklichen zu können. Es ergibt sich also bei der Betrachtung der Daten ein unvollständiges Bild. Die Streuung muß bei dieser Darstellung der Daten zunehmen, da die wahre Streuung der Schätzung nicht berücksichtigt wird. Auch wenn Probanden im Mittel eher die Einschätzung bevorzugten, der Sachverhalt sei „gestiegen“, lagen doch eine erhebliche Menge an

Daten unter der 50%-Marke. Betrachtet man die vollständigen Verteilungen, so nahmen die Streuungen auch ab: bei „synonymer“ Rückmeldung von 22.00 auf 19.48, bei „antonymer“ Rückmeldung von 21.08 auf 17.17 und ohne Rückmeldung von 20.11 auf 17.81.

Ein anderes Bild ergibt sich auch, wenn man die Daten wie üblich nur nach der Rückmeldung analysiert. Lautete die Rückmeldung „gestiegen“, nahm die Konfidenz von 56.19% für „gestiegen“ auf 58.10% zu. Lautete die Rückmeldung „gefallen“, nahm die Konfidenz für „gestiegen“ von 53.79% auf 52.15% leicht ab. Ohne Rückmeldung änderte sich die Konfidenz für gestiegen kaum. Von 55.18% für gestiegen bei der Schätzung auf 55.84% bei der Erinnerung. Die Streuung wurde jeweils enger. Bei der Rückmeldung „gestiegen“ fiel die Streuung von 21.61 bei der Schätzung auf 17.94 bei der Erinnerung. Bei der Rückmeldung „gefallen“ fiel die Streuung von 21.16 bei der Schätzung auf 18.26 bei der Erinnerung. Und ohne Rückmeldung verminderte sich die Streuung von 20.11 bei der Schätzung auf 17.81 bei der Erinnerung.

Insgesamt zeigte sich also bei der deskriptiven Betrachtung der Daten ein uneinheitliches Bild. Während in der Bedingung mit „antonymer“ Rückmeldung eine Verschiebung der Konfidenzen in Richtung der Rückmeldung auftrat, zeigte die „synonyme“ Rückmeldung keinerlei Wirkung. Es zeigte sich im Gegenteil eine Verschiebung der Erinnerung zur 50%-Marke, die auch bei Sachverhalten ohne Rückmeldung zu erkennen war. Analysiert man die Daten traditionell nur nach der erfolgten Rückmeldung, deutete sich eine Verschiebung in Richtung der Rückmeldung nur bei der Rückmeldung „gefallen“ an. Schon die Schätzungen deuteten a priori auf eine Präferenz für die Rückmeldung „gestiegen“. Die Mittelwerte für die Schätzungen liegen alle über der 50%-Marke (55.05% über alle Schätzungen). Betrachtet man daher die Rückmeldung „gestiegen“ als eine Bestätigung und „gefallen“ als einen subjektiv erlebten Widerspruch, zeigte auch die traditionelle Rückmeldung ein konsistentes Ergebnis. Es deutet sich ein Rückschaufehler nur im Falle „antonymer“, widersprechender Rückmeldung an.

Die Analyse des Anteils korrekter Erinnerungen zeigte im vierten Experiment Unterschiede zwischen Sachverhalten mit „synonymer“, „antonymer“ und ohne Rückmeldung, die darauf zurückzuführen waren, daß über den Faktor „Rückmeldung“ hinweg mehr korrekte Erinnerungen auftraten, wenn die Schätzung „gestiegen“ gelautet hatte. Das Ergebnis deutet auf eine bessere Erinnerungsleistung bei Sachverhalten, die als „gestiegen“ eingeschätzt wurden. Möglicherweise sind positive Einschätzungen leichter und besser erinnerbar als

negative. Die Anteile korrekter Erinnerungen lagen mit durchschnittlich 32.66% in diesem Experiment recht hoch. Der Grund dafür lag an der Verwendung eines Fragebogens, in dem eine explizite diskrete Prozentangabe zur Angabe der Konfidenz gefordert war. Dies ermöglichte gegenüber den vorangegangenen Experimenten, in denen die Konfidenz durch Verschiebung eines Balkens auf einer analogen Skala, ohne Rückmeldung des konkreten Konfidenzwerts, erfasst wurde, überhaupt erst eine korrekte Erinnerung.

Der signifikante Haupteffekt der Schätzung in der Differenz $\Delta\%$ der Konfidenzurteile zeigte ein deutliches systematisches Abweichen der Erinnerungen von den Schätzungen. Die Interaktion zwischen dem Faktor „Rückmeldung“ und „Schätzung“ erlaubt eine genauere Analyse des Effekts. Es zeigte sich in allen Rückmeldungen ein signifikanter Unterschied zwischen der Schätzung „gestiegen“ respektive „gefallen“. Nur im Falle „antonymer“ Rückmeldung entsprach die Richtung der Abweichung auch der Rückmeldung und damit einem Rückschaufehler. Die Rückmeldung, der Sachverhalt wäre „gestiegen“, erhöhte die erinnerte Sicherheit für „gestiegen“, während die Rückmeldung, der Sachverhalt wäre „gefallen“, die erinnerte Sicherheit für gestiegen verminderte. Im Falle synonyme Rückmeldung respektive ohne Rückmeldung entsprach das Abweichen einer Tendenz zur Mitte. Das legt den Schluß nahe, daß sich ein Rückschaufehler nur bei Sachverhalten zeigte, die entgegen der Schätzung rückgemeldet wurden. Dieser Befund deutete sich bereits im dritten Experiment an. Das könnte bedeuten, daß eine Rückmeldung, welche im Kontrast zur eigenen Einschätzung steht, einen stärkeren Effekt auszuüben vermag. Möglicherweise ist dieser Befund jedoch nur auf die verwendete Methode zurückzuführen und entsteht nur bei Konfidenzskalen mit einer diskreten numerischen Antwortskala. Es ist schließlich immer möglich solche Skalen in einfache Kategorien, wie bei diesem vierten Experiment, und in noch viel stärkerem Maße bei der analogen Skala aus den vorausgegangenen Experimenten, in „gefallen“, „unentschieden“ und „gestiegen“ zu unterteilen. Die Konsequenz wäre, daß eine Rückmeldung erst dann Anlaß zur Verarbeitung gäbe, wenn diese in eine andere Kategorie fiel. Mit anderen Worten: eine Rückmeldung hätte nur dann eine Konsequenz, wenn die (eigene) Schätzung beispielsweise in die Kategorie „gestiegen“ gefallen wäre und die Rückmeldung, im Widerspruch dazu „gefallen“ lautete. Eine Rückmeldung „gestiegen“ würde als konform mit der (eigenen) Schätzung erlebt und veranlasste dann auch keine Korrektur. Die Ergebnisse dieses und des vorausgegangenen Experimentes ließen sich so interpretieren.

Wie auch schon im vorangegangenen dritten Experiment erlaubt der Faktor „Rückmeldung“ eine genauere Analyse der beteiligten Effekte. Auch hier ist eine

Abschätzung des Rückschaufehlers und des Regressionseffekts möglich. Im Falle ohne Rückmeldung erwarten wir, daß der Differenzwert $\Delta\%$ nur die Regression zur Mitte anzeigt. Im Falle „synonymer“ Rückmeldung läuft der Effekt des Rückschaufehlers entgegen der Tendenz zur Mitte, während bei „antonymer“ Rückmeldung der Rückschaufehler in die gleiche Richtung, wie die Regression, weißt. Das erlaubt natürlich eine nachträgliche Einschätzung der Größe der Effekte.

Tabelle 6.2: $\Delta\%$ getrennt nach Schätzung (gestiegen/ gefallen) und Rückmeldung (synonym/ antonym/ keine).

| Schätzung | Rückmeldung | beteiligte Effekte | $\Delta\%$ |
|-----------|-------------|--------------------|------------|
| gestiegen | synonym | RegrE - RF | -2.60 |
| gefallen | synonym | RegrE - RF | 2.98 |
| gestiegen | antonym | RegrE + RF | -6.05 |
| gefallen | antonym | RegrE + RF | 5.39 |
| gestiegen | keine | RegrE | -3.15 |
| gefallen | keine | RegrE | 3.83 |

RegrE= Regressionseffekt; RF= Rückschaufehler

Daraus ergibt sich ein Regressionseffekt aus dem Mittel der Sachverhalte ohne Rückmeldung von $3.49 \Delta\%$ und ein Rückschaufehler von $1.47 \Delta\%$ ⁸. Der Regressionseffekt ist damit mehr als doppelt so groß wie der Rückschaufehler.

Die Analyse der freien Wiedergabe zeigte ebenfalls signifikante Zusammenhänge zur Rückmeldung. Der Haupteffekt der Schätzung, daß mehr Argumente bei der Schätzung „gefallen“ erinnert wurden, läßt sich kaum sinnvoll interpretieren. Es ist kein Grund einsichtig, warum generell die Erinnerung bei einer Rückmeldung, laute sie nun „gefallen“ oder „gestiegen“, besser sein sollte. Auch die Interaktion von „Rückmeldung“ und „Argument“ war wenig aufschlußreich. Der Effekt beruhte auf einem Unterschied in der Bedingung „keine“ Rückmeldung. Daß ohne Rückmeldung Argumente, die für „fallen“ sprachen, besser erinnert wurden, als solche, die für „steigen“ sprachen, zeigt einen Behaltensvorteil für solche

⁸ Es ergibt sich der Regressionseffekt (RegrE) aus dem Mittel für $\Delta\%$ für die Schätzung (S) „gestiegen“ und „gefallen“: $(3.15 + 3.83)/2=3.49$. Der Rückschaufehler (RF) bei synonyme Rückmeldung (RM) ergibt sich aus der Differenz des RegrE minus dem Mittelwert von $\Delta\%$ für die S „gestiegen“ und „gefallen“: $(3.49-(2.60+2.98)/2=0.70)$. Der RF bei antonymer RM ergibt sich aus der Differenz des Mittelwerts von $\Delta\%$ für die S „gestiegen“ und „gefallen“ minus dem RegrE: $((2.60+2.98)/2-3.49=2.23)$. Der RF ergibt sich dann aus dem Mittel von „synonymer“ und „antonymer“ RM: $((0.70+2.23)/2=1.47)$.

Argumente, ist aber inhaltlich nicht weiter von Belang. Die Interaktion zwischen „Schätzung“ und „Argument“ hingegen zeigte deutlich einen selektiven Effekt der Schätzung: Lautete die Schätzung „gestiegen“, wurden auch mehr Argumente, die für „steigen“ sprachen, erinnert. Lautete die Schätzung „gefallen“, wurden ebenfalls mehr Argumente erinnert, die für „fallen“ sprachen. Die signifikanten Ergebnisse der Einzelvergleiche bestätigten die Interaktion in vollem Umfang. Die Schätzung wirkte also selektiv auf die Erinnerungsleistung. Dies ist ein erster Hinweis auf selektive Aktivierung für den Rückschaufehler. In der Interaktion steckte neben der Schätzung auch der Faktor Rückmeldung. Die Differenzierung zwischen der Entscheidung zum Zeitpunkt der Schätzung und dem Effekt der Rückmeldung kann erst in der dreifach Interaktion zum tragen kommen.

In der signifikant gewordenen dreifach Interaktion zwischen „Rückmeldung“, „Schätzung“ und „Argument“ zeigte sich eine differenzierte Wirkung der Rückmeldung auf die Erinnerung der Argumente. Bei Argumenten, die für „steigen“ sprachen, differierte die Erinnerungsleistung stärker zwischen der Schätzung „gestiegen“ und „gefallen“ und den Stufen des Faktors „Rückmeldung“ als bei Argumenten, die für „fallen“ sprachen. Bei „antonymer“ Rückmeldung wurden mehr Argumente bei der Schätzung „gefallen“ als „gestiegen“ erinnert, aber die erwarteten Unterschiede zwischen Argumenten, die für „steigen“ respektive „fallen“ sprechen, sind geringer als bei „synonymer“ Rückmeldung oder „keine“ Rückmeldung. Die Befunde sprechen dafür, daß sowohl die Entscheidung zum Zeitpunkt der Schätzung als auch die Rückmeldung eine selektive Wirkung auf die Erinnerungsleistung haben. Die Unterschiede in der Erinnerungsleistung bei Sachverhalten ohne Rückmeldung zeigen, daß die Entscheidung, die bei der Abgabe der Schätzung getroffen wurde, zu einer besseren Enkodierung solcher Argumente führte, die der eigenen Entscheidung entsprachen. Bei „synonymer“ Rückmeldung hätten die Unterschiede noch ausgeprägter sein müssen. Das Fehlen eines deutlicheren Unterschieds könnte so interpretiert werden, daß durch die Rückmeldung keine neuen Argumente mehr aktiviert wurden und nur die bereits bei der Abgabe der Schätzung aktivierten nochmals gestärkt wurden. Bei „antonymer“ Rückmeldung kam es dann zu einer ausgeglichenen Aktivierung von Argumenten, weil zum Zeitpunkt der Schätzung Argumente aktiviert wurden, die diese Entscheidung befürworteten und zum zweiten Zeitpunkt durch die gegensätzliche, „antonyme“ Rückmeldung Argumente aktiviert wurden, die gegen die erste Entscheidung sprachen. Es bleibt daher festzuhalten, daß die erste Schätzung eine deutliche selektive Wirkung auf die Erinnerungsleistung der Argumente ausgeübt hat, die durch die Rückmeldung teilweise noch verstärkt wurde. Im Falle einer antonymen Rückmeldung wurden durch die Entscheidung

zum Zeitpunkt der Abgabe der Schätzung und später bei der Rückmeldung gleichmäßig die Stichworte zu den Argumente (*Images*) des Image-Sets aktiviert.

Die Wirkung „antonymer“ Rückmeldungen kann auch aus Befunden zum „counterfactual reasoning“ abgeleitet werden (Arkes et al. 1988; Davies, 1992; Hoch, 1985; Koriati et al., 1980; Lord, Lepper & Preston, 1984, siehe auch Kap. 2.2). Die systematische Wirkung einer Lösung konnte verringert oder gar neutralisiert werden, wenn über ein zur Lösung gegenteiliges Ergebnis nachgedacht wurde. Dies entspricht dem Fall zweier symmetrischer Lösungen, deren Wirkung sich gegenseitig aufheben.

Dennoch läßt sich festhalten, daß sich in der Interaktion zwischen „Schätzung“ und „Argument“ deutlich eine selektive Aktivierung von Argumenten gezeigt hat. Damit stellt dies vierte Experiment den ersten Nachweis *selektiver Aktivierung* dar.

6.2 Experiment 5: Freie Wiedergabe der Argumente im Urteilsdesign

EINLEITUNG

Dieses fünfte Experiment verwendete ebenfalls eine freie Wiedergabe zum Nachweis *selektiver Aktivierung*. Das Design wechselte in diesem Experiment vom Gedächtnisdesign zum Urteilsdesign. Eine selektive Wirkung der vorgegebenen Lösungen auf Argumente würde neben einer Bestätigung der Befunde aus dem vorangegangenen Experiment auch eine Erweiterung des Konzepts selektive Aktivierung darstellen. Wie bereits erwähnt (siehe Kap. 3), wäre es durchaus denkbar, daß sich die selektive Wirkung eines Ankers auf zweierlei Weise einstellt. Zum einen durch selektive Aktivierung von Wissens-elementen bei der Enkodierung des Ankers, zum anderen dadurch, daß der vorliegende (oder erfolgreich abgerufene) Anker als zusätzlicher Abrufreiz von Wissens-elementen aus dem Langzeitgedächtnis wirkt. Bei einem Experiment im Urteilsdesign ist davon auszugehen, daß die Wirkung des Ankers als Abrufreiz für Informationen aus dem Langzeitgedächtnis eine größere Rolle spielt als bei einem Experiment im Gedächtnisdesign, wo der Anker zum Zeitpunkt der Erinnerung nicht mehr vorliegt. Es ergibt sich daraus, daß der Rückschafehler größer sein sollte, wenn der Anker zum Erinnerungszeitpunkt vorliegt als wenn er auch erinnert werden muß. Im Einklang mit dieser Vorhersage fällt der Rückschafehler im Urteilsdesign auch größer aus als im Gedächtnisdesign (Campbell & Tesser, 1983; Davies, 1992; Fischhoff, 1977; Powell, 1988; Wood, 1978, Exp. 2; siehe auch Kap. 2.1). Die Anzahl zu bearbeitender Aufgaben könnte eine zusätzliche Rolle spielen. Wie im vorherigem Experiment 4 sollte jeweils in einem nachfolgenden Test die Erinnerung an die präsentierten Argumente in einer freien Wiedergabe getestet werden. Von besonderem Interesse war hierbei auch die Menge zu bearbeitender Sachverhalte. Es nahmen zwei Gruppen mit 9 beziehungsweise 18 Sachverhalten an dem Experiment teil. Denkbar wäre, daß mit zu erwartender schwächerer Erinnerungsleistung bei der doppelten Menge an zu erinnernden Sachverhalten und Argumenten, auch eine stärkere Tendenz einher gehen könnte die vorgegebene Lösung als Abruffilter zu nutzen. Möglicherweise moderiert die Erinnerungsleistung auch den Grad an Selektivität.

METHODE

Design Der Versuchsplan bestand in einem 2x3x2 Design mit dem Gruppenfaktor „Menge“ (9 Sachverhalte/18 Sachverhalte) und den Meßwiederholungsfaktoren „Lösung“ (gestiegen/gefallen/keine) und „Argument“ (fallen/steigen) (siehe Abb.

6.9). Probanden wurden randomisiert den Gruppen mit 9 bzw. 18 Sachverhalten des Gruppenfaktors „Menge“ zugewiesen. Die Zuordnung zu dem Faktor „Lösung“ wurde durch systematische Permutation einer einmaligen Zufallsreihenfolge realisiert, so daß alle Fragen über alle Bedingungen gleich häufig verwirklicht wurden. Eine mit „gestiegen“ bezeichnete Lösung zeigte an, daß der zu beurteilende Sachverhalt in Wirklichkeit gestiegen wäre. Entsprechend zeigte eine mit „gefallen“ bezeichnete Lösung an, daß der zu beurteilende Sachverhalt in Wirklichkeit gefallen wäre. Die Bedingung „keine“ Lösung diente als Kontrollbedingung. Der Faktor „Argument“ gruppiert die Wiedererkennensleistung der Argumente in Argumente, die dafür sprachen, daß der entsprechende Sachverhalt gestiegen („steigen“) und Argumente, die dafür sprachen, daß er gefallen („fallen“) war. Die Argumente sprachen jeweils zur Hälfte dafür, daß der entsprechende Sachverhalt gestiegen respektive gefallen war. Die Darbietung erfolgte blockweise, wobei die Reihenfolge der Argumente als auch der Blöcke selbst einer Zufallsreihenfolge folgte, so daß alle Argumente gleich häufig an jeder Position präsentiert wurden.

Als abhängige Maße wurden neben den Konfidenzurteilen bezüglich der Sicherheit, daß der zu beurteilende Sachverhalt gestiegen wäre, die Anzahl erinnerter Argumente in einem Wiedererkennenstest erhoben.

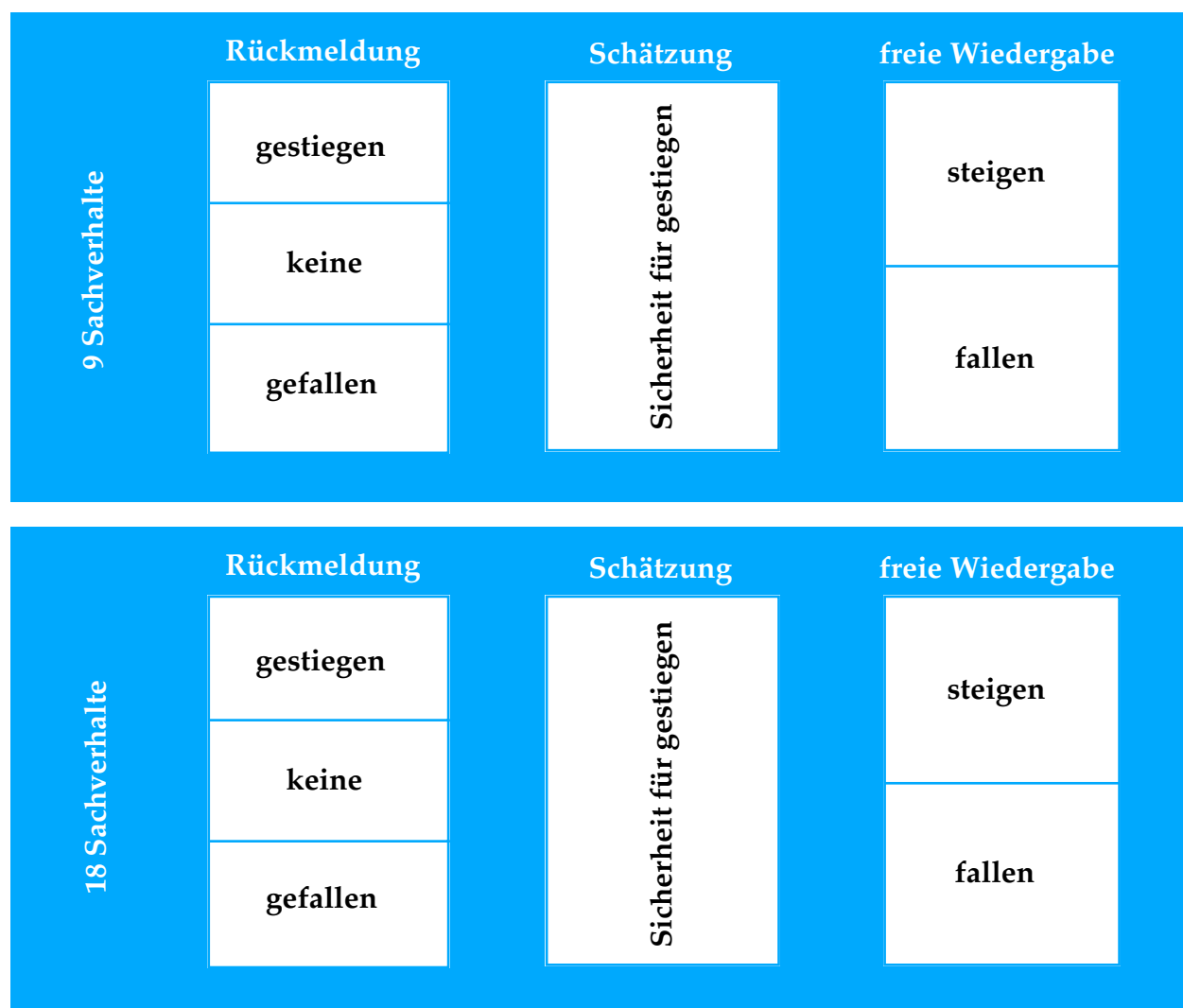


Abbildung 6.9: Versuchsplan von Experiment 5. Urteilsdesign in 2 Gruppen mit jeweils 9 bzw. 18 Sachverhalten und Meßwiederholung auf den Faktoren „Lösung“ und „Argumente“.

Probanden An der Untersuchung nahmen 104 Personen, vorwiegend Studierende aus verschiedenen Fachbereichen der Universität Trier, teil. Davon waren 69 weiblichen und 35 männlichen Geschlechts. Das Altersspektrum reichte von 18 bis 40 Jahren. Der Altersdurchschnitt lag bei 22.9 Jahren. Die Probanden erhielten 12 DM für die Teilnahme am Experiment.

Material Je nach Gruppe gab es 9 oder 18 Verifikationsaufgaben, die danach beurteilt werden sollten, wie sicher die Probanden jeweils waren, ob ein betreffender Sachverhalt gestiegen oder gefallen war. Die Beurteilung der Sachverhalte erfolgte jeweils durch Verschiebung einer Markierung auf einer analogen Skala auf dem Computerbildschirm, die von „gefallen“ zu „gestiegen“ reichte. Die Probanden bekamen den konkreten numerischen Skalenwert nicht angezeigt.

Um die Präsentation übersichtlicher zu gestalten und weil das Platzangebot auf dem Bildschirm begrenzt war, wurden die Argumente auf dem Bildschirm in Stichworten zusammengefaßt. Zunächst wurde den Probanden jedes Argument vollständig ausformuliert zusammen mit einem Stichwort dargeboten. Die Probanden sollten sich das Argument durchlesen und anschließend die Maustaste betätigen, damit das nächste Argument auf dem Bildschirm erscheinen konnte. Nachdem alle Argumente präsentiert worden waren, blieben auf dem Bildschirm, in der Reihenfolge ihrer Darbietung, nur die Stichwörter zu den Argumenten sichtbar. Die Probanden konnten aber jederzeit durch Klicken auf ein Stichwort das entsprechende Argument nochmals lesen.

Ablauf Die Durchführung folgte dem Urteilsdesign und fand in einer Sitzung statt. Der gesamte Versuch erfolgte am Rechner. Es bestand jeweils für vier Probanden gleichzeitig die Möglichkeit teilzunehmen.

Die Probanden wurden zunächst in einer ausführlichen Instruktion am Bildschirm aufgeklärt, daß sie im folgenden, je nach Gruppenzuordnung, 9 oder 18 Sachverhalte danach beurteilen sollten, ob der betreffende Sachverhalt gestiegen oder gefallen wäre. Die Probanden wurden informiert, daß zu einigen Sachverhalten die Lösungen dargeboten würden. Sie sollten aber trotzdem in jedem Fall versuchen, ein unabhängiges Urteil abzugeben, als kennten sie diese Information nicht. Sie wurden informiert, daß sie ihre Antwort in einer analogen Skala am unteren Bildschirmrand eintragen sollten, wobei der jeweilige Abstand zu den Extrempunkten der Skala, die mit „gestiegen“ bzw. „gefallen“ bezeichnet waren, die subjektive Sicherheit ihres Urteils anzeigen sollte. Eine Antwort nahe einem Extrempunkt drücke hohe Sicherheit, wohingegen eine Antwort in der Nähe der Mitte der Skala Unsicherheit ausdrücke. Sie wurden informiert, daß sie zu jedem Sachverhalt acht Argumente präsentiert bekämen. Sie wurden angehalten, diese sorgfältig zu lesen, da sie für den weiteren Verlauf des Experimentes sehr wichtig wären und ihnen ihre Entscheidung erleichtern sollten. Die Probanden wurden darüber informiert, daß aus Platzgründen auf dem Bildschirm jeweils nur noch Stichwörter zu den Argumenten erscheinen würden, sie aber jeweils durch Klicken auf ein entsprechendes Stichwort die Gelegenheit hätten, das zugehörige Argument nochmals zu sehen. Vor der Präsentation der ersten Frage wurde ein Übungsdurchgang (mit nicht ausgewerteten Sachverhalten) durchgeführt.

Zu Beginn eines Versuchsdurchgangs erschien am oberen Teil des Bildschirms jeweils ein Sachverhalt. Anschließend wurden nacheinander in zufälliger Reihenfolge, jeweils einzeln, die acht Argumente dargeboten. Es erschien am

Bildschirm jeweils das Argument, gefolgt von einem textlich abgehobenen Stichwort. Dann erschien ein kleines Dialogfenster mit der Aufforderung, sich das Argument und Stichwort gut durchzulesen, zu merken und dann auf „OK“ zu klicken, damit das nächste Argument präsentiert werden könnte. Nach dem letzten Argument waren der Sachverhalt und, falls es sich um einen Sachverhalt mit Lösung handelte, eine Mitteilung, ob der Sachverhalt in Wirklichkeit gefallen respektive gestiegen wäre und in der Reihenfolge ihres Erscheinens die Stichworte zu den Argumenten zu sehen. Den Probanden war es freigestellt, durch Klicken auf ein Stichwort nochmals das ausformulierte Argument zu sehen. Unterhalb der Stichworte wies im Fall einer Lösung ein Dialogfeld darauf hin, daß die Probanden ihr Urteil unabhängig von der präsentierten Lösung fällen sollten. Auf der analogen Skala am unteren Bildschirmrand, deren jeweilige Endpunkte mit „gestiegen“ und „gefallen“ gekennzeichnet waren, konnten die Probanden anschließend ihre Sicherheit angeben, mit der der präsentierte Sachverhalt gestiegen, respektive gefallen war (siehe Abb. 6.10).

Durch Klicken auf die Skala wurde an dieser Stelle ein Schieberegler sichtbar, der mit der Maus verschoben werden konnte. Gleichzeitig erschien eine Taste auf dem Bildschirm mit der Aufschrift „weiter“, mit der die Probanden zum nächsten Sachverhalt gelangen konnten. Es bestand keinerlei Zeitlimit. Die Probanden konnten solange bei einem Sachverhalt verweilen, wie es ihnen beliebte. Dieser Teil des Experimentes dauerte je nach Gruppe zwischen 15 oder 30 Minuten pro Proband.

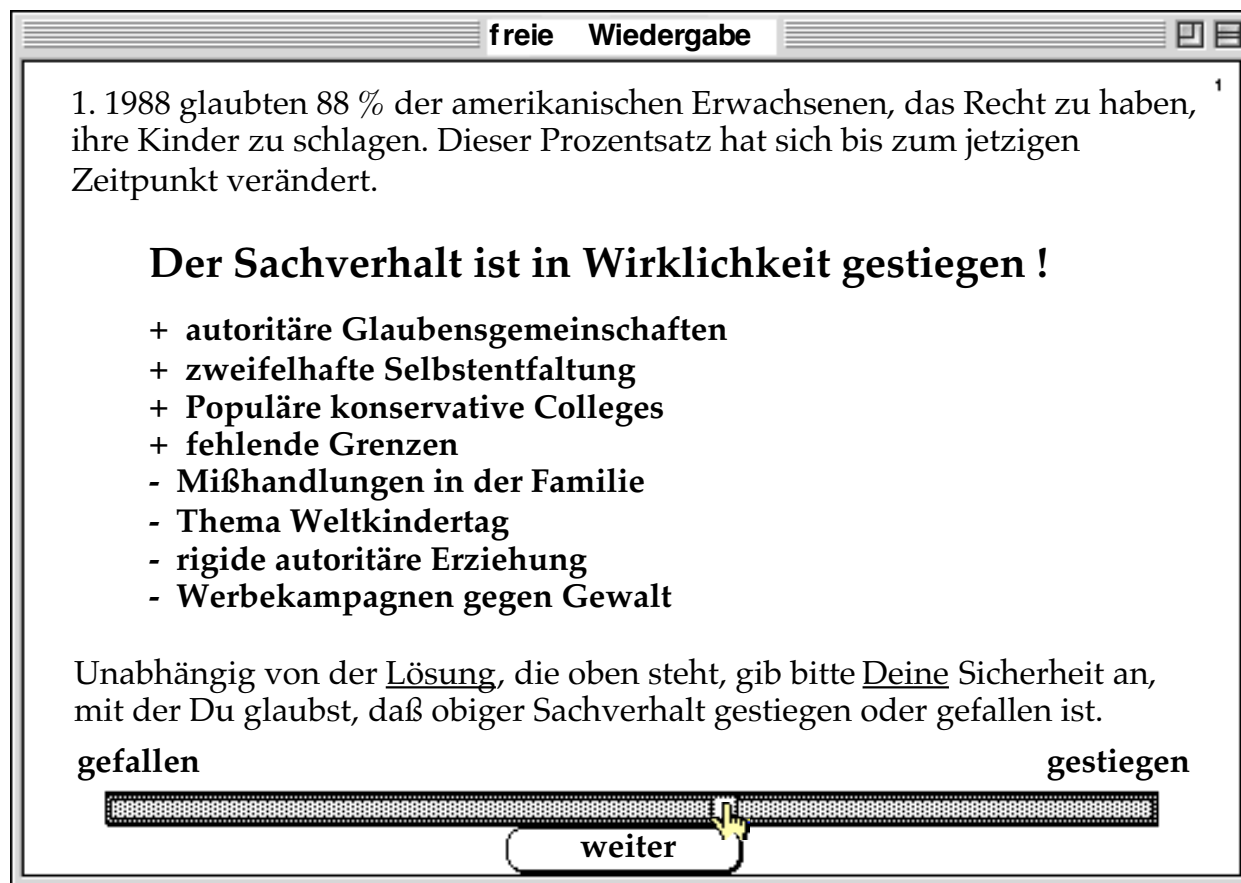


Abbildung 6.10: Bildschirm in Experiment 5 mit den Stichworten zu den Argumenten und der Skala zur Einschätzung der Konfidenz, ob der Sachverhalt gestiegen oder gefallen ist.

Hatten die Probanden zu allen Sachverhalten ihr Urteil abgegeben, erschien eine zweite Instruktion, die sie informierte, daß ihnen im folgenden nochmals alle Sachverhalte präsentiert würden. Ihre Aufgabe wäre es, möglichst schnell viele der präsentierten Argumente zu erinnern. Darauf erschien dann der erste Sachverhalt zusammen mit der Lösung und einem Dialogfenster, welches sie aufforderte, das erste Argument anzugeben, was ihnen einfiel. Hatten die Probanden ein Argument eingetippt, erschien ein weiteres Dialogfenster mit der Aufforderung zu entscheiden, ob das gerade eingetippte Argument dafür spräche, daß der Sachverhalt steige oder eher dafür, daß er falle. Zur Entscheidung dieser Frage konnten die Probanden einen entsprechend bezeichneten Button anklicken. Hatten die Probanden nach Durchlaufen des Übungsbeispiels keinerlei Fragen mehr, gelangten sie zum ersten Sachverhalt. Die Reihenfolge der Sachverhalte und der Bildschirmaufbau wurde beibehalten. Dieser Teil des Experimentes dauerte je nach Gruppe zwischen 20 und 40 Minuten pro Proband.

Auswertung Ein Proband mußte aus der Analyse ausgeschlossen werden, da er den zweiten Teil des Experiments nicht absolviert hat. Insgesamt nahmen also 103

Probanden am Experiment teil, wobei 51 Probanden auf die Gruppe mit 9 Sachverhalten und 52 auf die Gruppe mit 18 Sachverhalten fielen.

Die Zuordnung der Argumente in die entsprechenden Kategorien in Argumente, die für „steigen“ respektive für „fallen“ sprachen, geschah durch die Zuordnung der Probanden selbst. Eine Überprüfung der zugeordneten Argumente durch zwei Rater entdeckte nur in zwei Fällen Unstimmigkeiten. Hier hatte ein Proband alte Argumente offensichtlich vertauscht. Die Argumente wurden aber dennoch in den Kategorien belassen, da die Probanden nach jedem Argument explizit dazu aufgefordert worden waren, dieses in die Kategorie „spricht dafür, daß der Sachverhalt gestiegen ...“ respektive „... gefallen ist“, einzuordnen. Neben alten, mit den Sachverhalten präsentierten, traten selbstverständlich auch hier neue, erfundene Argumente auf, die ebenfalls in die Kategorien „steigen“ und „fallen“ eingeordnet wurden. Der Anteil neuer Argumente lag jedoch nur bei 2.69% („steigen“: 99.0 % alte und 1.0% neue Argumente; „fallen“: 98.3% alte und 1.7% neue Argumente). Es konnte daher auf getrennte Analysen mit alten und neuen Argumenten verzichtet werden.

ERGEBNISSE

Verteilung der Daten Die Verteilung von Schätzungen getrennt nach der Lösung (gestiegen/gefallen/keine) zeigt, daß in Abhängigkeit von der Lösung die Mittelwerte in Richtung der Lösung wanderten: Bei der Lösung „gestiegen“ lag der Mittelwert der Sicherheit dafür, daß der Sachverhalt gestiegen wäre, bei 64.16% und bei der Lösung „gefallen“ bei 39.85%, wohingegen der Mittelwert der Verteilung mit „keiner“ Lösung bei 55.07% lag (siehe Abb. 6.11). Die Standardabweichungen zwischen den Lösungen unterschieden sich nicht wesentlich voneinander (gestiegen: 16.52; gefallen: 13.83; keine: 15.53).

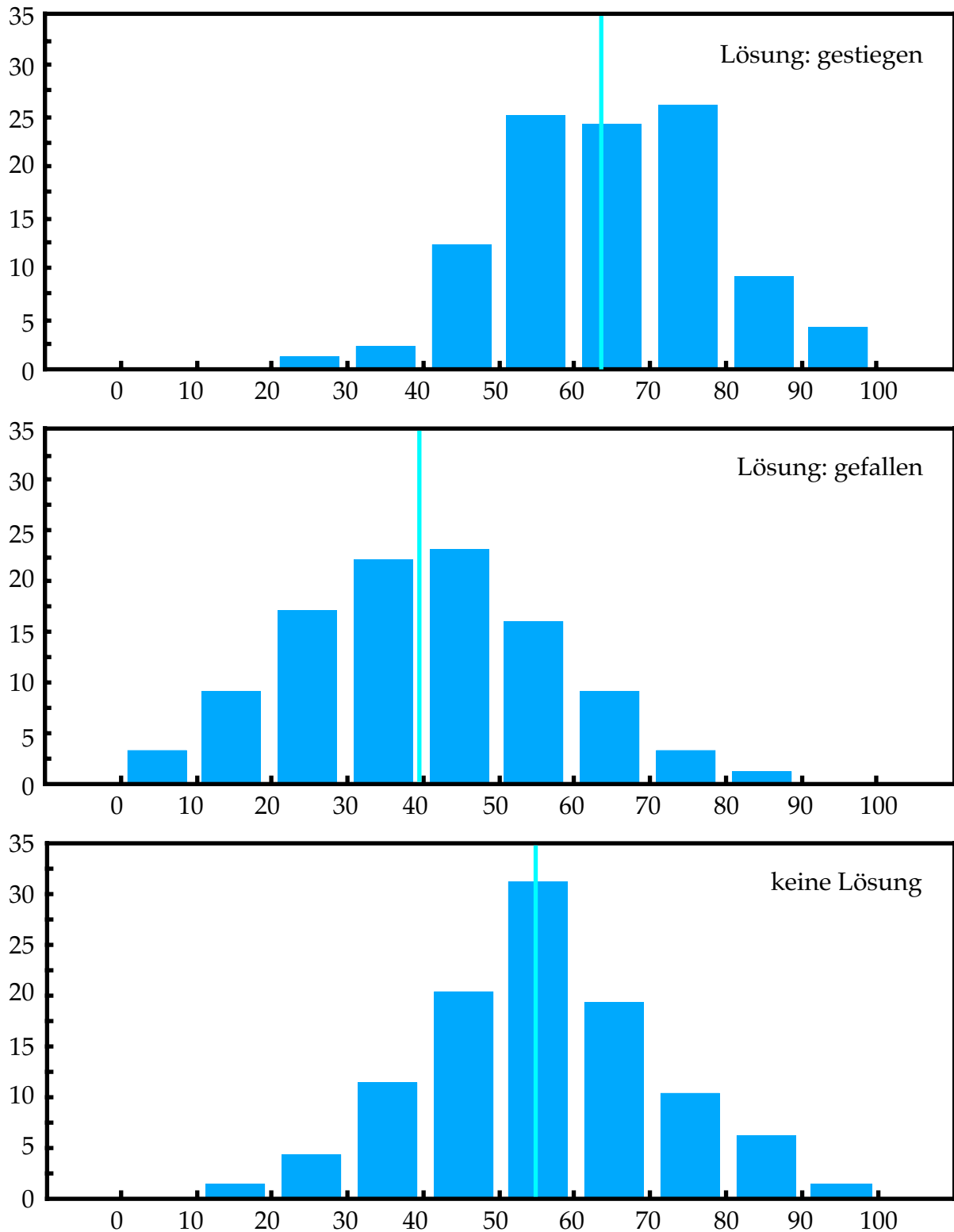


Abbildung 6.11: Relative Häufigkeiten für Schätzungen pro Zehner-Kategorie im Urteilsdesign.

Konfidenzurteile Die Analyse der Unterschiede in den Konfidenzurteilen in Abhängigkeit von der Lösung erlaubt die Beurteilung, ob ein Rückschaufehler vorliegt. Da es sich hier um ein Experiment im Urteilsdesign handelte und somit nur

ein Meßzeitpunkt vorlag, konnte für diese Analyse der Wert des Konfidenzurteils direkt ohne weitere Transformation herangezogen werden.

Bei einer zweifaktoriellen Varianzanalyse mit dem Gruppenfaktor „Menge“ (9 Sachverhalte/18 Sachverhalte) und dem Meßwiederholungsfaktor „Lösung“ (gefallen/gestiegen/keine) über die Sicherheit, daß der Sachverhalt gestiegen wäre, wurde der Gruppenfaktor „Menge“ ($F_{(1,101)}=.65$) nicht signifikant. Es zeigte sich jedoch ein Haupteffekt der „Lösung“ ($F_{(2,202)}=58.55$; Power (1-beta)=0.95): Ein Tukey Test zeigte für die Lösung „gestiegen“ (64.16%) eine höhere Sicherheit dafür, daß der Sachverhalt gestiegen war, als bei der Lösung „gefallen“ (39.85%) Die Sachverhalte mit „keiner“ Lösung (55.07%) lagen nahe der 50%-Marke (gestiegen/gefallen: $p=.00002$; gestiegen/keine: $p=.0002$; gefallen/keine: $p=.00002$). Die Interaktion zwischen „Menge“ und „Lösung“ wurde nicht signifikant ($F_{(2,202)}=.38$).

Freie Wiedergabe Die Analyse der Anzahl erinnerter Argumente in einer freien Wiedergabe erlaubt die Beurteilung eines direkten Effekts auf die Güte der Erinnerung durch die experimentelle Manipulation. Eine dreifaktorielle Varianzanalyse mit dem Gruppenfaktor „Menge“ (9 Sachverhalte/18 Sachverhalte) und den Meßwiederholungsfaktoren „Lösung“ (gefallen/gestiegen/keine) und „Argument“ (fallen/steigen) über die Anzahl erinnerter Argumente in einer freien Wiedergabe zeigte einen signifikanten Effekt des Gruppenfaktors „Menge“ ($F_{(1,101)}=17.78$; Power (1-beta)=0.97): In der Gruppe mit 8 Sachverhalten wurden mehr Argumente (2.21) erinnert als in der Gruppe mit 18 Sachverhalten (1.63). Weder der Meßwiederholungsfaktor „Lösung“ ($F_{(2,202)}=2.87$) noch der Faktor „Argument“ ($F_{(1,101)}=1.12$) zeigten signifikante Unterschiede in der Anzahl erinnerter Argumente.

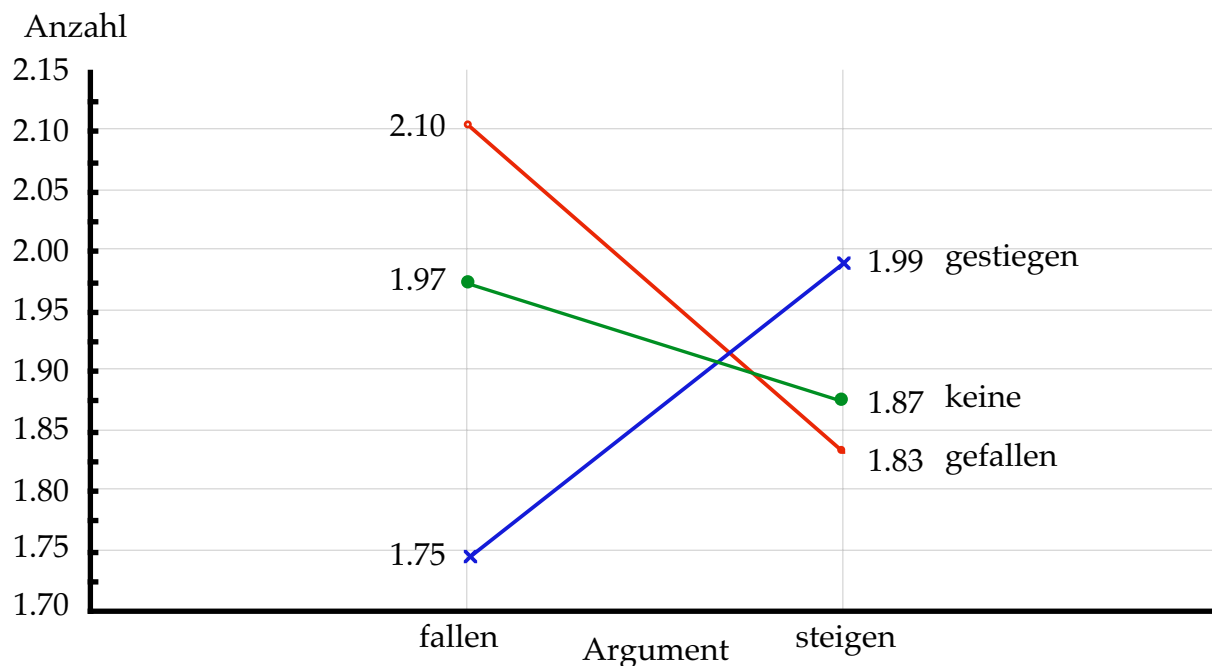


Abbildung 6.12: Interaktion Lösung (gestiegen/ gefallen/ keine) mit Argument (fallen/ steigen) über die Anzahl erinnerter Argumente in der freien Wiedergabe.

Die Interaktion zwischen „Lösung“ und „Argument“ ($F_{(2,202)}=12.40$; Power (1-beta)=0.95) wurde signifikant (siehe Abb. 6.12): Bei der Lösung „gestiegen“ wurden signifikant mehr Argumente, die für „steigen“ (1.99) sprachen, als für „fallen“ (1.75) erinnert ($t_{102}=3.48$, $p=.0007$). Bei der Lösung „gefallen“ wurden signifikant mehr Argumente, die für „fallen“ (2.10) sprachen, als für „steigen“ (1.83) erinnert ($t_{102}=3.48$, $p=.0007$). Argumente, die für „fallen“ sprachen, wurden öfter bei der Lösung „gefallen“ (.47) als bei der Lösung „gestiegen“ (.33) erinnert ($F_{(2,202)}=14.74$, $p=.000001$): Ein Tukey Test zeigte, daß mit der Lösung „gestiegen“ (1.75) weniger Argumente für „fallen“ erinnert wurden als bei „keiner“ (1.97) oder bei der Lösung „gefallen“ (2.10) (gestiegen/gefallen: $p=.00002$; gestiegen/keine: $p=.002$; gefallen/keine: $p=.13$). Die restlichen Einzelvergleiche wurden nicht signifikant (alle $F<3.02$; $t<1.35$). Die anderen zweifachen Interaktionen sowie die dreifach Interaktion wurden nicht signifikant (alle $F<1.64$).

DISKUSSION

Die Verteilung der Daten zeigte Verschiebungen in Richtung der vorgegebenen Lösungen. Die Lösung „gestiegen“ führte zu einer Konfidenz für „gestiegen“ von 64.16%, während bei „keine“ Lösung die Konfidenz bei 55.07% und mit der Lösung „gefallen“ bei nur 39.85% lag. Damit zeigte sich schon auf dieser deskriptiven Ebene ein systematischer Einfluß der vorgegebenen Lösungen, der auf einen Rückschaufehler hindeutete.

Der signifikante Haupteffekt der „Lösung“ in den Konfidenzurteilen zeigte eine deutliche systematische Verschiebung der Konfidenz in Richtung der Lösung und damit einen Rückschafehler. Die signifikanten Unterschiede in den Konfidenzurteilen in Abhängigkeit von der Lösung zeigten wie in Experiment 4 einen Rückschafehler: Die Lösung, der Sachverhalt wäre „gestiegen“, erhöhte die Sicherheit für gestiegen, die Lösung, der Sachverhalt wäre „gefallen“, verminderte hingegen die Sicherheit für gestiegen. Ohne Lösung ergab sich hier eine Konfidenz nahe der 50%-Marke. Einen Effekt des Faktors „Menge“ auf die Konfidenzen war aus inhaltlichen Gründen weder als Haupteffekt noch in der Interaktion zu erwarten. Wie schon im vorangegangenen vierten Experiment, zeigte sich in diesem fünften Experiment ein deutlicher Rückschafehler.

Die Analyse der freien Wiedergabe zeigte ebenfalls signifikante Zusammenhänge zur Lösung. Der Haupteffekt des Faktors „Menge“ mit besserer Erinnerungsleistung bei geringerer Aufgabenmenge, erstaunt wenig. Es ist ein gut gesicherter Befund, daß die Erinnerungsleistung sich umgekehrt proportional zur Anzahl zu bearbeitender Aufgaben verhält (vgl. Ebbinghaus, 1885). Die Interaktion zwischen „Lösung“ und „Argument“ zeigte, wie auch schon im vierten Experiment, einen deutlichen selektiven Effekt der Lösung: Lautete die Lösung „gestiegen“, wurden auch mehr Argumente, die für „steigen“ sprachen, erinnert. Lautete die Lösung „gefallen“, wurden ebenfalls mehr Argumente erinnert, die für „fallen“ sprachen. Die signifikanten Ergebnisse der Einzelvergleiche bestätigten die Interaktion fast in vollem Umfang. Lediglich der Einzelvergleich über Argumente, die für „steigen“ sprachen, wurde nicht signifikant. Die Lösung wirkte also selektiv auf die Erinnerungsleistung. Damit wurde in diesem fünften Experiment ein weiterer Nachweis *selektiver Aktivierung* für den Rückschafehler erbracht.

Eine weitere Varianzanalyse, die nur die Entscheidung der Probanden entgegen der vorgegebenen Lösung mitberücksichtigte fand keine signifikanten Unterschiede zwischen der Entscheidung und der Anzahl erinnerter Argumente (*alle* $F < 1.59$). Das Lösungsmuster sah aber vollkommen verschieden aus. Es wurden passend zur Entscheidung und entgegen der Rückmeldung mehr Argumente erinnert: lautete die Entscheidung „gestiegen“ wurden mehr Argumente für „steigen“ erinnert („steigen“ 1.72; „fallen“ 1.57). Lautete die Entscheidung „gefallen“ wurden mehr Argumente für „fallen“ erinnert („steigen“ 1.70; „fallen“ 1.53). Diese Analyse gibt einen Hinweis, daß die Entscheidung auch unabhängig von der Rückmeldung einen Effekt auf die Erinnerung der Argumente hatte. Es wurden passend zur Entscheidung mehr Argumente erinnert. Wirkt der Anker nur als selektiver Cue sollte dieser Effekt der Entscheidung auf die Erinnerungsleistung der Argumente

ausbleiben und nur der Effekt der vorgegebenen Lösung zu sehen sein. Dieses Befundmuster spricht daher sehr für einen selektiven Abruf der Argumente in Abhängigkeit von der getroffenen Entscheidung. Die vorgegebene Lösung löst daher nicht automatisch einen selektiven Abruf der Argumente aus, sondern erst die gewählte Entscheidung sorgt für die selektive Erinnerungsleistung. Da in diesem Experiment die meisten Entscheidungen konform mit der vorgegebenen Lösung waren, fand sich auch ein Nachweis selektiver Erinnerungsleistung in Abhängigkeit von der Lösungsvorgabe. Der Effekt der Lösung wäre vermutlich noch stärker ausgefallen, wenn zum Zeitpunkt der Erinnerung die Lösungen nochmals präsentiert worden wären und so als Abrufcues fungiert hätten. Das hätte eventuell in manchen Fällen, in denen die tatsächlich getroffene Entscheidung entgegen der Lösung gelaftet hatte, dafür gesorgt, daß dennoch mehr Argumente erinnert worden wären, die zur Lösung passen.

Auch in diesem fünften Experiment ließ sich also ein selektiver Effekt auf die Erinnerungsleistung mittels freier Wiedergabe erzielen. Abhängig von der vorgegebenen Lösung ließen sich mehr Argumente erinnern, die die Lösung favorisierten. Damit erweist sich die selektive Aktivierung solcher Elemente des Wissens, die dem präsentierten Anker ähnlich sind, als die zentrale Erklärung für den Rückschaufehler: Dargebotene Informationen (z. B. die Lösung) entfalten eine aktivierende Wirkung auf ihnen „ähnliche“ Elemente im Gedächtnis, die zur Generierung der Schätzung herangezogen wurden. Dies geschieht durch eine automatische Aktivationsausbreitung im semantischen Netzwerk (Anderson, 1983). Informationen, wie die Lösung (Anker) aktivieren, durch ihre semantische Ähnlichkeit Informationseinheiten (Images) aus dem Image-Set. Die stärker aktivierten Informationen sind in einer anschließenden Gedächtnissuche leichter zugänglich (McNamara, 1992), was zu einem vermehrten Abruf dieser Gedächtnisinhalte und als Konsequenz letztlich zu einer Verschiebung der Antwort in Richtung der dargebotenen Information führt (siehe Kap. 3.2 und Kap. 4)

Fazit

7.1 Abschließende Diskussion

Die Befunde aus den Experimenten lassen sich folgendermaßen zusammenfassen: Es fand sich ein Rückschaufehler in drei von fünf Experimenten (Experimente 3-5). Die Konfidenz für den rückgemeldeten Ausgang nahm zu. In Experiment 3 und 4, mit der Variation der Rückmeldung, „synonyme“ und „antonyme“ Rückmeldung, findet sich der Effekt lediglich bei „antonymer“ Rückmeldung. Die möglichen Gründe, weshalb in den ersten beiden Experimenten kein Einfluß der Rückmeldungen auf die Höhe der Konfidenz, sowie warum sich dieser Einfluß in den Experimenten 3 und 4 nur bei „antonymer“ Rückmeldung nachweisen ließ, sind jeweils schon in den Diskussionen zu den Experimenten besprochen worden. Es hatte zur Konsequenz, daß die Probanden in Experiment 2 dazu angehalten wurden, ihre Entscheidungen schneller zu treffen. In Experiment 3 wechselte der Inhalt der Sachverhalte, von neutralen sehr ausgewogenen, aber vielleicht auch etwas langweiligen Sachverhalten, hin zu Sachverhalten zu denen möglichst schon eine präexperimentelle Meinung vorhanden war. In Experiment 4 wurde das Experiment nicht mehr am Computer, sondern schriftlich, in Form eines Fragebogens durchgeführt. Schließlich, in Experiment 5, wechselte das Untersuchungsdesign von einem Gedächtnis- zu einem Urteilsdesign. Auf den Rückschaufehler wirkte sich merklich der Wechsel, weg von neutralen Sachverhalten, aus. Den stärksten Effekt auf den Rückschaufehler hatte aber der Wechsel vom Gedächtnis- zum Urteilsdesign. In den Experimenten 4 und 5 ließ sich eine systematische Beeinflussung der Erinnerung für Argumente in freier Wiedergabe nachweisen, abhängig von der mit der Schätzung getroffenen Wahl und der Rückmeldung. Diese Befunde lassen sich als ersten Nachweis *selektiver Aktivierung* interpretieren. In den Experimenten mit Wiedererkennenstest (Experiment 1 – 3) ließ sich kein selektiver Einfluß auf die Wiedererkennensleitung alter, durch den Anker favorisierter Argumente nachweisen.

Die zentrale Frage mit der sich die folgende Diskussion beschäftigen wird ist, warum hier in zwei von fünf (streng genommen in vier von fünf) Experimenten gelungen ist, was sonst - folgt man Christensen-Szalanski und Willham (1991; siehe Kap. 2.2) - bisher nur in 6 von 126 Arbeiten auftrat: Eine Unterdrückung des Rückschaufehlers.

7.1.1 Material

Konfidenzskala Einen wesentlichen Aspekt dieses Phänomens stellt das in allen Experimenten verwendete Material dar. Durch die gewählte Aufgabe, zu schätzen, ob ein bestimmter Sachverhalt gestiegen oder gefallen ist, und einer Konfidenzskala von 0 - 100%, bewegt man sich von offenen Almanachfragen hin zu Verifikationsaufgaben, da hier die Endpunkte der Skala zugleich auch die Anker sind (z.B. Fischhoff, 1977; siehe Kap. 2.2). Die Wahrscheinlichkeit für das Auftreten eines Rückschuffehlers ist damit aber auch abhängig von der Lage der Schätzungen. Liegen die Probanden mit ihrer Schätzung in der Nähe des Ankers, so gibt es einen geringeren Spielraum für eine Annäherung der Erinnerungen an den Anker. Liegen sie dagegen mit ihrer Schätzung jenseits der 50%-Marke der Skala, gibt es mehr Möglichkeiten, sich in der Erinnerung dem Anker anzunähern, als sich von ihm zu entfernen. Das hat zur Konsequenz, daß in Experiment 3 und 4 systematisch eine „synonyme“ Rückmeldung weniger und eine „antonyme“ Rückmeldung dagegen mehr Spielraum für die Annäherung der Erinnerung an den Anker ließ. Bedenkt man noch, daß weiter entfernte Anker eine stärkere absolute Verfälschung⁹ hervorrufen (Kohnert, 1996; Pohl, 1998), so sprechen diese Überlegungen für einen stärkeren (absoluten) Effekt des Ankers bei „antonymer“ Rückmeldung.

Kausalstruktur Ein weiterer Aspekt ist, daß zwar - ähnlich wie bei Almanachfragen - in allen hier vorgestellten Experimenten viele unterschiedliche Sachverhalte zu beurteilen waren, die Struktur der Sachverhalte jedoch sehr von einer normalen Almanachfrage abwich. Die hier verwendeten Sachverhalte mit jeweils einer Anzahl von Argumenten für zwei mögliche Ausgänge besaßen alle eine Kausalstruktur: Der Sachverhalt konnte gestiegen sein, weil beispielsweise die Argumente eins bis vier diesen (hypothetischen) Ausgang stützten, oder der Sachverhalt konnte gefallen sein, weil die Argumente fünf bis acht diesen (hypothetischen) Ausgang stützten. In jedem Fall waren zu jedem Sachverhalt zwei mögliche Ausgänge mit ihren zugehörigen Argumenten vorgegeben. Eine derartige Kausalstruktur ist aber eher ein Merkmal episodischen Materials. In den Experimenten wurden daher keine singulären Fakten, sondern mit Argumenten verknüpfte Sachverhalte dargeboten. Wie bei episodischen Material ist also auch hier eine wesentlich elaboriertere Enkodierung zu erwarten (Fischhoff, 1975; Wasserman et al., 1991; siehe Kap. 2.2).

⁹ Der Effekt eines von der Schätzung weiter entfernten Ankers ist absolut größer, aber relativ gesehen kleiner: Relativiert man die absoluten Distanz der Erinnerung vom Anker an der Distanz Schätzung zu Anker (Hell et al., 1988; Kohnert, 1996; Pohl, 1993), wirkt beispielsweise der absolut kleinere Effekt einer Erinnerung von 80 Jahren bei einem Anker von 82 Jahren und einer Schätzung von 78 Jahren, stärker, als der absolut stärkere Effekt einer Erinnerung von 70 Jahren, bei einer Schätzung von 64 Jahren (absoluter Effekt: 2 zu 6 Jahren; relativer Effekt: 50% zu 33%)!

Gemeint ist hier, daß in den Sachverhalten, durch die gestellte Aufgabe, zu beurteilen, ob der Sachverhalt gestiegen oder gefallen sei, zwangsläufig eine mehr oder minder tiefe Elaboration der präsentierten Argumente stattfinden mußte. Es war schließlich die Funktion der Argumente, von den Probanden zur Lösung dieser Aufgabe herangezogen zu werden. Die Befunde zu episodischem Material lassen dann auch hier kleinere Effekte als bei Almanachfragen erwarten. Christensen-Szalanski und Willham (1991) berichten eine Effektgröße für Episoden, die nur ein Drittel derjenigen bei Almanachfragen beträgt. Die Diskussion zu den prinzipiellen Gründen, warum dieser Effekt gegenüber Almanachfragen soviel geringer ist, trifft daher auch auf die hier in allen Experimenten als Material verwendeten Sachverhalte zu.

Auch hier gelten die Argumente von Fischhoff (1975) und von Wasserman et al. (1991), daß die möglichen alternativen Ausgänge gegebenenfalls schon antizipiert wurden und damit der Spielraum für eine verfälschende Wirkung des Ankers sehr viel kleiner wurde. Die Sachverhalte waren in ihrer Fragestellung geradezu darauf angelegt, daß Probanden die alternativen Ausgänge nicht aus den Augen verlieren sollten. Schließlich war das die Aufgabe zu entscheiden, ob der Sachverhalt gestiegen oder gefallen sei. Diese Fragestellung zog sich durch alle Sachverhalte hindurch. Ein Abwägen von Argumenten, die dafür sprechen, daß der Sachverhalt gefallen sein könnte, und der Argumente, die dafür sprechen, daß dieser Sachverhalt gestiegen ist, bedeutet bereits eine Antizipation des nicht gewählten Ausganges. Daher trifft diese Überlegung in hohem Maß auf alle hier präsentierten Experimente zu.

Dichotome Entscheidung Weiter kommt hinzu, daß nicht, wie in Fischhoffs Experimenten, vier, sondern nur zwei Ausgänge zu beurteilen waren. Bei vier möglichen Ereignissen beträgt die Eintretenswahrscheinlichkeit jeweils $p=.25$, bei zwei Ereignissen jedoch $p=.5$. Daher führt das Eintreten eines Ereignisses im ersten Fall zu einem größerem Ungleichgewicht gegenüber den alternativen Ereignissen, da die Eintretenswahrscheinlichkeit jedes einzelnen Ereignisses stark sinkt. Das gewählte Ereignis wird daher gegenüber den restlichen drei Ausgängen stärker hervorgehoben. Bei zwei Ereignissen hat der alternative nicht-gewählte Ausgang immer noch eine höhere Eintretenswahrscheinlichkeit als im ersten Fall die einzelnen alternativen Ereignisse. In Episoden mit mehreren Alternativen, wie etwa bei Fischhoffs Experimenten (Fischhoff, 1975; Wasserman et al., 1991), führte diese Rückmeldung dann zu einem viel stärkeren Effekt auf die Kausalstruktur der Episode. Eine Rekonzeptualisierung der Episode im Sinne von Wasserman et al. (1991), die den Ausgang mit seinen kausalen Ursachen verknüpft und ihn somit zu einem festen Bestandteil des Ereignisses selbst werden läßt, hat in Episoden mit

mehreren alternativen Ausgängen stärkere Konsequenzen. Findet eine Rekonzeptualisierung statt, wie es Wasserman et al. (1991) und auch Fischhoff (1975) nahelegen, so hat das zur Konsequenz, daß möglichst viele Aspekte der Geschichte mit dem rückgemeldeten Ausgang verknüpft und eventuell auch umgedeutet werden. Den alternativen Ausgängen bleiben dann nur noch wenige Aspekte der Geschichte, die für sie sprechen, übrig, welche sie sich zudem noch mit den anderen Ausgängen teilen müssen. Schon allein dieser Grund – fehlende bzw. ungenügende argumentative Stützung der alternativen Ausgänge durch Aspekte der Episode – sollte als Konsequenz gerade in Episoden, wie sie von Fischhoff (1975) oder Wasserman et al. (1991) verwendet wurden, auch zu einer vermindert eingeschätzten Wahrscheinlichkeit dieser Ausgänge führen.

Der rückgemeldete Ausgang einer Episode bewirkt ferner, daß die alternativen Ausgänge nicht mehr differenziert und einzeln, sondern als kompletter Block wahrgenommen werden: rückgemeldeter (wahrer) Ausgang gegen nicht eingetretene (falsche) Ausgänge. Die Konsequenzen sind, wie oben bereits erläutert, eine Konfidenzzunahme für den rückgemeldeten Ausgang sowie eine Abnahme für die restlichen, nicht rückgemeldeten, alternativen Ausgänge.

Bei den Sachverhalten in den hier vorgestellten Experimenten hingegen mit nur zwei alternativen Ausgängen, „Sachverhalt gestiegen“ respektive „gefallen“, sollte die nicht eingetretene Alternative immer noch klar durch Argumente gestützt bleiben. Schließlich waren die Experimente absichtlich so angelegt, daß die Hälfte der präsentierten Argumente jeweils einen Ausgang stützen sollte. Eine Restrukturierung im Sinne von Wasserman et al. (1991) oder Fischhoff (1975) hätte hier andere Konsequenzen. Die Struktur des Materials war von vornherein auf zwei Kausalketten angelegt: jeweils vier Argumente stützten einen Ausgang. Damit ist eine Rekonzeptualisierung des gesamten Sachverhalts mit beiden Kausalketten ausgeschlossen. Es ist sehr wohl möglich, daß zur Rückmeldung passende Argumente nochmals tiefer elaboriert wurden, aber das sollte nicht unbedingt zu Nachwirkungen auf die alternativen Argumente und deren Ausgang führen. Die Konsequenz in diesem Fall: eine drastische Verminderung der Wirkungsmöglichkeit für den rückgemeldeten Anker sowie ein deutlich verminderter Rückschaufehler.

Im Sinne der *selektiven Aktivierung* (Pohl, 1993; Pohl & Eisenhauer, 1997) – zu deren Nachweis die hier vorgestellten Experimente ja angelegt waren – hat die paritätische Verteilung der Argumente dazu geführt, daß diejenigen Argumente, die die jeweiligen Ausgänge – „Sachverhalt gefallen“ respektive „gestiegen“ – stützten, zu

einer Zweiteilung¹⁰ des Image-Sets der Sachverhalte führte: Jeweils ein Block an Images stützte jeweils einen Ausgang. Unterstützt wurde die Wahrnehmung dieser Blöcke auch durch den Versuchsaufbau, d.h. die Darbietung der Argumente in Blöcken für jeweils einen Ausgang. Das hatte zwar den Vorteil, daß keinerlei Mißverständnisse entstanden, welchen Ausgang des Sachverhalts ein Argument stützte, führte aber auch zu einer deutlichen Zweiteilung des Image-Sets. Die Konsequenz im Sinne SARA's (Pohl et al., in press) und auch der *selektiven Aktivierung* ist, daß der rückgemeldete Ausgang nur Argumente seines zugehörigen Blocks aktiviert hat. Dies wäre im Sinne der *selektiven Aktivierung* auch höchst wünschenswert, erhöhte es doch gezielt die Assoziationsstärken von Argumenten für den rückgemeldeten Ausgang und führte dann in seiner Konsequenz auch zu einer erhöhten Wahrscheinlichkeit, daß Images, die mit dem rückgemeldeten Ausgang assoziiert waren, bei der Erinnerung an die ursprüngliche Schätzung gefunden werden. Bei den Experimenten mit freier Wiedergabe gelang dieser Nachweis nur zum Teil. Es fand sich im Gedächtnisdesign der Rückschaufehler nur bei „antonymer“ - widersprechender Rückmeldung, aber der Nachweis *selektiver Aktivierung* der Argumente fehlte hier *scheinbar*¹¹.

Vieles spricht für die bereits geäußerte Interpretation des Befundes (siehe Kap. 6.1), daß hier paritätisch beide Blöcke an Argumenten aktiviert wurden: Zuerst der Block an Argumenten, die die Schätzung stützten, dann, durch die „antonyme“ Rückmeldung, genau der andere Block an Argumenten, der die Rückmeldung stützte. Das wiederum führte zu keinem großen Unterschied in der Assoziationsstärke der Images. Dieser Befund kann daher eher mit einigem Recht als ein Nachweis für den Effekt von „counterfactual reasoning“ (z.B. Koriat et al., 1980; siehe auch Kap. 2.2 und Kap. 6.1) auf die Images im Image-Set angesehen werden. Gerade dieser Befund spricht dafür, daß hier tatsächlich eine *selektive Aktivierung* stattgefunden haben könnte. Zunächst erhöhte die Schätzung selektiv die Assoziationsstärke der Argumente (Images), die zu dieser Entscheidung („gestiegen“ respektive „gefallen“) gehörten, anschließend führte die Rückmeldung des „antonymen“ Ausgangs zur selektiven Zunahme der Assoziationsstärke der Argumente des entgegengesetzten Ausgangs. Bei antonymer Rückmeldung sollten daher beide Blöcke an Argumenten des Image-Sets aktiviert worden sein.

¹⁰ Denkbar sind in SARA zwei Möglichkeiten: Ein Image-Set mit zwei polarisierten Blöcken an Images (Argumenten), oder aber zwei Image-Set mit jeweils den zugehörigen Images (Argumenten) eines Blocks. Die Konsequenz im zweiten im Gegensatz zum ersten Fall ist, daß eine Wechselbeziehung zwischen den zwei Image-Sets nicht angenommen werden muß. Meiner Meinung nach, sprechen die Ergebnisse ,der hier diskutierten Experimente, für den zweiten Fall (siehe weiter unten im Text).

¹¹ Es fand sich kein Behaltensvorteil für Argumente, die durch den Anker gestützt wurden. Diesen Befund müßte man als einen fehlenden Nachweis *selektiver Aktivierung* deuten, betrachtet man den Fall jedoch genauer, ergibt sich ein anderes Bild (siehe weiter unten im Text).

Gleichhohe Assoziationsstärken der Images für die jeweiligen Ausgänge sollten aber keinen Rückschafehler verursachen. Die Voraussetzung für diese Schlußfolgerung ist aber, daß die Prozesse alle auf das gleiche Image-Set zurückgreifen. Diese Prämisse war aber bei den hier berichteten Experimenten eventuell nicht erfüllt. So ist es sehr wohl möglich, daß die im Material angelegten Blöcke an Argumenten für je einen von zwei sich ausschließenden Ausgängen auch zur Bildung zweier Image-Sets führte. Ein Image-Set mit Argumenten für den Ausgang „der Sachverhalt ist gefallen“ und ein Image-Set für den Ausgang „der Ausgang ist gestiegen“. Wenn die Rückmeldung dazu geführt hat, daß bei der Rekonstruktion der Schätzung jeweils der zur Rückmeldung passende Block favorisiert wurde, sollte das trotz ähnlicher Assoziationsstärken der Images in beiden Blöcken zu einem Rückschafehler bei „antonymer“ Rückmeldung führen, da nur das zur Rückmeldung passende Image-Set genutzt wurde.

Wert der Images Ein weiterer wichtiger Aspekt zur selektiven Aktivierung der Argumente bei den hier verwendeten Sachverhalten ist, daß bei einem Sampling von Images, die einen diskreten Wert (und keinen kontinuierlichen numerischen Wert, wie in dem Beispiel zu *selektiver Aktivierung*; siehe Abb. 4.4 und Kap. 4.4) haben, jedes gezogene Image (Argument) entweder für „Sachverhalt gefallen“ oder „gestiegen“ sprach. Die Folge ist, daß das Ergebnis der Integration ebenfalls dichotom und nicht kontinuierlich sein kann. Mit anderen Worten, es macht keinen Unterschied, ob man lediglich ein Argument oder aber alle vier Argumente, die dafür sprechen, daß der „Sachverhalt gefallen“ ist, heranzieht. Das Ergebnis wird in jedem Fall eine Entscheidung zugunsten des Ausganges „der Sachverhalt ist gefallen“ sein. Das ist aber in dem numerischen Beispiel in Kapitel 4.4, bei der Frage nach dem Sterbealter Goethes, ganz anders. Hier hat es große Auswirkung auf die Schätzung und auch Erinnerung, ob man z.B. aufgrund der allgemeinen Lebenserwartung im 18. Jahrhundert das Todesalter Goethes auf 38 Jahre schätzt, oder auch noch andere Informationen heranzieht, wie beispielsweise, daß Goethes Werk aus über 140 Bänden besteht, und diese Information wiederum für ein Sterbealter von etwa 90 Jahren spricht. Die Konsequenz wäre, wie bereits in Kapitel 4.4 angedeutet, eine sehr unterschiedliche Schätzung: Im ersten Fall läge die Schätzung bei 38 Jahren, im letzteren Fall hingegen bei 64 Jahren.

Wichtigkeit der Argumente Hinzu kommt, daß der Rückschafehler nicht direkt an der durch die Argumente betroffenen Entscheidung, sondern erst mittelbar an der Konfidenz für die getroffene Entscheidung gemessen wurde. An der Entscheidung alleine ist nur sehr schlecht bzw. sehr grob ein Rückschafehler zu messen. Eine unterschiedliche Auswirkung der Argumente auf den Wert der Konfidenzeinschätzung, wenn die Argumente dichotom nur für den einen oder den anderen Ausgang sprechen, ist aber nur über eine Wertung der Argumente denkbar. In den Experimenten 1 und 2 hat eine solche Einschätzung der Wichtigkeit der Argumente, abgesehen von einer unproblematischen Bevorzugung von Argumenten, die für „gestiegen“ sprachen, ergeben, daß alle Argumente mit einer ähnlichen Wichtigkeit (>55%) eingestuft wurden. Das würde dann aber bedeuten, daß ein selektiver Einfluß der Images (Argumente) nur über das Verhältnis abgerufener Images auszumachen wäre: Sprechen zwei Argumente dafür, daß der Sachverhalt „gestiegen“ ist und nur eines dagegen, sollte die Sicherheit der Entscheidung zunehmen. Die numerische Entsprechung auf der Konfidenzskala ist allerdings nicht eindeutig. Die Sicherheit sollte zwar zunehmen, unklar ist jedoch, um wieviel Prozent.

Andererseits ist auch denkbar, daß die Einschätzung der Wichtigkeit der Argumente analog auf die Konfidenzskala umgesetzt wird: Die Ergebnisse aller Experimente, mit Ausnahme des letzten Experiments, legen nahe, daß die Sicherheit nicht sehr stark zugenommen hat. Möglicherweise lag das aber auch daran, daß die Argumente im Schnitt mit einer Wichtigkeit von etwa 55% eingestuft wurden, was, analog auf die Konfidenzskala übertragen, auch einen Wert nahe der Mitte erwarten ließe. Zwei Argumente, beide mit einer eingeschätzten Wichtigkeit von 50%, sprächen dann auch für eine Konfidenz von 50%. Ein Gegenargument mit einer Wichtigkeit von ebenfalls 50%, würde daran nichts ändern. Die Konfidenz aufgrund dieser drei Argumente betrüge 50%. Diese Erwartung deckt sich mit den tatsächlichen Befunden aus den Experimenten 1 bis 4: In der überwiegenden Mehrzahl lagen die Schätzungen nahe der 50%.

Wie bereits besprochen (siehe Kap. 5.1) wirkt sich aber ein Wert von 50% wegen seiner Salienz auf einer analogen, nicht-numerischen Skala sehr günstig auf die Erinnerungsleistung aus. Der genaue Wert der Konfidenz für die Schätzung ist wegen der feinen Auflösung der analogen Skala nur per Zufall erreichbar, aber daß der Wert in etwa bei 50% lag, ist noch sehr gut zu erinnern. Die Folge ist eine starke Regression zur Mitte. Die Forderung, die sich daraus ergibt, ist, daß die subjektive Wichtigkeit der Argumente erhöht werden muß, um einen selektiven Effekt der Argumente erreichen zu können. In SARA (Pohl et al., in press) wäre es denkbar,

daß Images mit dichotomen Inhalt dann nicht mit ihrem dichotomen Wert, sondern mit ihrer Wichtigkeit als Wert in das Image-Set aufgenommen würden. Es ließen sich dann auch differenziertere Effekte auf die Konfidenzskala vorhersagen.

7.1.2 Allgemeine Aspekte

Selektive Aktivierung versus selektiver Abruf Es findet sich durchaus ein Effekt *selektiver Aktivierung* als Folge „synonymer“ Rückmeldung in Experiment 4. Der gleiche Effekt findet sich allerdings auch für Sachverhalte ohne Rückmeldung. Zudem legt auch die Varianzanalyse von Experiment 5 (siehe Kap. 6.2), mit Entscheidungen entgegen der Rückmeldung, nahe, daß der selektive Effekt hier vor allem auf der Entscheidung zum Zeitpunkt der Schätzung beruhte.

In Experiment 5 deutet daher vieles darauf hin, daß der selektive Effekt eher auf Urteils-, denn auf Gedächtnisprozesse zurückzuführen ist. Jedoch läßt sich dem entgegenhalten, daß die Probanden auch in diesem fünften Experiment ihre Entscheidungen erinnern mussten, da die freie Wiedergabe erst nach Abgabe ihrer Urteile erfolgte. So sollten nach SARA (Pohl et al., in press) sowohl *selektive Aktivierung* als auch *selektiver Abruf* möglich gewesen sein. Dieses Experiment erlaubt aber keine genaue Trennung der jeweiligen Anteile. Allerdings sollte das Simulationsmodell SARA erlauben, diesen Fall zu entscheiden. Es ist möglich, den Anker entweder nur als Abrufhinweis zu nutzen, oder aber auch seine Enkodierung zuzulassen, und dann diese beiden Fälle gegeneinander zu testen (siehe Kap. 4).

Regressionseffekte und Rückschaufehler Ein weiterer Aspekt, warum in Experiment 3 und 4 bei „antonymer“ Rückmeldung eine Verschiebung der Konfidenz in Richtung der Rückmeldung nachweisbar war, soll hier verdeutlicht werden. Die Frage, die ich hier verfolgen möchte, ist, ob in den Experimenten 3 und 4 überhaupt ein substantieller Rückschaufehler zu beobachten war.

Der Schlüssel zur Interpretation dieses Befundes liegt möglicherweise in der Reanalyse zur Abschätzung der Größe des Regressionseffekts gegenüber dem Rückschaufehler. Der Regressionseffekt war in Experiment 4 mit 3.46% gegenüber nur 1.47% für den Rückschaufehler mehr als doppelt so groß. Die absolute Größe des Rückschaufehlers ist, vergleicht man diesen Effekt beispielsweise mit dem sehr ähnlichen Experiment 3, wo er 2.70% betrug, nur annähernd halb so groß. Das einzige der fünf Experimente mit einem klaren Rückschaufehler, Experiment 5, hatte einen mittleren Effekt von 12.15% (9.09% für die Rückmeldung „gestiegen“ und 15.22% für „gefallen“). Der Effekt war daher in den Experimenten 3 und 4 sehr klein. Auch wenn man in Rechnung stellt, daß der Effekt im Urteilsdesign in der Regel größer ist als im Gedächtnisdesign (Campbell & Tesser, 1983; Davies, 1992; Fischhoff, 1977; Pennington, 1981; Pennington et al., 1980; Powell, 1988; Wood, 1978, Exp. 2; siehe Kap. 2.1), und der Effekt des Retentionsintervalls von einer Woche bei den Experimenten 1-4 hinzukommt (Hell et al., 1988; Pennington, 1981; siehe Kap. 2.1), so sind Unterschiede um den Faktor 4.5 oder gar 8.3 schon sehr groß. Die signifikante Interaktion zwischen „Rückmeldung“ und „Schätzung“ beruhte in Experiment 4 daher zu zwei Dritteln auf einem Regressionseffekt und in Experiment 3 sogar zu über drei Vierteln. Der Effekt des Rückschaufehlers konnte daher nur sehr gering sein.

Hinzu kommt weiterhin eine bereits erwähnte Besonderheit der verwendeten Prozentskala: Die Endpunkte der Skala sind zugleich die Anker. Das hat zur Konsequenz, daß eine Regression zur Mitte in einem Fall entgegen der Richtung des Ankers (bei „synonymer“ Rückmeldung), in dem anderen Fall in Richtung auf den Anker wirkt (bei „antonymer“ Rückmeldung). Dies bildete ja bereits den Ausgangspunkt für die Reanalyse und Abschätzung des Regressionseffekts. Bedenkt man aber die Anwendung folgenden Arguments von Kohnert (1996):

„Nimmt man nämlich an, daß die Schätzungen der Probanden nicht zu 100% reliabel sind, so müssen die Erinnerungsverteilungen, wenn sich die Probanden an nichts mehr erinnern und daher einfach noch mal schätzen, zum Mittelwert der Schätzverteilung regredieren. Da dieser Mittelwert in der Regel nicht extrem weit vom Anker entfernt liegt, regredieren die Erinnerungen nicht nur gegen den Mittelwert der Schätzverteilung, sondern auch gegen den Anker selbst“ (Kohnert, 1996, S.60),

so folgt daraus in unserem Fall ein stärkerer Regressionseffekt für „antonyme“ Rückmeldungen. Demnach ist es zu bezweifeln, ob der Rückschaufehler überhaupt noch substantiell war.

Reaktionszeiten *Selektive Aktivierung* ließ sich in den Experimenten 1 - 3 im Wiedererkennensparadigma nicht nachweisen. Bei den Reaktionszeiten kann die Varianz der Lesezeiten den Effekt überdeckt haben, da es leider nicht möglich war, alle Argumente in einem Stichwort zusammenzufassen (viele Argumente waren in zwei bis drei Stichworten zusammengefaßt). Die dadurch stark erhöhte Varianz hat möglicherweise die vorhandene Varianz durch die selektive Aktivierung überdeckt. Da sich aber auch in der signalentdeckungstheoretischen Auswertung keinerlei selektive Effekte der Rückmeldung nachweisen ließen, liegt der Schluß nahe, daß in den Experimenten 1 - 3 keine nachweisbare selektive Aktivierung aufgetreten ist.

In den Experimenten 4 und 5 mit freier Wiedergabe ließ sich ein selektiver Effekt der Schätzung nachweisen, wobei eine Entscheidung zwischen gedächtnisbasierten versus urteilsbasierten Prozessen nicht sicher getroffen werden konnte. Es ist daher bedauerlich, daß bei dem hier gewählten Material die Stichworte zu so unterschiedlichen Lesezeiten geführt haben, daß die Effekte möglicherweise verdeckt wurden. Für eine Entscheidung zwischen gedächtnis- versus urteilsbasierten Prozessen wären Reaktionszeiten ein sehr sensibles und valides Maß. Die Erhöhung der Assoziationsstärken in SARA (Pohl et al., in press) sollte zunehmend zu besserem und schnellerem Abruf führen. Die Voraussetzung ist allerdings, daß der Anker enkodiert wurde und das Image-Set verändert hat. Handelt es sich allerdings nur um einen urteilsbasierten Prozeß, sollten durch Enkodierung erhöhte Assoziationsstärken keinen weiteren differenzierenden Einfluß auf die Images haben.

Widerspruch Eine weitere interessante Frage, die bereits in der Diskussion zu Experiment 3 und 4 auftauchte, ist, ob erlebter Widerspruch (zwischen eigener Schätzung und Anker) ein konstituierendes Element des Rückschaufehlers sein könnte. Die Daten aus diesen beiden Experimenten ließen sich jedenfalls so interpretieren. Hier ließe sich eine Verbindung zu Erklärungsansätzen des Rückschaufehlers machen, die stark die konstituierende Rolle von Metakognitionen betonen (Strack & Mussweiler, 1997; Werth, 1998). Möglicherweise spielen bei den hier gewählten Sachverhalten Metakognitionen ebenfalls eine wichtige Rolle.

Die Sachverhalte und vor allem die Konfidenzskala waren so angelegt, daß leicht eine nicht beabsichtigte Veränderung der Skaleneigenschaften der Konfidenzskala stattfinden konnte (siehe Kap. 6.1): Die Skala von 0-100% Sicherheit, daß ein Sachverhalt „gestiegen“ ist, konnte beispielsweise in eine Skala mit nur drei Kategorien für „gefallen“, „unentschieden“ und „gestiegen“ verändert worden sein. Die Konsequenz wäre, wie bereits erläutert, daß nur bei einer „antonymen“, als

diskrepanz erlebten Rückmeldung, überhaupt ein Bedarf zur Anpassung des eigenen Urteils zum Zeitpunkt der Schätzung entstehen konnte und nur in diesem Fall die Rückmeldung überhaupt enkodiert und mit dem Image-Set verknüpft wurde. Das würde ferner die viel stärkeren selektiven Effekte der Schätzung erklären können. Wie bereits erwähnt, stützt auch die traditionelle Auswertung der Experimente 3 und 4 (nur nach erfolgter Rückmeldung) diese Interpretation, da der Rückschaufehler nur bei der weniger favorisierten Rückmeldung auftrat und somit auch dieser Befund als Beleg für subjektiv erlebten Widerspruch gedeutet werden kann.

7.2 Abschluß und Ausblick

7.2.1 Fazit

Die hier vorgestellten Experimente lieferten einen sehr profunden Einblick in die Prozesse, die den Rückschaufehler beeinflussen. Die Experimente ohne den typischen robusten Effekt lassen sich positiv als erfolgreicher Nachweis präventiven *counterfactual reasonings* interpretieren. Andererseits werfen diese Befunde auch die Frage auf, ob die Annahme hauptsächlich automatischer Prozesse der Informationsverarbeitung für den Effekt ausreicht. Die grundlegende Beobachtung, daß sowohl Ankereffekt als auch Rückschaufehler äußerst robuste und intentional nur geringfügig zu beeinflussende Phänomene sind (Christensen-Szalanski & Willham, 1991; Pohl & Hell, 1996), stellen die Befunde dieser Arbeit in Zweifel. Die Befunde sprechen sehr für die Beteiligung intentionaler, strategischer oder metakognitiver Einflüsse (Hawkins & Hastie, 1990; Kohnert, 1996; Mussweiler et al., 1997; Schwarz & Stahlberg, 1999; Werth, 1998). Vieles spricht dafür, daß bei dem hier verwendeten Material evaluative Prozesse eine wichtige Rolle gespielt haben.

Die Experimente mit freier Wiedergabe zeigten erfolgreich einen selektiven Effekt der Rückmeldung und auch der Entscheidung zum Zeitpunkt der Schätzung. Dies ist der erste Nachweis *selektiver Aktivierung* von Informationen eines experimentell vermittelten Image-Sets durch die Rückmeldung und die Entscheidung zum Zeitpunkt der Schätzung. Für das Prozessmodell SARA (Pohl et al., in press) ist dieser Nachweis zur Erklärung der Prozesse, die zum Rückschaufehler führen, zentral. Das in Kapitel 4 ausführlich vorgestellte Modell SARA liefert erstmals eine umfassende Theorie mit breitem Anwendungsbereich im Bereich verfälschter Urteile und Erinnerungen. Ein Vorteil des Modells ist, daß es einerseits weitgehend auf klassischen Annahmen der kognitiven Psychologie beruht, andererseits aber flexible Prozeßmodule postuliert, die für viele Arten experimenteller Abläufe nutzbar gemacht werden können. Die entscheidende Frage für das Modell und auch für die Erklärung des Rückschaufehlers ist aber, ob die selektive Wirkung bereits durch die Enkodierung des Ankers, oder erst später zum Zeitpunkt des Abrufs stattgefunden hat? Mit anderen Worten, beruht der Effekt allein auf urteilsbasierten Prozessen oder basiert er auch auf gedächtnisbasierten Prozessen? SARA erlaubt es, diesen Fall zu entscheiden. Es ist möglich, den Anker entweder nur als Abrufhinweis zu nutzen, oder aber auch eine Enkodierung zuzulassen, und dann diese beiden Fälle gegeneinander zu testen (siehe Kap. 4).

Die Wahl des Materials in den hier vorgestellten Experimenten stellte sich als problematisch heraus. Die künstliche Vermittlung eines Image-Sets zu Sachverhalten mit paritätischen Argumenten zugunsten der nur zwei Alternativen umfassenden Ausgänge, die Wahl von zunächst sehr neutralen Sachverhalten, führte zu Antworten nahe der Mitte der Konfidenzskala. Rückmeldungen zu den Sachverhalten wurden eventuell nur im Falle eines erlebten Widerspruchs enkodiert und führten dann zu einem kleinen Effekt.

7.2.2 Ausblick

Fraglich zur Lösung des Problems mit der Art von Material, wie es in den hier vorgestellten Experimenten verwendet wurde, ist jedoch, ob ein Ungleichgewicht der Argumente die grundsätzliche Lösung dieses Problem darstellte. Vielmehr erscheint es mir konsequent, selektive Aktivierung an Material mit mehr als zwei Ausgängen zu untersuchen. Ein Rückgriff auf episodisches Material, wie es in den ersten Experimenten von Fischhoff (1975) benutzt wurde, oder aber eine komplexe wirtschaftliche Entscheidungssituation, wie bei Walster (1967), erlaubte eine bessere Einbindung von Argumenten für mehrere Ausgänge. Walster (1967) fand, daß bei einem Hauskauf der Effekt sehr stark abhängig von der Höhe des erzielten Gewinns oder Verlustes variierte: Je extremer der Ausgang, desto höher der Rückschafehler. Sie interpretierte dieses Ergebnis mit dem Bedürfnis ihrer Vpn, die Welt bei allzu großen Diskrepanzen wieder stimmig und vorhersagbar zu machen. SARA würde den Befund als eine sehr elaborierte Enkodierung des Ankers deuten, die besonders die Assoziationsstärke von kausal mit dem Anker verknüpften Argumenten erhöhte. Mit diesem Paradigma wären daher sehr viel stärkere Effekte des Ankers zu erwarten, als bei den Sachverhalten in den hier vorgestellten Experimenten. Eine weitere Veränderung des Materials mit der Einführung mehrerer möglicher Ausgänge sollte vermeiden, daß innerhalb des Image-Sets die Assoziationsstärke der Argumente bereits zu hoch mit der Frage und ihren möglichen alternativen Ausgängen verbunden wären, und ein selektiver Einfluß des Ankers nur noch schwer nachweisbar wäre. Mehr alternative Ausgänge sollten zu einer verminderten Abrufwahrscheinlichkeit einzelner Images führen.

Zwei Anker Wirkt der Anker lediglich als Abrufcue, so sollte seine selegierende Wirkung nur durch seine präexperimentelle Lerngeschichte bestimmt sein. In SARA würden wir annehmen, daß seine Wirkung als Cue auf seiner numerischen Ähnlichkeit zu Images des Image-Sets besteht. Wirkt der Anker aber auch durch seine Enkodierung in das Image-Set selegierend, so sollte dieser Samplingprozeß

gezielt die Assoziationsstärken der Images zueinander erhöht haben, die gemeinsam im Arbeitsgedächtnis waren (Siehe Kap. 4.4). Dies sollte aber für manche Images zu einer deutlich höheren Assoziationsstärke geführt haben. Ein abgewandeltes Urteilsexperiment mit Reaktionszeiten als abhängige Variable erlaubte dann Evidenz für oder wider gedächtnisbasierter Prozesse zu finden:

In einem einführenden Text zu kausalen Entscheidungen gibt man den Probanden mehrere (mindestens aber drei) vom Inhalt sehr ähnliche Episoden mit jeweils unterschiedlichem Ausgang zu lesen. Dieser Text betont die Argumente, die gerade zu diesem Ausgang des Ereignisses geführt haben. In der nächsten, ähnlichen Episode, die die Probanden zu lesen bekommen, wird ein anderer Ausgang favorisiert. Wieder werden die Argumente, die zu diesem Ausgang geführt haben, betont. Es folgt eine weitere Episode mit einem weiteren Ausgang und den zugehörigen Argumenten.

In der Testphase wird dann wieder eine ähnliche Episode mit einem Ausgang präsentiert - aber ohne Argumente. Die Probanden werden gebeten, die Wahrscheinlichkeit für alle möglichen Ausgänge einzuschätzen. Anschließend folgt ein Wiedererkennenstest mit alten - aus dem einführenden Text stammenden - und neuen, noch nicht präsentierten Argumenten. Der rückgemeldete Ausgang ist aber jetzt ein anderer.

Aus urteilsbasierter Sicht sorgt der neue Anker als Abrufhinweis jetzt für den selektiven Abruf von Images, die mit diesem Anker assoziiert sind. Aus gedächtnisbasierter Sicht sollte der erste Anker ebenfalls enkodiert worden sein und auch zu einer zusätzlich erhöhten Abrufwahrscheinlichkeit für alte Argumente sorgen, die mit der ersten Rückmeldung assoziiert sind. Findet sich ein Datenmuster mit schnelleren Reaktionszeiten nur für Argumente, die mit dem zweiten Anker assoziiert sind, wäre das ein starkes Argument für einen rein urteilsbasierten Prozeß. Finden sich hingegen auch schnellere Reaktionszeiten für Argumente, die mit dem ersten Anker assoziiert sind, spräche dieser Befund stark für einen gedächtnisbasierten Effekt *selektiver Aktivierung*.

Abschließen möchte ich mit einem Zitat Jorge Luis Borges:

Ich glaube, man sollte nicht vergessen, was der französische Philosoph Bergson sagte, der behauptete, daß das Gedächtnis selektiv sei, das heißt, das Gedächtnis wählt aus. Wenn die Menschen von Natur oder Temperament her eher pathetisch sind, neigen sie dazu, sich an ihr Unglück zu erinnern, da das Unglück ihnen als Vorwand für pathetische Reden dient. Da ich aber keinen Hang zum Pathos habe, oder versuche, nicht pathetisch zu sein, vergesse ich das Schlechte und die Erinnerung an Kummer... Also, ich glaube, daß die Erinnerung des Vergessens bedarf. (Jorge Luis Borges und Oswaldo Ferrari, 1990; Lesen ist Denken mit fremden Gehirn. Gespräche über Bücher & Borges. Übersetzung von Gisbert Haefs. Zürich: Arche Verlag.)

Literatur

- Alba, J. W. & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, 93, 203-231.
- Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22, 261-295.
- Anderson, J. R. and R. Milson (1989). Human memory: An adaptive perspective. *Psychological Review* 96(4): 703-719.
- Anderson, J. R. & Schooler, J. W. (1991). Reflections of the environment in memory. *Psychological Science*, 2, 396-408.
- Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.
- Anderson, N. H. (1986). Algebraic rules in psychological measurement. In H. R. Arkes & K. R. Hammond (Eds.), *Judgment and decision making: An interdisciplinary reader* (pp. 77-94). Cambridge: Cambridge University Press.
- Anderson, J. A., & Mozer, M. C. (1981). Categorization and selective neurons. In G. E. Hinton & J. A. Anderson (Eds.), *Parallel models of memory* (pp. 213-236). Hillsdale, NJ: Erlbaum.
- Arkes, H. R. (1988). Commentary on the article by Verplanken and Pieters. *Journal of Behavioral Decision Making*, 1, 146.
- Arkes, H. R., Faust, D., Guilmette, T. J., & Hart, K. (1988). Eliminating the hindsight bias. *Journal of Applied Psychology*, 73(2), 305-307.
- Arkes, H. R., Wortmann, R., Saville, P. D., & Harkness, A. R. (1981). Hindsight bias among physicians weighing the likelihood of diagnoses. *Journal of Applied Psychology*, 66, 252-254.
- Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, 54, 569-579.

- Bellezza, F. S. (1981). Mnemonic devices: Classification, characteristics, and criteria. *Review of Educational Research, 51*(2), 247-275.
- Belli, R. F. (1989). Influences of misleading postevent information: Misinformation interference and acceptance. *Journal of Experimental Psychology: General, 118*(1), 72-85.
- Bem, D. J., & McConnell, H. K. (1970). Testing the self-perception explanation of dissonance phenomena: On the salience of premanipulation attitudes. *Journal of Personality and Social Psychology, 3*(3), 205-211.
- Buchman, T. A. (1985). An effect of hindsight on predicting bankruptcy with accounting information. *Accounting Organizations and Society, 10*, 267-285.
- Bukszar, E. W., & Connolly, T. (1988). Hindsight bias and strategic choice: Some problems in learning from experience. *Academy of Management Journal, 31*(3), 628-641.
- Carrette, T. R., & Moreland, R. L. (1983). The direct and indirect effects of inadmissible evidence. *Journal of Applied Social Psychology, 13*, 291-309.
- Casper, J. D., Benedict, K., & Kelly, J. R. (1988). Cognitions, attitudes, and decision making in search and seizure cases. *Journal of Applied Social Psychology, 18*, 93-113.
- Campbell, J. D., & Tesser, A. (1983). Motivational interpretations of hindsight bias: An individual difference analysis. *Journal of Personality, 51*(4), 605-620.
- Christensen-Szalanski, J. J. J., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes, 48*(1), 147-168.
- Clark, L. F., & Woll, S. B. (1981). Stereotype biases: A reconstructive analysis of their role in reconstructive memory. *Journal of Personality and Social Psychology, 3*(3), 205-211.
- Connolly, T., & Bukszar, E. W. (1990). Hindsight bias: Self-flattery or cognitive error? *Journal of Behavioral Decision Making, 3*(3), 205-211.

- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Davies, M. F. (1987). Reduction of hindsight bias by restoration of foresight perspective: Effectiveness of foresight-encoding and hindsight-retrieval strategies. *Organizational Behavior and Human Decision Processes*, 40(3), 50-68.
- Davies, M. F. (1992). Field dependence and hindsight bias: Cognitive restructuring and the generation of reasons. *Journal of Research in Personality*, 26(1), 58-74.
- Dawson, N. V., Arkes, H. R., Siciliano, C., Blinkhorn, R., Lakshmanan, M., & Petrelli, M. (1988). Hindsight bias: An impediment to accurate probability estimation in clinicopathologic conferences. *Medical Decision Making*, 8, 259-264.
- Devine, P. G., & Ostrom, T. M. (1985). Cognitive mediation of inconsistency discounting. *Journal of Personality and Social Psychology*, 49, 5-21.
- Ebbinghaus, H. (1885). *Über das Gedächtnis*. Leipzig: Dunker.
- Edwards, W. (1984). How to make good descisions. *Acta Psychologica*, 56, 5-27.
- Eisenhauer, M. (1996). Selective activation as explanation of the hindsight bias. In M. Conway, C. Cornoldi & S. Gathercole (Ed.), *Second Conference on Memory* (14-19 July 1996) (pp. 36). Padova.
- Eisenhauer, M. (2000). Rückschaufehler unter Zeitbeschränkung. In D. Vorberg, A. Fuchs, U. Heinrich, U. Mattler, & S. Töllner (Ed.), *Experimentelle Psychologie - Beiträge zur 40. Tagung experimentell arbeitender Psychologen und Psychologinnen*, Braunschweig 2000 (pp. 122). Berlin: Pabst Verlag.
- Eller, F. (1984) *Der Knew-it-all-along-Effekt: Eine Untersuchung zum intuitiven Urteils- und Entscheidungsverhalten*. Diplomarbeit, Institut für Psychologie, Universität Kiel.
- Erdfelder, E., & Buchner, A. (1998). Decomposing the hindsight bias: An integrative multinomial processing tree model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(2), 387-414.

- Fama, E. O. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25, 383-417.
- Fischhoff, B. (1975). Hindsight \neq foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1(3), 288-299.
- Fischhoff, B. (1977). Perceived informativeness of facts. *Journal of Experimental Psychology: Human Perception and Performance*, 3(2), 349-358.
- Fischhoff, B. (1980). For those condemned to study the past: Reflections on historical judgment. In R. A. Shweder & D. W. Fiske (Eds.), *New directions for methodology of behavioral science: Fallible judgment in behavioral research* (pp. 79-93). San Francisco: Jossey-Bass.
- Fischhoff, B., & Beyth, R. (1975). "I knew it would happen." Remembered probabilities of once-future things. *Organizational Behavior and Human Performance*, 13, 1-16.
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A rebuttal to Kahneman and Tversky (1996). *Psychological Review*, 103(3), 592-596.
- Goethals, G. R., & Reckman, R. F. (1973). The perception of consistency in attitudes. *Journal of Experimental Social Psychology*, 9, 491-501.
- Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, 35(7), 603-618.
- Hasher, L., Attig, M. S., & Alba, J. W. (1981). I knew it all along: Or, did I? *Journal of Verbal Learning and Verbal Behavior*, 20, 86-96.
- Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, 107(3), 311-327.
- Hell, W. (1993). Gedächtnistäuschungen. In W. Hell, K. Fiedler, & G. Gigerenzer (Eds.), *Kognitive Täuschungen* (pp. 13-38). Heidelberg: Spektrum der Wissenschaften.

- Hell, W., Gigerenzer, G., Gauggel, S., Mall, M., & Müller, M. (1988). Hindsight bias: An interaction of automatic and motivational factors? *Memory and Cognition*, 16(6), 533-538.
- Hennessey, J. E., & Edgell, S. E. (1991). Hindsight bias in a very sparse environment. *Bulletin of the Psychonomic Society*, 29(5), 433-436.
- Hoch, S. J., & Loewenstein, G. F. (1989). Outcome feedback: Hindsight and information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(4), 605-619.
- Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning, and discovery*. Cambridge, MA: MIT Press.
- Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, 21(11), 1161-1166.
- Janoff-Bulman, R., Timko, C., & Carli, L. L. (1985). Cognitive biases in blaming the victim. *Journal of Experimental Social Psychology*, 21, 161-177.
- Kahneman, D. (1992). Reference points, anchors, norms, and mixed feelings. *Organizational Behavior and Human Decision Processes*, 51, 296-312.
- Kohnert, A. (1996). *Grenzen des Rückschaufehlers: Die Verzerrung von Erinnerungen an früheres Wissen durch neue Informationen*. Bonn: Holos.
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 107-118.
- Leary, M. R. (1981). The distorted nature of hindsight. *Journal of Social Psychology*, 115, 25-29.
- Leary, M. R. (1982). Hindsight distortion and the 1980 presidential election. *Personality and Social Psychology Bulletin*, 8(2), 257-263.
- Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7, 560-572.

- Loftus, E. F. (1977). Shifting human color memory. *Memory and Cognition*, 5(6), 696-699.
- Loftus, E. F., Korf, N. L., & Schooler, J. W. (1989). Misguided memories: Sincere distortions of reality. In J. C. Yuille (Eds.), *Credibility assessment* (pp. 155-173). Dordrecht: Kluwer.
- Loftus, E. F., & Loftus, G. R. (1980). On the permanence of stored information in the human brain. *American Psychologist*, 35, 409-420.
- Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, 4(1), 19-31.
- Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and "nonparametric" indexes. *Psychological Bulletin*, 107, 401-413.
- Macmillan, N.A. (1992). Signal Detection Theory as Data Analysis Method and Psychological Decision Model. In Keren, G., & Lewis, C. (Hrsg.), *A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.
- McCloskey, M. & Zaragoza, M. S. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, 114, 1-16.
- McNamara, T. P. (1992). Priming and constraints it places on theories of memory and retrieval. *Psychological Review*, 99, 650-662.
- McNicol, D. (1972). *A primer of signal detection theory*. London: George Allen & Unwin.
- Mensink, G.-J., & Raaijmakers, J. G. W. (1988). A model of interference and forgetting. *Psychological Review*, 95(4), 434-455.
- Mensink, G.-J., & Raaijmakers, J. G. W. (1989). A model of contextual fluctuation. *Journal of Mathematical Psychology*, 33, 172-186.

- Metcalfe, J., & Bjork, R. A. (1991). Composite models never (well, hardly ever) compromise: Reply to Schooler and Tanaka (1991). *Journal of Experimental Psychology: General*, 120(2), 203-210.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.
- Mitchell, T. R., & Kalb, L. S. (1981). Effects of outcome knowledge and outcome valence on supervisors' evaluations. *Journal of Applied Psychology*, 66, 604-612.
- Neisser, U. (1981). John Dean's memory: A case study. *Cognition*, 9, 1-22.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Padaver-Singer, A. M., & Barton, A. H. (1975). The impact of pretrial publicity on jurors' verdicts In R. J. Simon (Ed.), *The jury system in America*, (pp. 112-127). Beverly Hills, CA: Sage.
- Pennington, D. C. (1981). Being wise after the event: An investigation of hindsight bias. *Current Psychological Research*, 1, 271-282.
- Pennington, D. C., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 521-533.
- Pennington, D. C., Rutter, D. R., McKenna, K., & Morley, I. K. (1980). Estimating the outcome of a pregnancy test: Women's judgements in foresight and hindsight. *British Journal of Social and Clinical Psychology*, 19, 317-324.
- Pichert, J. W., & Anderson, R. C. (1977). Taking different perspectives on a story. *Journal of Educational Psychology*, 69(4), 309-315.
- Pohl, R. F. (1992). Der Rückschau-Fehler: systematische Verfälschung der Erinnerung bei Experten und Novizen. *Kognitionswissenschaft*, 3(1), 38-44.

- Pohl, R. F. (1993) *Der Rückschaufehler - Ein Modell zur Analyse und Erklärung systematisch verfälschter Erinnerungen*. Habilitationsschrift, FB I - Psychologie der Universität Trier.
- Pohl, R. F. (1998). The effects of feedback source and plausibility of hindsight bias. *European Journal of Cognitive Psychology*, 10 (2), 191-212.
- Pohl, R. (1999). Der Einfluß suggestiver Faktoren auf den Ankereffekt. Experimentelle und klinische Hypnose [Themenheft: *Suggestion und Suggestibilität*], 15, 29-53.
- Pohl, R. & Eisenhauer, M. (1997). Ein kognitives Prozeßmodell zur Erklärung systematisch verzerrter Urteile und Erinnerungen. In H. Mandl (Hrsg.), *Bericht über den 40. Kongreß der Deutschen Gesellschaft für Psychologie in München 1996* (S. 807-812). Göttingen: Hogrefe.
- Pohl, R. & Gawlik, B. (1995). Hindsight bias and misinformation effect: Separating blended recollections from other recollection types. *Memory*, 3, 21-55.
- Pohl, R. F., Hardt, O. & Eisenhauer, M. (2000). SARA – Ein kognitives Prozeßmodell zur Erklärung von Ankereffekt und Rückschaufehler. *Kognitions-wissenschaft*, 9, 77-92.
- Pohl, R. F., & Hell, W. (1996). No reduction of hindsight bias with complete information and repeated testing. *Organizational Behavior and Human Decision Processes*, 67(1), 49-58.
- Pohl, R., Ludwig, M. & Ganner, J. (1999). Kein Zusammenhang zwischen Grad der Elaboration und Ausmaß des Rückschaufehlers. *Zeitschrift für Experimentelle Psychologie*, 4, 275-287.
- Powell, J. L. (1988). A test of the "knew-it-all-along" effect in the 1984 presidential and statewide elections. *Journal of Applied Social Psychology*, 18(9), 760-773.
- Ratcliff, R. and G. McKoon (1981). "Automatic and strategic priming in recognition." *Journal of Verbal Learning and Verbal Behavior* 20, 204-215.

- Ross, L., Lepper, M. R., Strack, F., & Steinmetz, J. (1977). Social explanations and social expectations: Effects of real and hypothetical explanations on subjective likelihood. *Journal of Personality and Social Psychology*, 35, 817-829.
- Ross, M., & Conway, M. A. (1986). Remembering one's own past: The construction of personal histories. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition* (pp. 122-144). New York: Guilford Press.
- Ross, M., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 32, 880-892.
- Ross, M., McFarland, C., & Fletcher, G. J. (1981). The effect of attitude on the recall of personal histories. *Journal of Personality and Social Psychology*, 40, 627-634.
- Schooler, J. W., & Tanaka, J. W. (1991). Composites, compromises, and CHARM: What is the evidence for blend memory representations? *Journal of Experimental Psychology: General*, 120(1), 96-100.
- Schul, Y., & Burnstein, E. (1985). When discounting fails: Conditions under which individuals use discredited information in making a judgment. *Journal of Personality and Social Psychology*, 49, 894-903.
- Schwarz, N., & Bless, H. (1992). Assimilation and contrast effects in attitude measurement: An inclusion/exclusion model. In J. F. Sherry & B. Sternthal (Eds.), *Advances in consumer research* (pp. 72-77). Provo, UT: Association for Consumer Research.
- Shank, R.C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.
- Sherman, S. J., Skov, R. B., Hervitz, E. F. & Stock, C. B. (1981). The effects of explaining hypothetical future events: From possibility to probability to actuality and beyond. *Journal of Experimental Social Psychology*, 17, 142-158.
- Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 544-551.

- Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance*, 6, 649-744.
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117, 34-50.
- Sporer, S. L. (1993). Der Falschinformationseffekt: Verschmelzung, Koexistenz oder Abruffehler. In L. Montada (Hrsg.), *Bericht über den 38. Kongreß der Deutschen Gesellschaft für Psychologie in Trier 1992* (Band 2; S. 345-356). Göttingen: Hogrefe.
- Stahlberg, D., Eller, F., & Frey, D. (1990). Effekte nachträglicher Informationen auf die Erinnerungsleistung - Zur Konvergenz zweier Forschungsrichtungen. In D. Frey (Eds.), *Bericht über den 37. Kongress der Deutschen Gesellschaft für Psychologie in Kiel 1990* (pp. 143). Kiel: Verlag für Psychologie.
- Stahlberg, D., Eller, F., Romahn, A., & Frey, D. (1993). Der Knew-it-all-along-Effekt in Urteilsituationen von hoher und geringer Selbstwertrelevanz. *Zeitschrift für Sozialpsychologie*, 24(2), 94-102.
- Stahlberg, D., & Maass, A. (1998). Hindsight bias: Impaired memory or biased reconstruction? In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (pp. 105-132). Chichester: Wiley.
- Sue, S., Smith, R. E., & Caldwell, C. (1973). Effects of inadmissible evidence on the decision of simulated jurors: A moral dilemma. *Journal of Applied Social Psychology*, 3, 345-353.
- Sue, S., Smith, R. E., & Gilbert, R. (1974). Biasing effects of pretrial publicity on judicial decisions. *Journal of Criminal Justice*, 2, 163-171.
- Sue, S., Smith, R. E., & Pedroza, G. (1975). Authoritarianism, pretrial publicity and awareness of bias in simulated jurors. *Psychological Reports*, 37, 1299-1302.
- Synodinos, N. E. (1986). Hindsight distortion: "I-knew-it-all-along and I was sure about it.". *Journal of Applied Social Psychology*, 16(2), 107-117.

- Snyder, M., & Uranowitz, S. W. (1978). Reconstructing the past: Some cognitive consequences of person perception. *Journal of Personality and Social Psychology*, 36, 941-950.
- Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, 73(3), 437-446.
- Trabasso, T., & Van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, 24, 612-630.
- Tufte, E. R., & Sun, R. A. (1975). Are there bellwether electoral districts?. *Public Opinion Quarterly*, 39, 1-18..
- Tulving, E. (1983). *Elements of episodic memory*. Oxford: Clarendon Press.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Verplanken, B., & Pieters, R. G. M. (1988). Individual differences in reverse hindsight bias: I never thought something like Chernobyl would happen. Did I? *Journal of Behavioral Decision Making*, 1(3), 131-147.
- Wasserman, D., Lempert, R. O., & Hastie, R. (1991). Hindsight and causality. *Personality and Social Psychology Bulletin*, 17(1), 30-35.
- Wippich, W. (1990). Erinnerungen an Gerüche: Benennungsmaße und autobiographische Erinnerungen zeigen Geruchsnachwirkungen an. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 37(4), 679-695.
- Wood, G. (1978). The "knew-it-all-along" effect. *Journal of Experimental Psychology: Human Perception and Performance*, 4(2), 345-353.
- Wyer, R. S., Jr., & Budesheim, T. L. (1987). Person memory and judgment: The impact of information that one is told to disregard. *Journal of Personality and Social Psychology*, 53, 14-29.

Wyer, R. S., Jr., & Unversagt, W. H. (1985). The effects of instructions to disregard information on ist subsequent recall and use in making judgements. *Journal of Personality and Social Psychology*, 48, 533-549.

Anhang

A.1 Sachverhalte zu Experiment 1 - 2

1. Die Stadtwerke führten jüngst eine Umfrage zur Benutzung der Busse unter den Bürgern der Stadt durch. Hat die Benutzung der öffentlichen Busunternehmen im vergangenen Jahr zugenommen oder abgenommen?

Argumente für gefallen:

weil schlechte Anschlußverbindungen bestehen.

weil immer mehr Bürger bei Fahrten innerhalb der Stadt ihr eigenes Auto benutzen.

weil die Busse häufig Verspätung haben.

weil die Busse zu unregelmäßig fahren.

weil preiswerte Reiseangebote reduziert wurden.

weil Billigtarife für Kinder, Studenten und Senioren abgeschafft wurden.

Argumente für gestiegen:

weil viele Bürger es aus Umweltschutzgründen vermeiden, bei Fahrten innerhalb der Stadt ihr Auto zu benutzen.

weil durch den Ausbau des Streckennetzes alle Stadtbezirke gut erreichbar sind.

weil die Busunternehmen seit einiger Zeit Städtereisen anbieten.

weil die Busse in letzter Zeit immer voll sind.

weil ein Wochenendticket eingeführt wurde.

weil die Fahrpreise verringert wurden.

2. Der TSV Wilmershausen ist der älteste Sportverein der Stadt. Ist die Zahl der Vereinsmitglieder (aktive und passive) im letzten Jahr gestiegen oder gefallen?

Argumente für gefallen:

weil die Mitgliedsbeiträge gestiegen sind.

weil von ehemaligen Mitgliedern der "Konkurrenzverein" VfL Wilmershausen gegründet wurde.

weil viele an Leichtathletik interessierte Jugendliche zu den in diesem Bereich wesentlich erfolgreicherer Vereinen in den Nachbarstädten gewechselt sind.

weil viele Mitglieder mit der Vereinsführung unzufrieden sind.

weil die Sportanlagen teilweise veraltet sind.

weil es in vielen Sportbereichen an qualifizierten Trainern mangelt.

Argumente für gestiegen:

weil ein breiteres Angebot an Sportarten realisiert wurde.

weil der TSV neue interessante Freizeitaktivitäten wie Wandern oder Fahrradtouren anbietet.

weil durch den Aufstieg der ersten Fußballmannschaft in die Oberliga das Fußballfieber ausgebrochen ist.

weil der TSV nun ein besonders familienfreundliches Programm hat.

weil der TSV nun auch Seniorensport betreibt.

weil der TSV eine groß angelegte "Bleib gesund-treibe Sport"-Kampagne gestartet hat.

3. Ist die Zahl der Verkehrsunfälle 1995 gegenüber dem Vorjahr gestiegen oder gefallen?

Argumente für gefallen:

weil in Wilmershausen unerwartete Radarkontrollen eingeführt worden sind.

weil nun in vielen Wohnbezirken nicht mehr schneller als Tempo 30 gefahren werden darf.

weil zusätzlich an vielen Stellen Fußgängerüberwege eingerichtet worden sind.

weil an vielen ehemals gefährlichen Kreuzungen der Verkehr jetzt durch Ampeln geregelt wird.

weil eine Umgehungsstraße gebaut wurde, um hohes Verkehrsaufkommen in der Stadt zu vermeiden.

weil zusätzlich in vielen Stadtbezirken Radfahrwege angelegt worden sind.

Argumente für gestiegen:

weil im letzten Jahr deutlich mehr Fahrer unter Alkoholeinfluß gefahren sind.

weil es in Wilmershausen immer öfter zu Spritztouren Minderjähriger mit zumeist gestohlenen Fahrzeugen kommt.

weil sich der Anteil an Unfallopfern unter den chirurgischen Patienten der Krankenhäuser erhöht hat.

weil sich das Verkehrsaufkommen erhöht hat.

weil in Wilmershausen neuerdings illegale Autorennen auf der Straße stattfinden.

weil es durch eine neue Verkehrsführung in Wilmershausen viele unübersichtliche Verkehrsbereiche gibt, wo immer wieder Fußgänger zu Schaden kommen.

4. Die Partei "Ökologische Initiative" verfehlte bei der letzten Wahl den Einzug in den Stadtrat. Nun steht die Wahl zum Stadtrat wieder an. Sind die Chancen der "Ökologischen Initiative" gegenüber der letzten Wahl gestiegen oder gefallen?

Argumente für gefallen:

- weil der Parteivorsitzende nach den letzten Pressemeldungen in eine Steueraffäre verwickelt sein soll.
- weil sich ökologische Themen nunmehr auch in den Programmen der anderen Parteien finden.
- weil die Partei nach einigen der letzten Umfragen nicht mit dem Einzug in den Stadtrat rechnen kann.
- weil im letzten Jahr ein Ortsverein der Grünen gebildet wurde, der in Konkurrenz zu der Partei steht.
- weil mit Klaus Petersen ein wichtiger Hoffnungsträger aus der Partei ausgetreten ist.
- weil sich der Gründer der Partei, der eine wichtige Identifikationsfigur für die Bürger darstellte, aus der Politik zurückgezogen hat.

Argumente für gestiegen:

- weil sich die Partei in der letzten Zeit verstärkt für den Erhalt der städtischen Grünanlagen eingesetzt hat, was ihr bei den Bürgern viel Sympathien eingebracht hat.
- weil viele Bürger mit der Arbeit des Stadtrats während der letzten Periode unzufrieden waren.
- weil die Partei gegenüber der letzten Wahl nun auch von vielen Bürgerinitiativen unterstützt wird.
- weil sich die Partei nun verstärkt für Verbesserungen im sozialen Bereich (Kindergartenplätze, Versorgung älterer Menschen usw.) einsetzt, die von Bürgern seit langem gefordert werden.
- weil sich für die bevorstehende Wahl viele Personen aus Kunst und Kultur der Stadt Wilmershausen für die Partei stark machen.
- weil die Partei für diese Wahl ein umfassendes Konzept mit einer langfristigen Perspektive entworfen hat, in dem neben ökologischen Gesichtspunkten auch andere Probleme der Stadt thematisiert werden.

5. Die Bäckerei Wolters stellt seit einigen Jahren Brot und Backwaren auf ökologischer Basis her. Ist der Gewinn der Bäckerei während der letzten zwei Jahre gestiegen oder gefallen?

Argumente für gefallen:

weil die Produkte im Vergleich zu konventionellen Erzeugnissen viel teurer sind.

weil die Konkurrenz zwischen den Bäckern in Wilmershausen durch die Eröffnung neuer Bäckereien in den letzten Jahren gestiegen ist.

weil die Einkaufszentren zunehmend eigene Bäcker haben, so daß viele Bürger aus Gründen der Zeitersparnis ihr Brot dort kaufen.

weil die Ökokette "Natur pur", die auch Brot und Backwaren verkauft, nun auch in Wilmershausen einen Ökoladen eingerichtet hat.

weil die Bäckerei wegen eines Brandes im letzten Jahr lange geschlossen war.

weil nach der anfänglichen Begeisterung für die Produkte viele Bürger doch wieder bei dem Bäcker in ihrer Nähe kaufen.

Argumente für gestiegen:

weil die Bäckerei inzwischen auch einige Reformhäuser mit ihren Produkten beliefert.

weil die Bäckerei viele neue Kunden hinzugewonnen hat.

weil nach einer Umfrage im "Lokalkurier" von Wilmershausen im letzten Jahr den meisten Bürgern auf ökologischer Basis hergestelltes Brot besser schmeckt.

weil die Nachfrage nach Bioprodukten allgemein gestiegen ist.

weil die Bäckerei unlängst in einem anderen Ort ein weiteres Geschäft eröffnet hat.

weil das Angebot der Bäckerei vielfältiger geworden ist.

6. Hat das diesjährige Programm der Volkshochschule (VHS) Wilmershausen gegenüber dem des Vorjahres hinsichtlich seines Umfangs (Anzahl der Angebote) zugenommen oder abgenommen?

Argumente für gefallen:

weil die VHS allgemein sparen muß.

weil es für die VHS schwer geworden ist, in vielen Bereichen Dozenten zu finden.

weil die VHS aufgrund des Umbaus ihrer Unterrichtsräume Probleme hat, die Kurse räumlich unterzubringen.

weil in Bereichen wie Philosophie oder Religion die Nachfrage in den letzten Jahren so gering geworden ist, daß Kurse zu diesen Themen gestrichen wurden.

weil die VHS ihr Programm auf bestimmte Schwerpunkte konzentrieren will.

weil für manche Fächer Dozenten von außerhalb verpflichtet werden müssen, was für die VHS zunehmend teurer wird.

Argumente für gestiegen:

weil die VHS in diesem Jahr auch Kinofilme zeigt.

weil im Bereich Fremdsprachen die Nachfrage noch größer geworden ist.

weil die VHS in seinem Programm nun auch auf das gestiegene Interesse an alternativer Medizin und alternativen Heilverfahren reagiert hat.

weil die Nachfrage nach Kursen wie Kochen oder Töpfern so groß geworden ist, daß nun mehrere Kurse parallel angeboten werden.

weil der Trend in den letzten Jahren verstärkt dahin geht, daß immer mehr Menschen ihre Allgemeinbildung auffrischen oder verbessern wollen.

weil die VHS ihr Angebot um Themen wie Freizeitgestaltung nach frühzeitiger Rente erweitert hat.

7. Ist die Zahl der Schüler der Musikschule Wilmershausen gegenüber dem Vorjahr gestiegen oder gefallen?

Argumente für gefallen:

weil im vergangenen Jahr mehr Schüler als in den Jahren davor aufgehört haben.

weil die finanzielle Unterstützung durch die Stadt gekürzt worden ist.

weil bei den klassischen Instrumenten ein Rückgang der Schülerzahl zu verzeichnen ist.

weil immer weniger Kinder die musikalische Früherziehung besuchen.

weil viele Jugendliche lieber Musik auf CD hören als Musik selbst zu machen.

weil für bestimmte Bereiche des Instrumentalunterrichts qualifizierte Musiklehrer fehlen.

Argumente für gestiegen:

weil der Ruf der Musikschule durch aufsehenerregende Erfolge bei Nachwuchswettbewerben wie "Jugend musiziert" gestiegen ist.

weil durch öffentliche Vorspiele im letzten Jahr neue Schüler gewonnen werden konnten.

weil das Interesse am gemeinsamen Musizieren in Orchestern der Musikschule deutlich gestiegen ist.

weil die Musikschule nun auch eigene Chöre hat.

weil sich die Schülerzahl in der Sparte "Popmusik" erhöht hat.

weil der Einzugsbereich der Musikschule größer geworden ist.

8. Ist die Anzahl verkaufter Exemplare des "Wilmershausener Lokalkuriers" im letzten Jahr gestiegen oder gefallen?

Argumente für gefallen:

weil der Einzelpreis gestiegen ist.

weil der "Lokalkurier" Konkurrenz durch das neue "Wilmershausener Tageblatt" bekommen hat.

weil immer mehr Leute die Bildzeitung kaufen, wie Zeitungshändler im letzten Jahr registrierten.

weil es im letzten Jahr zu Protesten wegen offensichtlich nationalistischer Kommentare des Chefredakteurs gekommen ist.

weil Abonennten im letzten Jahr oft von Zustellungsproblemen betroffen waren.

weil die Anzahl kritischer Leserbriefe, in denen Leser ihre Unzufriedenheit mit der Zeitung äußern gestiegen ist.

Argumente für gestiegen:

weil die Anzahl der Abonnenten während des letzten Jahres gestiegen ist.

weil die Redaktion den Wünschen vieler Leser nach einem ausführlicheren Lokalteil nachgekommen ist.

weil der Kulturteil mit der Übersicht über Theater-, Kino- und Musikveranstaltungen verbessert wurde.

weil der Anzeigenteil umfangreicher gestaltet wurde und die Aufgabe von Anzeigen günstiger geworden ist.

weil die Zeitung durch das verbesserte Layout übersichtlicher und ansprechender aussieht.

weil die Zeitung jetzt auch in anderen Orten des Landkreises erscheint.

9. Der Möbelmarkt Heinemann bietet preiswerte Möbel an, die man leicht selbst zusammenbauen kann. Ist der Umsatz des Möbelmarktes in den letzten zwei Jahren gestiegen oder gefallen?

Argumente für gefallen:

weil in den letzten zwei Jahren bei den Einzelteilen der Möbel häufig Verarbeitungsfehler vorkamen.

weil viele Kunden anspruchsvoller geworden sind und exklusivere Möbelstücke suchen.

weil der Lagerbestand an Möbeln in den letzten zwei Jahren oftmals sehr gering war, so daß es bei den Lieferungen zu langen Wartezeiten für die Kunden kam.

weil die Konkurrenz durch andere Möbelhäuser gestiegen ist.

weil sich die Klagen über die "chaotische Organisation" des Möbelmarkts in den letzten Jahren gehäuft haben.

weil der Markt wegen Problemen beim Import von billigem Holz zeitweise mit den Produkten nicht nachgekommen ist.

Argumente für gestiegen:

weil die Anzahl der Studierenden, die in Wilmershausen wohnen und günstig Möbel brauchen, gestiegen ist.

weil die Möbel teilweise noch preiswerter geworden sind.

weil die Nachfrage nach einfachen, leicht aufzustellenden Regalen gestiegen ist.

weil kleinere Unternehmen und Betriebe vermehrt ihre Büromöbel dort kaufen.

weil die Auswahl größer geworden ist.

weil Beratung und Kundenservice in den letzten zwei Jahren erheblich verbessert worden sind.

10. Ist der Umsatz des Computergeschäfts "COMPISOFT" in den letzten zwei Jahren gestiegen oder gefallen?

Argumente für gefallen:

weil das Geschäft im Vergleich mit den steigenden Werbeaufwendungen der großen Computergeschäfte nicht auffällt.

weil die Konkurrenz durch immer mehr kleinere Computer-Shops beständig größer geworden ist.

weil bekannte Computerhändler Geschäfte direkt im Einkaufszentrum, d.h. in günstigerer Lage eröffnet haben.

weil es durch längere Lieferzeiten immer häufiger zu Verzögerungen gekommen ist.

weil es in den letzten zwei Jahren Probleme mit fehlerhaften Geräten gab.

Argumente für gestiegen:

weil sich immer mehr Leute für Computer interessieren.

weil sich die Beratung im Vergleich zu den großen Geschäftsketten beständig verbessert hat und man sich mehr Zeit für die Kunden nimmt.

weil durch die Einführung eines mobilen Reparaturdienstes viele Kunden hinzugewonnen wurden.

weil COMPISOFT die Rechner nun auch dem Privatmann direkt nach Hause liefert.

weil COMPISOFT seit zwei Jahren auch Software für Unternehmen liefert.

weil die Firma angefangen hat, Computer-Einführungskurse anzubieten.

11. Die berühmte "Wilmerhäuser Buchmesse" findet jedes Jahr statt. Ist die Anzahl an Besuchern dieses Jahr gegenüber dem vergangenen Jahr gestiegen oder gefallen?

Argumente für gefallen:

weil das Programm diesmal pointierter, aber gleichzeitig auch reduzierter ausgefallen ist.

weil die Eintrittspreise allgemein erhöht worden sind.

weil es für Studenten und Senioren diesmal keinen ermäßigten Eintritt gegeben hat.

weil der diesjährige Themenschwerpunkt "fernöstliche Literatur" auf nicht so breites Interesse gestoßen ist.

weil die Messe zeitgleich mit anderen bedeutenden Veranstaltungen in der Region stattgefunden hat.

weil viele Besucher in den letzten Jahren darüber geklagt haben, daß ihnen die Messe zu unübersichtlich organisiert sei und ihnen die Orientierung fehle.

Argumente für gestiegen:

weil die Leute allgemein wieder mehr lesen.

weil dieses Mal auch bekannte deutsche Autoren ihr Kommen zugesagt haben.

weil die Messe überregionale Bedeutung bekommen hat.

weil in ganz Deutschland für die Messe geworben wurde.

weil die Messe jetzt für die ganze Familie etwas bietet; so sind jetzt auch Kinderbuchautoren vertreten.

weil die Messe dieses Jahr ein Medienspektakel zu werden versprach.

12. Ist die Arbeitslosigkeit in der Region um Wilmershausen im letzten Jahr gegenüber dem Vorjahr gestiegen oder gefallen?

Argumente für gefallen:

weil aufgrund des größeren Stellenangebots in bestimmten Bereichen wie dem Handwerk wieder Arbeitsplätze vermittelt werden konnten.

weil es in bestimmten Bereichen dazu gekommen ist, daß sich Angestellte eine Arbeitsstelle teilen.

weil sich einige Betriebe zusammengeschlossen haben, um wirtschaftlich konkurrenzfähig zu bleiben und so viele Arbeitsplätze wie möglich zu erhalten.

weil neue Unternehmen sich in der Region niedergelassen haben.

weil im letzten Jahr ein neuer Erlebnispark mit vielen neuen Arbeitsplätzen in Betrieb genommen worden ist.

weil neue Einrichtungen und Arbeitsplätze für Langzeitarbeitslose gefunden werden konnten, z.B. im Lagerbereich.

Argumente für gestiegen:

weil die in Wilmershausen ansässige Schuhfabrik wegen Sparmaßnahmen einzelne Abteilungen geschlossen hat.

weil die Zahl der Mitarbeiter bei der Stadt reduziert worden ist.

weil ein Zulieferbetrieb für einen großen Automobilkonzern pleite gegangen ist, da der Konzern sich andere Zulieferer gesucht hat.

weil sich die wirtschaftliche Lage im Einzelhandel verschlechtert hat.

weil es im letzten Jahr weniger Großaufträge für das Baugewerbe gab und es zu Stellenkürzungen gekommen ist.

weil das kleine Kaufhaus "IDEAL" wegen Konkurrenz durch neue Supermärkte geschlossen worden ist.

13. *Das Antiquariat "Heinrich Müller" bietet alte Bücher und Restauflagen sowie Kunststiche und Kunstpostkarten an. Lief der Verkauf im vergangenen Jahr besser oder schlechter als im Jahr davor?*

Argumente für gefallen:

- weil alte Schul- und Sachbücher besonderes schlecht gegangen sind.
- weil sich zunehmend Kunden darüber beschwert haben, daß die Preise überhöht sind.
- weil teurere "Sammlerstücke", für die das Antiquariat bekannt ist, immer seltener gekauft worden sind.
- weil das Gesamtangebot im letzten Jahr gegenüber dem Vorjahr reduziert war, da weniger brauchbare Bücher hereingekommen sind.
- weil alte Stiche, Bilder und Kunstpostkarten kaum beachtet worden sind.
- weil das Antiquariat im letzten Jahr in ein Gebäude außerhalb der Innenstadt umgezogen ist.

Argumente für gestiegen:

- weil besonders im letzten Jahr immer mehr Studenten das Antiquariat aufsuchten, um günstig an Restexemplare von Lehrbüchern älterer Auflagen zu kommen, die gerade neu aufgelegt worden sind.
- weil das Interesse an älteren Werkausgaben namhafter Autoren wegen einer großen literaturwissenschaftlichen Vortragsreihe der Stadtbibliothek im letzten Jahr gestiegen ist.
- weil das Antiquariat Neuland betreten hat und seit dem letzten Jahr auch auf Buchbörsen mit einem Stand vertreten ist.
- weil der Zustand der neu angekauften antiquarischen Bücher im letzten Jahr noch besser geworden ist.
- weil sich das Antiquariat seit dem letzten Jahr auf alte Comics spezialisiert hat.
- weil das Antiquariat seit dem letzten Jahr auch ganze Jahrgänge alter Zeitschriften im Angebot hat.

14. Ist der Umsatz der Wilmershausener Schokoladen- und Keksfabrik "Schokolina" im vergangenen Jahr im Vergleich zum Jahr davor gestiegen oder gefallen?

Argumente für gefallen:

weil der Verkauf zu Ostern im letzten Jahr gegenüber den Jahren zuvor stark abgenommen hat.

weil die Leute allgemein bewußter leben und weniger Süßigkeiten essen.

weil es im letzten Jahren Probleme mit der Haltbarkeit der Kekse gab und sich viele Kunden darüber beschwert haben, daß die Kekse schon nach kurzer Zeit trocken waren.

weil viele Konsumenten in der letzten Zeit doch eher wieder Produkte mit bekannten Handelsnamen kaufen.

weil die Verpackungen anderer Produkte immer ansprechender geworden sind, während die Fabrik ihre Produkte in einfachen Tüten mit Etikett verkauft.

weil der Trend eher von den Keksen weggeht; die Leute essen lieber Gummibärchen.

Argumente für gestiegen:

weil Weihnachten des letzten Jahres deutlich mehr Produkte verkauft worden sind als zu den Weihnachtsfesten der Jahre davor.

weil die Fabrik nun auch Gebäck und Schokolade für Diabetiker herstellt.

weil das Sortiment um neue, leckere Kekssorten erweitert worden ist.

weil es seit einem Jahr "Vorratspackungen" zu einem günstigen Preis gibt.

weil sich die teuren Pralinenmischungen im letzten Jahr so gut wie nie zuvor verkauft haben.

weil die Fabrik jetzt auch einen Extra-Laden in der Innenstadt eingerichtet hat.

15. Ist der Umsatz des Getränkehandels Jakobson im letzten Jahr gestiegen oder gefallen?

Argumente für gefallen:

weil im letzten Jahr bei einer großen Lieferung der Biersorte "Olympia Pils" das Verfallsdatum überschritten war.

weil das vorher außergewöhnlich umfassende und vielfältige Sortiment aus finanziellen und Lagerplatzgründen begrenzt werden mußte.

weil die Geschäftskette "TOP Markt" als ein Hauptabnehmer abgesprungen ist.

weil die Marke "Timos tollster Fruchtsaft" boykottiert wurde, da der Hersteller rechtsradikale Gruppierungen unterstützen soll.

weil sich bei einer Erhebung über alle Getränke- und Supermärkte in der Region zeigte, daß der Getränkekonsum pro Kopf im letzten Jahr zurückgegangen ist.

weil in der Region einige neue Getränkemärkte eröffnet worden sind.

Argumente für gestiegen:

weil es im letzten Jahr viel mehr Feierlichkeiten (wie das Stadtgründungsfest) als in den Jahren zuvor gab, für die der Getränkehandel die Getränke stellte.

weil der Getränkehandel nun als Sponsor auf den Trikots der Fußballmannschaft des TSV Wilmershausen vertreten ist.

weil der Getränkehandel nun auch Alters- und Studentenwohnheime beliefert.

weil seit dem letzten Jahr eine "Party-Aktion" läuft, in deren Rahmen es Mengenrabatt beim Kauf vieler Getränke gibt.

weil der Getränkehandel ab dem letzten Jahr sämtliche Kneipen in der Region mit nichtalkoholischen Getränken versorgt.

weil immer mehr Vereine in Wilmershausen ihre Getränke für Feste beim Getränkehandel holen.

16. Ist die Zahl der Bewohner des Wohnstifts "Lindenhof" in den letzten beiden Jahren gestiegen oder gefallen?

Argumente für gefallen:

weil sich in der Stadt immer mehr Wohngemeinschaften älterer Menschen gebildet haben.

weil immer mehr Leute ihre Eltern, wenn sie alt werden, daheim pflegen.

weil immer mehr Menschen im Alter unabhängig sein und selbständig leben wollen.

weil in den letzten beiden Jahren neue Wohnstifte in schönerer Umgebung (z.B. Wald in der Nähe) gebaut worden sind, während die Umgebung des "Lindenhofs" eher eintönig ist.

weil die Zahl der Todesfälle in den letzten beiden Jahren gestiegen ist.

weil sich wegen der gestiegenen Kosten immer weniger Leute einen Platz im Wohnstift leisten können.

Argumente für gestiegen:

weil das Wohnstift durch eine verbesserte medizinische Betreuung für sich werben kann.

weil die Wohnungen in den letzten beiden Jahren schöner gestaltet worden sind.

weil immer mehr Leute über die gute Stimmung unter den Pflegepersonen und Bewohnern berichten.

weil für die Bewohner attraktive Angebote wie z.B. ein Literaturkreis oder eine Gesprächsrunde eingerichtet worden sind.

weil der Wunsch nach geselligem Zusammensein und die Angst vor sozialer Isolierung immer größer wird.

weil das Wohnstift im letzten Jahr ausgebaut wurde und nun mehr Wohneinheiten zur Verfügung stehen.

17. Ist die Anzahl der Ratsuchenden bei der ELFE (Erziehungs-, Lebens-, Familien- und Eheberatung) im letzten Jahr gegenüber dem Vorjahr gestiegen oder gefallen?

Argumente für gefallen:

weil die Zahl von Kindern mit Sprachstörungen als Anlaß zum Aufsuchen der Beratungsstelle weiter abgenommen hat.

weil mehr Beratungsgespräche pro KlientIn notwendig geworden sind.

weil sexuelle Probleme im letzten Jahr seltener Anlaß für eine Ehe- oder Lebensberatung waren.

weil die Wartelisten gerade im letzten Jahr völlig überfüllt waren, so daß "neue" Ratsuchende lange auf eine Beratung warten mußten.

weil sich die Anzahl von Kindern und Jugendlichen, die zur Erziehungsberatung mit Problemen bei Arbeit- und Leistungsanforderungen gekommen sind, reduziert hat.

weil im letzten Jahr die Anzahl der Ratsuchenden im Rahmen der Lebensberatung gegenüber dem Vorjahr deutlich abgenommen hat.

Argumente für gestiegen:

weil die Stellenzahl der Mitarbeiter erhöht worden ist.

weil die Anzahl der behandelten Kinder im letzten Jahr deutlich gestiegen ist.

weil die Bahndlungsintensität (gemessen in Beraterstunden pro Fall) gefallen ist.

weil die Anzahl der ratsuchenden Paare zugenommen hat.

weil mehr ältere Menschen wegen Einsamkeit die Beratungsstelle aufsuchten.

weil gegenüber dem Vorjahr im letzten Jahr viel mehr Kinder und Jugendliche wegen Genußmittel-, Drogen- und Medikamentenmißbrauch behandelt worden sind.

18. Ist der Gewinn des Hotels "Excelsior" im letzten Jahr gestiegen oder gefallen?

Argumente für gefallen:

weil die Preise für eine Übernachtung stark gestiegen sind.

weil das Hotel immer mehr auf Prominente ausgerichtet ist; einfache Hotelzimmer gibt es nicht mehr.

weil durch den Bau des City-Hotels, der im letzten Jahr abgeschlossen wurde und sich direkt in der Innenstadt befindet, starke Konkurrenz aufgekommen ist.

weil wegen eines großen Umbaus im letzten Jahr Teile des Hotels geschlossen werden mußten.

weil die Zahl der Touristen und Besucher der Stadt von außerhalb im letzten Jahr zurückgegangen ist.

weil der Ruf des Hotels unter einer Serie von Diebstählen im letzten Jahr gelitten hat.

Argumente für gestiegen:

weil im letzten Jahr mehrere internationale Kongresse in Wilmershausen stattgefunden haben.

weil seit dem letzten Jahr eine Fernsehserie in dem Hotel gedreht wird.

weil der Service verbessert wurde.

weil der Ruf des Hotels v.a. wegen der exquisiten Küche enorm gestiegen ist.

weil im letzten Jahr viele Politiker dort abgestiegen sind.

weil ein Veranstaltungssaal angegliedert worden ist: Seit dem letzten Jahr treten dort bekannte Künstler auf.

A.2 Sachverhalte zu Experiment 3 - 5

1. 1988 glaubten 88 % der amerikanischen Erwachsenen, das Recht zu haben, ihre Kinder zu schlagen. Dieser Prozentsatz hat sich bis zum jetzigen Zeitpunkt verändert.

Argumente für gefallen:

1. weil die autoritäre Erziehung vor allem wegen ihrer Rigidität zunehmend kritisiert wird.
2. weil Werbekampagnen in den Medien über die Folgen von Gewalterziehung aufklären.
3. weil in den Medien zunehmend über Mißhandlungen in der Familie berichtet wird.
4. weil Gewalt in der Familie zu einem wichtigen Thema des Weltkindertages wurde.

Argumente für gestiegen:

1. weil die Prämissen der antiautoritären Erziehung angezweifelt werden. Z.B.: die angeblich notwendige und ungehinderte Selbstentfaltung des Kindes.
2. weil autoritär ausgerichtete Glaubensgemeinschaften in den USA wieder mehr Zulauf bekommen.
3. weil strenge und konservative Colleges in den USA an Popularität gewinnen.
4. weil die Zusammenhänge zwischen Jugendkriminalität und den fehlenden aber notwendigen Grenzen in der „Laissez-faire“-Erziehung immer häufiger diskutiert werden.

2. 1975 gaben 6 % aller Frauen zwischen 22 und 35 Jahren Liebe als Hauptmotiv für Sex an. Dieser Prozentsatz hat sich bis zum jetzigen Zeitpunkt verändert.

Argumente für gefallen:

1. weil Liebe als heres Motiv („moralisch einwandfreie" Begründung) für Sex entmythologisiert worden ist und ein moralischer Begründungsbedarf entfällt.
2. weil viele öffentliche Diskussionen über Sex zur Konsequenz haben, daß Spaß am Sex immer mehr als ausreichendes Motiv anerkannt wird.
3. weil der Anteil weiblicher Kunden in Sexshops in den letzten Jahren deutlich gestiegen ist.
4. weil die zunehmende Emanzipation der Frau ihr eine freiere Einstellung zum Sex ermöglicht.

Argumente für gestiegen:

1. weil der reine Sex ohne Liebe in Verruf gekommen ist: Er wird oft als leer empfunden.
2. weil im Zeitalter von AIDS die Gefahren des Sex sehr deutlich vor Augen liegen.
3. weil Werte wie Treue in der Partnerschaft wieder einen höheren Stellenwert gefunden haben.
4. weil die in den Siebzigern propagierte „freie Liebe“ heute kein Thema mehr ist.

3. Im Jahr 1970 ermittelte das BKA, daß 53 % aller Kriminellen mit Haftstrafen zwischen 3 und 18 Monaten erneut straffällig geworden waren. Dieser Prozentsatz hat sich bis zum jetzigen Zeitpunkt verändert.

Argumente für gefallen:

1. weil neue Formen der Haft, weg von der Aufbewahrung, Erfolg zeigen.
2. weil auf soziale Dienste und Betreuung, auch in den Haftanstalten, mehr Wert gelegt wird,
3. weil die Sozialdienste der Haftanstalten sich zunehmend um berufliche Wiedereingliederung bemühen.
4. weil durch die Möglichkeit, als Freigänger das Gefängnis schon während der Haftzeit zeitweise zu verlassen, soziale Kontakte aufrecht erhalten und eine Isolation vermieden werden können.

Argumente für gestiegen:

1. weil Kleinkriminalität und Delikte, die zu kurzen Haftstrafen führen, allgemein stark zunehmen.
2. weil die Maßnahmen sozialer Dienste in der Haftanstalt oft auf langfristige Veränderungen ausgelegt sind und bei Insassen mit kurzen Haftstrafen gerade in der kurzen Zeit nicht greifen können.
3. weil der Kontakt zum alten Milieu durch Maßnahmen wie Freigang noch gefördert wird. So kann es auch schon während der Haftzeit zu weiteren Delikten kommen.
4. weil die Bewährungsaufgaben in den letzten Jahren verschärft wurden und sich dadurch mehr Konflikte mit dem Gesetz ergeben.

4. In einer Studie aus dem Jahre 1950 wählten 65 % der Jungen geschlechtstypisches Spielzeug. Dieser Prozentsatz hat sich bis zum jetzigen Zeitpunkt verändert.

Argumente für gefallen:

1. weil der Erziehungsstil sich verändert hat, und nicht mehr so viel Wert auf geschlechtstypisches Spielzeug gelegt wird.
2. weil die im Zuge der Emanzipation angefachte Diskussion über geschlechtsspezifische Rollen die Aufmerksamkeit zunehmend auf die rollenbildende Funktion des Spielzeugs gelenkt hat.
3. weil auch außerhalb des Elternhauses, z. B. in Kindergärten, heute geschlechtsuntypische Rollen stärker gefördert werden.
4. weil viele Eltern mit ökologisch-alternativer Einstellung bewußt auch geschlechtsuntypisches Spielzeug kaufen.

Argumente für gestiegen:

1. weil Studien belegen, daß bei freier Auswahl des Spielzeugs Jungen und Mädchen häufiger zu geschlechtstypischem Spielzeug greifen, unabhängig von ihrer Erziehung.
2. weil die erneute Zunahme konservativer Werte in der Gesellschaft zu einem Trend pro geschlechtstypischem Spielzeug führt.
3. weil die zunehmenden Leistungsanforderungen in der Gesellschaft zu einer Überbetonung männlich aggressiver Werte und so auch zu einer Aufwertung geschlechtstypischen Spielzeugs führen.
4. weil die Werbung, vor allem in Funk und Fernsehen, wieder stark auf geschlechtstypisches Spielzeug setzt.

5. Im Jahre 1970 betrug der Anteil an Anrufern bei der Telefonseelsorge, die zwischen 20 und 30 Jahren alt sind, 14 Prozent. Dieser Prozentsatz hat sich bis zum jetzigen Zeitpunkt verändert.

Argumente für gefallen:

1. weil die Möglichkeiten, sich seit den Siebzigern neugebildeten Selbsthilfegruppen anzuschließen, gerade von der Altersgruppe zwischen 20 und 30 Jahren genutzt wird.
2. weil eine Themenverlagerung stattfindet. Themen wie Sexualität, die gerade die Altersgruppe zwischen 20 und 30 Jahren betreffen und früher oft bei der Telefonseelsorge angesprochen wurden, jetzt bevorzugt in längeren persönlichen Beratungen besprochen werden.
3. weil die Diskussion über soziale Dienste dafür gesorgt hat, daß neben der Telefonseelsorge Alternativen zur Beratung bekannt geworden sind.
4. weil gerade in der Altersgruppe zwischen 20 und 30 Jahren der Telefonseelsorge der Ruf von Antiquiertheit anhaftet.

Argumente für gestiegen:

1. weil die Möglichkeit, die Anonymität zu wahren, die Attraktivität der Telefonseelsorge gegenüber anderen Diensten erhöht hat.
2. weil die starke Betonung materieller Werte in der Gesellschaft und der zunehmende Konkurrenzdruck zu einer starken Zunahme an Problemen, gerade in der Altersgruppe zwischen 20 und 30 Jahren, geführt hat.
3. weil die Angst vor Vorurteilen gegenüber der Inanspruchnahme psychosozialer Dienste abgenommen hat.
4. Der Prozentsatz ist gestiegen, weil die Telefonseelsorge von Werbekampagnen der Schwulenbewegung für das „rosa Telefon“ profitiert.

6. Der prozentuale Anteil von ökologisch produzierten Lebensmitteln am Gesamtnahrungsmittelangebot, also das Verhältnis ökologisch zu konventionell produzierten Lebensmitteln, hat sich von 1990 bis 1994 geändert.

Argumente für gefallen:

1. weil die Preise für Bioprodukte noch bedeutend höher sind.
2. weil das Vertriebssystem für Bioprodukte, gerade im Vergleich zum immer besser ausgebautem Vertrieb konventioneller Produkte, noch schlecht entwickelt ist.
3. weil die Verbraucher durch falsche „Bioprodukte“ verunsichert wurden.
4. weil die Angebotsmenge an nicht-ökologischen, konventionell produzierten Lebensmitteln gestiegen ist, und damit sich der Anteil ökologischer Produkte vermindert haben könnte.

Argumente für gestiegen:

1. weil die Angebotsvielfalt an ökologisch produzierten Lebensmitteln sich erhöht hat.
2. weil die Tierseuchenskandale der letzten Zeit die Aufmerksamkeit auf Risiken der industriellen Produktionsweise von Lebensmitteln lenken.
3. weil die Nachfrage nach Bioprodukten allgemein gestiegen ist.
4. weil durch die EU-Subventionspolitik immer mehr Bauern auf ökologische Produktionsweise umsteigen.

7. Im Jahre 1970 nahmen 22 % der Adoptionswilligen in der BRD ein Kind aus dem außereuropäischen Raum auf. Dieser Prozentsatz hat sich bis zum jetzigen Zeitpunkt geändert.

Argumente für gefallen:

1. weil die Ausländerfeindlichkeit allgemein zugenommen hat.
2. weil die neuen Bundesländer mit ihrem höheren Potential an Ausländerfeindlichkeit dazugekommen sind.
3. weil die Probleme, auf die Ausländer\Innen in Deutschland treffen, Adoptionswillige abschrecken, die sonst gern ein ausländisches Kind adoptiert hätten. Sie fürchten sich vor späteren Konsequenzen für ihr Kind.
4. weil die Kosten für eine Adoption aus fernen Ländern in Zeiten wirtschaftlicher Rezession wieder eine stärkere Rolle spielen.

Argumente für gestiegen:

1. weil bekanntgewordene Beispiele von prominenten Schauspieler\Innen mit Adoptivkindern wie Witta Pohl (Drombuschs) einen positiven Effekt auf die allgemeine Bereitschaft zur Adoption ausländischer Kinder ausüben.
2. weil die Zahl unfreiwillig kinderlos Gebliebener steigt. Gleichzeitig steigt die Wartezeit bei der Adoption europäischer Kinder.
3. weil viele Bürger\Innen in der Adoption eines außereuropäischen Kindes einen Beitrag gegen die zunehmende Ausländerfeindlichkeit sehen.
4. weil die Berichterstattung aus von Katastrophen betroffenen Gebieten höhere Bereitschaft zu helfen weckt.

8. Wie hat sich die Zufriedenheit der Frauen bezüglich ihrer gesellschaftlichen Stellung in Deutschland zwischen 1978 und 1993 verändert?

Argumente für gefallen:

1. Die Zufriedenheit hat abgenommen, weil es immer noch zuwenig Initiativen gibt, die der Frau erlauben, Familie und Beruf zu vereinbaren.
2. Die Zufriedenheit hat abgenommen, weil in der Berufswelt die Normen immer noch durch Männer geprägt werden, so daß sich die Frau für ihre Karriere anpassen muß.
3. Die Zufriedenheit hat abgenommen, weil Frauen für die gleiche Arbeit immer noch weniger Geld als Männer erhalten.
4. Die Zufriedenheit hat abgenommen, weil der rapide Abbau von sozialen Einrichtungen in Ostdeutschland zu einer Mehrbelastung der Familien führt, die in den allermeisten Fällen auf die Frau zurückfällt.

Argumente für gestiegen:

1. Die Zufriedenheit hat zugenommen, weil der technische Fortschritt sowie das größere Angebot an Fertiggeräten die Haushaltsführung erleichtern.
2. Die Zufriedenheit hat zugenommen, weil Frauen Errungenschaften wie zum Beispiel die Quotenregelung erreicht haben, die in einigen Parteien Anwendung findet.
3. Die Zufriedenheit hat zugenommen, weil moderne Männer keine festen Rollenvorstellungen mehr haben und zunehmend bereit sind, neue Aufgaben zu übernehmen.
4. Die Zufriedenheit hat zugenommen, weil sich die Sensibilität für Frauenfragen in der Gesellschaft erhöht hat.

9. Die durchschnittliche Anzahl der tatsächlich mit dem Fahrrad zurückgelegten Kilometer pro Kopf, hat sich zwischen 1973 und 1993 geändert.

Argumente für gefallen:

1. Die Anzahl hat sich verringert, weil das Streckennetz der öffentlichen Verkehrsmittel in den Städten als bequeme Alternative zum Fahrrad besser ausgebaut wurde.
2. Die Anzahl hat sich verringert, weil immer mehr Menschen ein eigenes Auto besitzen und das auch benutzen.
3. Die Anzahl hat sich verringert, weil Aufgrund des gestiegenen Verkehrsaufkommens eine höhere Unfallgefahr für Fahrradfahrer besteht, die viele abschreckt.
4. Die Anzahl hat sich verringert, weil Aufgrund des Smogs und der hohen Ozonwerte viele Menschen körperliche Belastung in der Stadt vermeiden.

Argumente für gestiegen:

1. Die Anzahl hat sich erhöht, weil das Wissen um die Bedeutung des Sports für die Gesundheit gestiegen ist.
2. Die Anzahl hat sich erhöht, weil man in der Großstadt aufgrund des hohen Verkehrsaufkommens mit dem Auto nicht mehr mobil ist.
3. Die Anzahl hat sich erhöht, weil das Fahrradfahren sich zum beliebten und modischen Freizeitsport entwickelt hat.
4. Die Anzahl hat sich erhöht, weil Fahrräder heute eine bessere Qualität haben und sich besser fahren.

10. Der Anbau von Coca-Pflanzen in Kolumbien hat sich von 1980 bis heute geändert. Ist er gestiegen oder gefallen?

Argumente für gefallen:

1. Der Coca-Anbau ist gefallen, weil die Drogenkampagnen ihre Auswirkungen auch in Kolumbien zeigen.
2. Der Coca-Anbau ist gefallen, weil die USA sehr starken Druck auf Länder wie Kolumbien ausübt, den Drogenanbau zu bekämpfen.
3. Der Coca-Anbau ist gefallen, weil die Drogenprozesse in Medellin und Erfolge in der Bekämpfung der Mafia im Land ihre positiven Effekte zeigen.
4. Der Coca-Anbau ist gefallen, weil die Aufklärung der indianischen Bevölkerung, die traditionellerweise Coca anbaut und konsumiert, über die gesundheitlichen Gefahren des Konsums ihre Wirkung zeigt.

Argumente für gestiegen:

1. Der Coca-Anbau ist gestiegen, weil der Verkauf von Cocain mittlerweile zum wichtigsten Wirtschaftsfaktor des Landes wächst.
2. Der Coca-Anbau ist gestiegen, weil die Drogenkartelle ihre Macht immer mehr gefestigt haben.
3. Der Coca-Anbau ist gestiegen, weil die Beliebtheit von Cocain gegenüber anderen Drogen ungebrochen ist.
4. Der Coca-Anbau ist gestiegen, weil gerade in einer Gesellschaft mit hohen Leistungsanforderung der Konsum von Cocain, wegen seiner aufputschenden Wirkung, gesteigert ist.

11. Nach statistischen Erhebungen starben im Jahre 1985 in Deutschland 8400 Menschen durch Verkehrsunfälle. Die Anzahl der jährlichen Verkehrstoten hat sich bis 1993 geändert.

Argumente für gefallen:

1. Die Anzahl hat sich verringert, weil die Sicherheitsstandards bei Autos gestiegen sind.
2. Die Anzahl hat sich verringert, weil der Einsatz des Airbags erste Ergebnisse zeigt.
3. Die Anzahl hat sich verringert, weil die Sicherheitssitzpflicht für Kleinkinder eingeführt wurde.
4. Die Anzahl hat sich verringert, weil bei der Verkehrsplanung in immer stärkerem Maße Sicherheitsgesichtspunkte berücksichtigt werden.

Argumente für gestiegen:

1. Die Anzahl hat sich erhöht, weil Laut Polizeistatistiken die Zahl alkoholisierter Fahrer\Innen gestiegen ist.
2. Die Anzahl hat sich erhöht, weil auch die Anzahl an Spritztouren Minderjähriger mit zumeist gestohlenen Fahrzeugen (sogenannte „joy-rides“) sich erhöhte.
3. Die Anzahl hat sich erhöht, weil der Anteil von Kleinwagen mit geringeren Unfallschutz, verglichen mit Wagen der Mittel- bzw. der gehobenen Klasse, an der gesamten Autoanzahl gestiegen ist.
4. Die Anzahl hat sich erhöht, weil nach Untersuchungen und Umfragen die Aggressivität im Straßenverkehr insgesamt zugenommen hat.

12. Die Anzahl der von Jugendlichen (Alter von 12 bis 18 Jahren) im Jahr durchschnittlich gelesenen Romane hat sich von 1975 bis 1993 geändert.

Argumente für gefallen:

1. Die Anzahl hat sich verringert, weil der zunehmende Leistungsdruck in der Schule zu Einschränkungen in der Freizeit führt.
2. Die Anzahl hat sich verringert, weil die Konkurrenz durch andere Medien, vor allem Fernsehen und Computer, sich stark erhöht hat.
3. Die Anzahl hat sich verringert, weil laut Umfragen viele Jugendliche das Lesen als zu zeitintensiv ansehen.
4. Die Anzahl hat sich verringert, weil die zunehmende Amerikanisierung der Jugendkultur zur Schwerpunktverlagerung der Freizeitgestaltung hin zu eher sportlicher Betätigung führt.

Argumente für gestiegen:

1. Die Anzahl hat sich erhöht, weil Die Verkaufszahlen für Literatur sich insgesamt erhöht haben.
2. Die Anzahl hat sich erhöht, weil die Werbestrategien des Buchmarktes aggressiver, vielfältiger und zielgruppenorientierter geworden sind.
3. Die Anzahl hat sich erhöht, weil der zunehmend freiere Unterrichtsstil in den Schulen, vor allem im Deutschunterricht, bei Jugendlichen verstärkt das Interesse für Literatur weckt.
4. Die Anzahl hat sich erhöht, weil gestiegene Besucherzahlen z.B. bei der Frankfurter Buchmesse für ein allgemein stärkeres Interesse an Büchern sprechen.

13. Der Anteil der deutschen Bevölkerung, der sich ein Eigentumshaus wünscht, hat sich zwischen 1960 und 1993 verändert.

Argumente für gefallen:

1. Der Anteil hat sich verringert, weil die Zahl der Single-Haushalte steigt. Diese bevorzugen kleine Mietwohnungen bzw. Appartements.
2. Der Anteil hat sich verringert, weil die Grundstückspreise in Relation zum Einkommen viel stärker gestiegen sind.
3. Der Anteil hat sich verringert, weil die immer stärker geforderte Flexibilität der Arbeitnehmer\Innen oft einen Ortswechsel verlangt und viele von dem Kauf oder Bau eines Eigenheims abhält.
4. Der Anteil hat sich verringert, weil der Mangel an Grundstücken die Attraktivität des Angebots stark einschränkt.

Argumente für gestiegen:

1. Der Anteil hat sich erhöht, weil die steigenden Mietpreise den Hausbau bzw. -kauf immer günstiger erscheinen lassen.
2. Der Anteil hat sich erhöht, weil der Wunsch, die eigenen vier Wände zu gestalten, zugenommen hat. Den Gestaltungsmöglichkeiten und persönlichen Freiheiten sind in Mietwohnungen oft zu enge Grenzen gesetzt.
3. Der Anteil hat sich erhöht, weil die Mietverhältnisse zunehmend unpersönlicher werden (Wohnungsbaugesellschaften; keine direkte Möglichkeit, den Vermieter anzusprechen).
4. Der Anteil hat sich erhöht, weil die zunehmende Bevölkerungsdichte in den Städten für eine zunehmende Attraktivität von Baugebieten in Randbezirken sorgt.

14. Der Anteil der Urlauber\Innen deutscher Nationalität auf Mallorca hat sich seit 1985 verändert.

Argumente für gefallen:

1. Der Anteil hat sich verringert, weil der Ruf Mallorcas als „Putzfraueninsel“ viele Urlauber\Innen abschreckt.
2. Der Anteil hat sich verringert, weil die Insel den Ruf hat, stark von Touristen überlaufen zu sein.
3. Der Anteil hat sich verringert, weil der Trend zu „Abenteuerurlaub“ und Fernreisen steigt.
4. Der Anteil hat sich verringert, weil Mallorca als Urlaubsziel in Deutschland bereits so bekannt ist, daß andere, unbekanntere Alternativen bevorzugt werden.

Argumente für gestiegen:

1. Der Anteil hat sich erhöht, weil eine Studie der Lufthansa belegt, daß die Preise der Flüge von Deutschland in Richtung Mallorca allgemein gesunken sind.
2. Der Anteil hat sich erhöht, weil die Anzahl und Qualität der Hotel- und Pauschalangebote deutscher Reiseunternehmer für Mallorca steigen.
3. Der Anteil hat sich erhöht, weil dem Trend zu Kurztrips und Wochenendreisen für Familien, Mallorca als Ziel z.B. wegen geringer Anreisezeiten besonders entgegenkommt.
4. Der Anteil hat sich erhöht, weil das vielfältige Programmangebot auf Mallorca, mit zahlreichen Diskotheken, Ferienclubs usw. besonders im internationalen Vergleich, den Erwartungen der deutschen Urlauber\Innen entgegenkommt.

15. Der durchschnittliche Alkoholkonsum pro Kopf auf von skandinavischen Ländern aus fahrenden Fähren hat sich seit 1988 verändert.

Argumente für gefallen:

1. Der Alkoholkonsum hat sich verringert, weil Aufgrund des gestiegenen Gesundheitsbewußtseins auf den Konsum alkoholischer Getränke verzichtet wird.
2. Der Alkoholkonsum hat sich verringert, weil erhöhte Sicherheitsmaßnahmen auf den Fähren den Ausschank von Alkohol einschränken.
3. Der Alkoholkonsum hat sich verringert, weil der Schock über die Fährunglücke in der letzten Zeit zur Selbsteinschränkung der Passagiere führt, um im Notfall handlungsfähig sein zu können.
4. Der Alkoholkonsum hat sich verringert, weil die Struktur der Urlauber auf den Fähren sich dahingehend geändert hat, daß anteilmäßig mehr Familien und minderjährige Jugendgruppen reisen.

Argumente für gestiegen:

1. Der Alkoholkonsum hat sich erhöht, weil die Alkoholsteuern auf dem Festland gestiegen sind.
2. Der Alkoholkonsum hat sich erhöht, weil durch die nach wie vor stark reglementierte Ausgabe von Alkohol in skandinavischen Ländern die Attraktivität des freien Alkoholkonsums auf den Fähren fördert.
3. Der Alkoholkonsum hat sich erhöht, weil der Anteil besonders preisgünstiger baltischer Fähren zugenommen hat.
4. Der Alkoholkonsum hat sich erhöht, weil der Alkoholkonsum pro Kopf in Skandinavien allgemein gestiegen ist.

16. Der Anteil der ausländischen Studierenden an der TU Darmstadt im Fachbereich Architektur hat sich zwischen 1990 und 1993 verändert

Argumente für gefallen:

1. Der Anteil hat sich verringert, weil die rigidere Asylpolitik der Bundesrepublik Deutschland auch potentielle ausländische Studierende betrifft.
2. Der Anteil hat sich verringert, weil die Berichterstattung in den Medien über Ausschreitungen gegen Ausländer\Innen, Deutschland als Studienort für ausländische Studierende unattraktiver macht.
3. Der Anteil hat sich verringert, weil die im Ausland ebenfalls sehr bekannte TU Aachen in einer Umfrage unter Studierenden bessere Beurteilungen bekommen hat. Was die Attraktivität Darmstadts gegenüber Aachen mindern könnte.
4. Der Anteil hat sich verringert, weil die Einsparungen in der Bildungspolitik sich auch auf die Anzahl der Stipendien auswirken.

Argumente für gestiegen:

1. Der Anteil hat sich erhöht, weil die Bemühungen der deutschen Wirtschaft um mehr internationale Zusammenarbeit in Europa sich auch in der steigenden Zahl von Austauschprojekten äußern.
2. Der Anteil hat sich erhöht, weil die Wiedervereinigung und die Bemühungen um das Zusammenwachsen Europas es für (besonders Ost-) europäische Studierende interessanter und wichtiger macht, Deutsch zu lernen.
3. Der Anteil hat sich erhöht, weil seit der Wiedervereinigung die Baubranche einen starken wirtschaftlichen Aufschwung verzeichnet. Die Arbeitsmarktsituation für Architekt\Innen hat sich verbessert.
4. Der Anteil hat sich erhöht, weil nach einer EG-Reform seit einigen Jahren Bürger\Innen aus allen EG-Staaten zu den gleichen Bedingungen wie Deutsche zum Hochschulstudium zugelassen werden können.

17. Die Zahl der sogenannten Kaffeefahrten in Deutschland hat sich zwischen 1985 und 1993 verändert.

Argumente für gefallen:

1. Die Anzahl hat sich verringert, weil die Angst vor Unfällen durch übermüdete Busfahrer\Innen vor Busreisen abschreckt.
2. Die Anzahl hat sich verringert, weil das Selbstverständnis älterer Menschen sich gewandelt hat. Sie bevorzugen eher individuelle Reisen mit dem eigenen Auto.
3. Die Anzahl hat sich verringert, weil das Ansehen solcher Veranstaltungen durch Diskussionen über unsaubere Geschäftsmethoden stark geschwächt worden ist.
4. Die Anzahl hat sich verringert, weil der durch Verkaufsveranstaltungen eng gesteckte Zeitplan es den Reisenden kaum ermöglicht, ihr Ausflugsziel zu genießen.

Argumente für gestiegen:

1. Die Anzahl hat sich erhöht, weil die Zahl der Rentner\Innen gestiegen ist. Diese bilden die Hauptzielgruppe für Kaffeefahrten.
2. Die Anzahl hat sich erhöht, weil die breitere Palette an Ausflugsorten Kaffeefahrten auch für andere Zielgruppen attraktiv macht.
3. Die Anzahl hat sich erhöht, weil das Angebot der veranstaltenden Unternehmen an Kaffeefahrten zugenommen hat.
4. Die Anzahl hat sich erhöht, weil der steigende Konkurrenzdruck viele Busunternehmen zwingt, ihre finanziellen Engpässe durch sogenannte Kaffeefahrten auszugleichen.

18. Die Bereitschaft zur Organspende hat sich in der deutschen Bevölkerung in den letzten 10 Jahren verändert.

Argumente für gefallen:

1. Die Bereitschaft hat sich verringert, weil die ethische Diskussion im Zusammenhang mit Organspenden, „Mensch als Ersatzteillager“ mit zunehmenden medizinischen Möglichkeiten immer brisanter wird.
2. Die Bereitschaft hat sich verringert, weil religiöse Gruppierungen, die Eingriffe in ihren Körper grundsätzlich ablehnen, mehr Zulauf bekommen.
3. Die Bereitschaft hat sich verringert, weil der illegale Handel mit gespendeten Organen immer häufiger in den Medien thematisiert wird und potentielle Spender abschreckt.
4. Die Bereitschaft hat sich verringert, weil die Entnahme von Organen ohne Zustimmung der Angehörigen in letzter Zeit häufig zu Skandalen führte.

Argumente für gestiegen:

1. Die Bereitschaft hat sich erhöht, weil die verbesserten Techniken und der hohe Wissensstand dazu geführt haben, daß das Vertrauen in diese medizinische Möglichkeit gestiegen ist.
2. Die Bereitschaft hat sich erhöht, weil die verbesserte medizinische Aufklärung zu einem erhöhtem Verständnis in der Bevölkerung geführt hat.
3. Die Bereitschaft hat sich erhöht, weil immer mehr Menschen über Angehörige selbst von der Problematik betroffen sind, und laut einer Umfrage steigt deren Bereitschaft, selbst zu spenden.
4. Die Bereitschaft hat sich erhöht, weil die ständig steigende Zahl bereits erfolgreich durchgeführter Transplantationen die Bereitschaft in der Bevölkerung verstärkt, selbst Organspender\In sein zu wollen.

19. Im Jahre 1983 wurden in der EG 19% mehr Milchprodukte hergestellt als verbraucht. Wie hat sich dieser Milchüberschuß bis heute entwickelt ?

Argumente für gefallen:

1. Die EG Landwirtschaftsministerien haben Maßnahmen zum generellen Abbau von Subventionen für Produkte aus dem Agrarwesen ergriffen.
2. Die Diskussion über die Sinnlosigkeit von landwirtschaftlichen Überschüssen und deren Vernichtung zeigt ihre Wirkung.
3. Viele Milchüberschüsse konnten durch neue Milchprodukte verarbeitet werden.
4. Die mangelnde Rentabilität von kleinen und mittleren bäuerlichen Betrieben hat zur Aufgabe vieler Betriebe geführt.

Argumente für gestiegen:

1. Der Absatz von überschüssigen Milchprodukten bei Hilfsmaßnahmen hat abgenommen und könnte damit den Überschuß wiederum erhöhen.
2. Die Subventionen für Molkereiprodukte sind immer noch hoch. (spricht für gleichbleibend!)
3. Der zunehmende Import von billigen Milchprodukten aus osteuropäischen Ländern hat den Verbrauch von EG-Milcherzeugnissen gesenkt, was zu einem höheren Überschuß führen könnte.
4. Viele landwirtschaftliche Betriebe haben von Ackerbau auf Milchwirtschaft umgestellt.

20. Im Jahre 1985 sind 17,8 % der Leistungssportler/innen von gesundheitlichen Folgeschäden betroffen gewesen. Wie hat sich dieser Prozentsatz verändert?

Argumente für gefallen:

1. Die Zahl der Dopingfälle hat zugenommen.
2. Die höheren Anforderungen im Sport durch den technischen Fortschritt bewirken einen schnelleren Verschleiß.
3. Das immer strenger werdende Konkurrenzsystem hat eine ungesunde Lebensweise, physisch wie psychisch, zur Folge.
4. Durch den medizinischen Fortschritt ist eine bessere medizinische Versorgung möglich.

Argumente für gestiegen:

1. Durch eine gegenüber 1985 gestiegene Anzahl von Trainingseinheiten entsteht eine Einseitigkeit in der Belastung des Körpers.
2. Verschärfte Dopingkontrollen halten den Mißbrauch von Medikamenten in Schach.
3. Es existieren erhöhte Sicherheitsmaßnahmen, z.B. durch verbesserte Ausrüstung oder Sportgeräte.
4. Die sich verbreitende Diskussion über Folgeschäden schärft das Bewußtsein der Sportler/innen.

21. Im Jahr 1985 betrug die Anzahl der Interessent/en/innen an asiatischen Kampfsportarten in der deutschen Bevölkerung 3,3%. Wieviel Prozent der Bevölkerung interessiert sich heutzutage für asiatischen Kampfsport?

Argumente für gefallen:

1. Der Anspruch an eine betriebene Sportart als ganzheitliche Auslastung des Menschen steigt.
2. Das Interesse an asiatischer Lebenskultur hat zugenommen.
3. Das Ausüben von asiatischen Kampfsportarten hat einen meditativen Nebeneffekt.
4. Immer mehr deutsche Bürger/innen wollen über die Verteidigungstechnik ihr Selbstvertrauen stärken.

Argumente für gestiegen:

1. Es existiert ein großes alternatives Angebot an Fitnesscentern. (kritisch!)
2. Asiatische Kampfsportarten sind sehr zeitaufwendig und trainingsintensiv.
3. Asiatische Kampfsportcenter haben teilweise inzwischen den schlechten Beigeschmack von Konsum- und Geschäftemacherei.
4. Kampfsport allg. hat oft den Ruf, von der rechtsradikalen Szene gelenkt zu sein (die "Schwarzen Sheriffs" in München).

22. Im Jahre 1970 benutzten 11% der bundesdeutschen Autofahrer/innen nie die Autobahn. Wieviel Prozent der Autofahrer vermeiden heutzutage die Autobahn?

Argumente für gefallen:

1. Es werden immer mehr Zweitwagen zugelassen. Der Zweitwagen wird meistens nicht zum Autobahnfahren genutzt.
2. Viele Autofahrer/innen nutzen ihren Wagen nur noch für Strecken, wo keine anderen Alternativen möglich sind.
3. Die Berichte über Massenkarambolagen auf Autobahnen schrecken viele Autofahrer/innen ab.
4. Das hohe Verkehrsaufkommen führt zu einer stärkeren Nutzung der Bahn und des Flugzeugs für weite Strecken.

Argumente für gestiegen:

1. Die hohen Kosten für Umwege über Landstraßen fallen bei den gestiegenen Benzinkosten mehr ins Gewicht.
2. Das Autobahnstreckennetz ist stark ausgebaut worden.
3. Das Verkehrsaufkommen insgesamt ist gestiegen.
4. Die schwierige Arbeitsmarktlage erfordert eine höhere Mobilität, die zu einem Anstieg des Pendlertums führt.

23. Wie hat sich die Tendenz der Unternehmen, Tierversuche durchzuführen, in den letzten 2 Jahren geändert?

Argumente für gefallen:

1. Es kommen immer mehr neue Produkte auf den Markt, die getestet werden müssen.
2. Die inzwischen anerkannten Alternativen zu Tierversuchen sind immer noch sehr teuer.
3. Es gibt keinen äußeren Druck (Gesetze), der die Unternehmen dazu zwingen könnte, Alternativtests zu Tierversuchen anzuwenden.
4. Die konservative Medizin beharrt immer noch auf die Effektivität der Tierversuche.

Argumente für gestiegen:

1. Immer mehr Firmen werben für ihre Produkte damit, daß sie ohne Tierversuche getestet worden seien.
2. Es gibt inzwischen Einzeller, die man als effektive Versuchsobjekte verwenden kann.
3. Tierschützer/innen erhalten immer mehr Beachtung in der Öffentlichkeit durch medienwirksame Aktionen.
4. Die Übertragbarkeit der Ergebnisse von Tierversuchen auf den Menschen wird zunehmend in Frage gestellt.

24. Eine Langzeitstudie von 1985 bis heute hat ergeben, daß sich die Einstellung junger Menschen in Frankreich in Bezug auf den Aufbau einer eigenen Familie geändert hat.

Argumente für gefallen:

1. Eine Studie im Jahr 1993 erwies, daß die Tendenz, sich jeweils nur für einen Lebensabschnitt an einen Partner/eine Partnerin zu binden, steigt.
2. Die jährlich steigenden Lebenshaltungskosten führen bei jungen Paaren oft zu dem Konflikt, ob eine Familie gegründet werden kann.
3. Aufgrund der zunehmenden Umweltzerstörung waren 1990 nur noch 60% der jung verheirateten Paare bereit, Kinder auf die Welt zu bringen.
4. Der allgemeine Stellenwert des beruflichen Erfolgs ist gestiegen, und immer weniger Menschen möchten heiraten.

Argumente für gestiegen:

1. Sozialwissenschaftler haben in Umfragen gezeigt, daß auch heute noch die Familie als unentbehrliches Auffangnetz gewertet wird.
2. Die Politik der Regierungspartei fördert das Familienbewußtsein.
3. Viele Menschen fürchten sich vor der Vereinsamung als Single.
4. In einer Studie von 1993 wollten wieder 70% der jungen Menschen in Frankreich heiraten.

B.1 Instruktion der Experimente am Computer (Experiment 1-3 und 4)

Instruktion 1

Liebe Versuchsteilnehmerin, lieber Versuchsteilnehmer,

Du wirst im oberen Bereich des Bildschirms nacheinander 24 Sachverhalte sehen. Deine Aufgabe wird es sein, zu jedem **Sachverhalt** Deine subjektive Einschätzung abzugeben, ob dieser im Laufe der Zeit **gestiegen** oder **gefallen** ist. Um Dir eine Grundlage für Deine Entscheidung zu liefern, werden hintereinander jeweils acht **Argumente** erscheinen. Diese Argumente sprechen entweder dafür, daß der entsprechende Sachverhalt gefallen, oder daß er gestiegen ist. Du sollst zunächst jedes Argument selbst bewerten, ob es dafür spricht, daß der Sachverhalt steigt oder fällt. Anschließend erscheint eine Skala auf der Du Deine subjektive Sicherheit eintragen kannst, ob der entsprechende Sachverhalt gestiegen, oder gefallen ist. Aus Platzgründen erscheinen auf dem Bildschirm nur **Stichworte** zu den Argumenten. Du kannst aber jederzeit durch einen Mausklick auf das entsprechende Stichwort das zugehörige Argument nochmals sehen. Solange Du die Maustaste gedrückt hältst bleibt das Argument zu sehen. Zunächst folgt ein **Übungsbeispiel!**

Merke Dir bitte die Stichworte und Argumente zu den jeweiligen Sachverhalten und auch Deine Einschätzung, da diese zum zweiten Termin in einer Woche abgefragt werden!

Instruktion 2

Liebe Versuchsteilnehmerin, lieber Versuchsteilnehmer,

Du wirst im oberen Bereich des Bildschirms wieder, wie im ersten Teil des Versuches, nacheinander 24 Sachverhalte sehen.

In dem letztem Teil des Experimentes geht es darum, daß Du Dich möglichst genau an **Deine ursprünglichen Schätzungen** erinnerst. Du sollst also nochmals Deine Einschätzung von **letzter Woche** wiedergeben, ob der betreffende Sachverhalt gestiegen oder gefallen ist.

Zu einigen Fragen werden Dir die Lösungen präsentiert werden. Laß Dich dadurch nicht beeinflussen und versuche möglichst genau, **Deine ursprüngliche Schätzung** zu erinnern.

Dir wird zunächst wieder ein kurzes **Beispiel** gezeigt.

Bitte frage die Versuchsleiterin bzw. den Versuchsleiter, falls Dir irgendetwas unklar sein sollte!

Good luck!

Instruktion 3

Liebe Teilnehmerin, lieber Teilnehmer,

Dir werden jetzt eine Reihe von Stichwörtern dargeboten werden, und Du sollst möglichst **schnell** entscheiden, ob es sich um ein „**altes**“ Stichwort handelt, daß schon Mal zu diesem Sachverhalt letzte Woche dargeboten wurde, oder, ob es sich um ein „**neues**“ Stichwort handelt, was letzte Woche noch nicht im Zusammenhang mit diesem betreffenden Sachverhalt genannt wurde.

Drücke bitte die

⇐ linke, **rote Taste** , wenn Du das Stichwort für „**alt**“ hältst.

Wenn Du das

Stichwort für „**neu**“ hältst, drücke bitte die rechte, **gelbe Taste** ! ⇒

Antworte bitte so **schnell** Du kannst und so korrekt wie möglich!

Du wirst im folgendem, nach jeder Erinnerung an Deine Entscheidung, einen Durchgang dargeboten bekommen, wo Du Stichwörter nach neu und alt einordnen sollst.

B.2 Instruktionen zum Fragebogen (Experiment 4)

Instruktion 1

Vpn-Nr.:

Code:

Liebe(r) TeilnehmerIn !

Zunächst vielen Dank dafür, daß Du Dich bereit erklärt hast, an unserem Experiment teilzunehmen.

Wir möchten Dich noch einmal darauf hinweisen, daß Du Dich in einer Woche zu einem weiteren Termin bei uns melden mußt, damit der Versuch auch seinen Zweck erfüllt !!!

Bitte trage Dich sich hierzu mit Deinem Code auf unserem Terminkalender für **heute in einer Woche** ein, oder sprich uns direkt an.

Dein Code setzt sich wie folgt zusammen:

1. + 2. Stelle: die letzten beiden Ziffern des Geburtsjahres der Mutter

3. Stelle: 1. Buchstabe des Vornamens der Mutter

4. Stelle: 1. Buchstabe des Vornamens des Vaters

z. B. 1940- Susanne- Herbert : 40-SH

Der Code dient dazu, die Anonymität der Daten zu gewährleisten.

Alter:

Geschlecht:

Nun beginnt die eigentliche Aufgabe:

Bitte lese Dir die folgenden Aussagen mit den dazugehörigen Argumenten durch und triff Deine Entscheidung, ob etwas gefallen oder gestiegen ist mit Hilfe von Wahrscheinlichkeitsaussagen.

Die Argumente spielen auch im letzten Teil des Experimentes eine Rolle. Lies Dir daher sorgfältig alle Argumente durch, merken Dir diese möglichst und triff Dein Urteil, wie sicher Du Dir bist, ob ein Sachverhalt gestiegen oder gefallen ist.

Hier ein Beispiel:

Die Tendenz zu Telefonieren hat sich innerhalb der letzten vier Jahre verändert.

- Es gibt insgesamt **mehr Haushalte** mit Telefonen.
- Der **Preis** pro Telefoneinheit ist **gestiegen**.
- Die **Erreichbarkeit** rund um die Uhr durch das Telefon ist heutzutage unabdingbar.
- Es gibt viele neue Möglichkeiten der direkten Verständigung durch die **Computer-Mailbox** und auch durch Faxgeräte
- Der **technische Fortschritt** macht Telefonieren attraktiver.
- Die allgemeinen **Lebenshaltungskosten** sind gestiegen
- Telefonieren ist **selbstverständlich** geworden
- Eine Studie hat ergeben, daß der **Anrufbeantworter** viele Personen vom Telefonieren **abschreckt**.

Wahrscheinlichkeit **für gestiegen** (in Prozent)

%

Das bedeutet Du bist nicht ganz sicher, daß in letzter Zeit mehr telefoniert wird als vor vier Jahren. Bist Du hingegen ganz sicher, daß das Telefonieren zugenommen hat, solltest Du auch 100% angeben. Mit anderen Worten, je sicherer Du bist desto näher gehst Du an die 100% heran, und je unsicherer Du bist desto näher gehst Du an die 50% ran!

Bitte keine Frage auslassen!

Instruktion 2

Vpn-Nr. Code:

Vielen Dank, daß Du wieder gekommen bist.

Bitte gib wieder Deinen Code an.

Zur Erinnerung:

- 1.+2. Stelle: die letzten beiden Ziffern des Geburtsjahres der Mutter
3. Stelle: Anfangsbuchstabe des Vornamens der Mutter
4. Stelle: Anfangsbuchstabe des Vornamens des Vaters

Zu einigen der Fragen, zu denen Du beim letzten Termin ein Urteil abgegeben hast, bekommst Du jetzt eine Rückmeldung darüber, ob Du richtig oder falsch geschätzt hast.

Deine Aufgabe ist es aber, Dich **möglichst genau** an Deine Einschätzung der **letzten Woche** in Prozentwerten zu erinnern. Schreib diese in die dafür vorgesehene Zeile.

Gib also bitte keine neue Einschätzung ab, sondern erinnern Dich genau an Deine ursprüngliche Einschätzung!

Bitte keine Frage auslassen!

Instruktion 3

Liebe(r) TeilnehmerIn !

geh bitte den Fragebogen noch einmal von vorne durch. Wie Du Dich sicherlich erinnern kannst, waren zu jeder Frage im ersten Durchgang acht Argumente aufgeführt. Versuch bitte, zu jeder Frage so viele Argumente wie möglich zu erinnern und sie stichwortartig aufzuschreiben. Schön wäre es, wenn Du die Stichworte auch noch danach sortieren könntest, ob sie dafür sprechen, daß der Sachverhalt in der Frage „gestiegen“ oder „gefallen“ ist. Tragen dazu bitte die Argumente in die folgenden Spalten ein.

Bitte keine Frage auslassen!

C Erklärung

Hiermit erkläre ich, Markus Eisenhauer, daß ich die vorliegende Dissertation selbständig verfaßt und keine anderen als die angegebenen Hilfsmittel benutzt und die aus fremden Quellen direkt oder indirekt übernommenen Gedanken als solche kenntlich gemacht habe.

Die Dissertation habe ich bisher keinem anderen Promotionsausschuß in gleicher oder vergleichbarer Form vorgelegt. Die Dissertation wurde bisher nicht veröffentlicht.

Trier, den 30. Juni 2000.