

**Quellenidentifikationsleistung in
variierenden Befragungskontexten**

-

**Instruktionen zur Verbesserung
kindlicher Erinnerungsleistung bei Zeugenaussagen**



Inauguraldissertation
zur
Erlangung des Doktorgrades
der Humanwissenschaftlichen Fakultät
der Universität zu Köln
vorgelegt von

Liesel Anna Maria Heiermann
aus Wesel

August 2018

Erstgutachterin:

Prof. Dr. Ellen Aschermann

Zweitgutachterin:

Prof. Dr. Birgit Träuble

Tag der mündlichen Prüfung: 13.12.2018

Die Rechtspsychologie beschäftigt sich mit der Anwendung psychologischer Methoden und psychologischen Wissens zur Beantwortung von Fragen des Rechtssystems (Wrightsmann, 2001). Insbesondere die Aussagepsychologie nimmt im Bereich der forensischen Psychologie einen bedeutenden Stellenwert ein. Zu Anfang des 20. Jahrhunderts wurde von William Marston der Vorläufer des Lügendetektors entwickelt, im Jahr 1954 hat Undeutsch mit der *Undeutsch-Hypothese* den Grundstein zur systematischen Erforschung von Zeugenaussagen gelegt. Er differenzierte erstmalig zwischen Glaubhaftigkeit als Aussage- und Glaubwürdigkeit als Personenmerkmal und postulierte, dass sich erlebnisbasierte Aussagen in ihrer inhaltlichen Qualität von erfunden systematisch unterscheiden (Steller, 1989). Zeugenaussagen sind im Rahmen von polizeilichen Ermittlungen und gerichtlichen Verfahren häufig das einzige zur Verfügung stehende Beweismittel (Fisher, 2010); die Aussagen unterliegen jedoch zahlreichen Einflüssen, welche auf personaler, situativer und kontextueller Ebene verortet werden können. Im polizeilichen, gerichtlichen sowie psychologischen Umfeld wurden zahlreiche Befragungsformate und Methoden entwickelt sowie evaluiert, mit dem Ziel, möglichst wahre und akkurate Angaben zu erhalten (Howitt, 2011). Hier setzt die Untersuchung der Arbeit an.

In der vorliegenden experimentellen Feldstudie mit $N = 105$ Kindern¹ ($M = 9,81$ Jahre) wurde überprüft, wie sich variierende Instruktionen, welche die Zuverlässigkeitsmotivation manipulieren, auf die Aussagegenauigkeit und Quellenidentifikationsleistung auswirken. In Weiterführung von Roebers, Moga und Schneider (2001), welche die positive Wirkung von Belohnung auf die Aussagegenauigkeit nachweisen konnten, wurde in der vorliegenden Untersuchung die Relevanz von Strafe in Kombination mit einer freien und einer forcierten Antwortbedingung untersucht. Dabei wurde verbal und bildlich präsentiertes Material in einer Originalsituation mit einer siebentägigen Latenz abgefragt. Es wurden vier Befragungsbedingungen realisiert, welche sich durch hohe oder niedrige Schwellen der Zuverlässigkeitsmotivation unterscheiden. Dies wurde über die Möglichkeit einer „Ich weiß nicht“ (*IWN*)-Antwortoption realisiert (Faktor **FORCIERUNG**). Es wird davon ausgegangen, dass damit Kooperationseffekten entgegen gewirkt und das Gedächtnismonitoring erleichtert wird (Hughes & Grieve, 1980; Koriat

¹ Zugunsten einer besseren Lesbarkeit wird auf die Unterscheidung der männlichen und weiblichen Bezeichnungen verzichtet und das generische Maskulin verwendet. Im Folgenden sind immer beide Geschlechter eingeschlossen.

& Goldsmith, 1996; Roebers et al., 2001). Des Weiteren wurde eine Bestrafung falscher Antworten im Sinne der operanten Konditionierung vorgenommen (Faktor STRAFE), welche die negativen Konsequenzen von Falschaussagen repräsentieren soll. Als abhängige Maße wurden die numerische Anzahl erinnerter Items im freien Bericht und traditionelle Maße der Quellenidentifikationsleistung erhoben. Es wurde festgestellt, dass durch eine obligatorische Antwort die akkuratesten Aussagen resultieren, die Bestrafung hingegen keinen Einfluss hatte. Die Bereitstellung einer *IWN*-Antwortoption zeigte keine positiven Effekte auf die Aussagequalität hinsichtlich Itemerkennung und Quellenidentifikation. Die Befunde stehen zunächst im Gegensatz zu den Arbeiten von Roebers et al. (2001) und legen die Vermutung nahe, dass Prozesse des Gedächtnismonitorings im Sinne von Koriat und Goldsmith (1996) bei Kindern der untersuchten Altersstufe noch nicht stattfinden. Eine weitere Erklärung zur Entstehung der inkonsistenten Ergebnisse könnte in der unvollständigen Kreuzung der Faktorstufen bei Roebers et al. (2001) zu verorten sein.

Inhaltsverzeichnis

A THEORIE UND GRUNDLAGEN	8
Einleitung	9
1 Gedächtnisprozesse und Modellvorstellungen	10
1.1 Quellengedächtnis.....	12
1.1.1 Reality Monitoring.....	12
1.1.2 Source Monitoring	13
1.2 Gedächtnisentwicklung und Source Monitoring bei Kindern.....	20
2 Kognitive Aspekte der Zeugenaussage	31
2.1 Validität von Augenzeugenaussagen	31
2.2 Augenzeugengedächtnis und autobiografisches Gedächtnis	31
2.2.1 Schemawissen	32
2.2.2 Pseudoerinnerungen und Konfabulationen.....	33
2.3 Kinder als Zeugen	34
2.3.1 Personale Aspekte.....	38
2.3.2 Situative Aspekte	40
3 Aussagepsychologische Begutachtung und Befragungstechniken	46
3.1 Glaubhaftigkeitsbegutachtung	46
3.2 Kriterienorientierte Inhaltsanalyse	47
3.3 Das kognitive Interview.....	52
3.4 Aktuelle Interviewtechniken und -verfahren	56
4 Zuverlässigkeitsmotivation	58
4.1 Strategische Regulation der Aussagegenauigkeit	60
4.2 Ansätze zur Operationalisierung.....	63
4.2.1 „Ich weiß nicht“-Antwortoption	64
4.2.2 Vermeidungslernen.....	69
B EMPIRIE	71
5 Fragestellung	72
6 Methode	73
6.1 Versuchsdesign	73
6.2 Hypothesen	77
6.3 Stichprobe, Material und Untersuchungsablauf.....	78
7 Ergebnisse	85
7.1 Deskriptive Statistik	85

7.1.1 Itemerkennung	85
7.1.2 Quellenidentifikation	87
7.1.3 <i>IWN</i> -Antwortoption.....	88
7.2 Inferenzstatistische Analyse	88
7.2.1 Itemerkennung	89
7.2.1.1 Untersuchungen des Faktors FORCIERUNG	89
7.2.1.2 Untersuchungen des Faktors STRAFE	91
7.2.2 Quellenidentifikation	92
7.2.2.1 Untersuchungen des Faktors FORCIERUNG	92
7.2.2.2 Untersuchungen des Faktors STRAFE	93
7.2.3 Zusammenfassung der inferenzstatistischen Ergebnisse.....	94
7.3 Qualitative Inhaltsanalyse.....	94
7.3.1 Der freie Bericht im Interview.....	95
7.3.2 Deskriptive Ergebnisse der Inhaltsanalyse	98
7.4 Explorative Analyse	101
8 Diskussion	101
8.1 Freier Bericht.....	102
8.2 Itemerkennung und Quellenidentifikation.....	105
8.3 Limitationen	110
C RESÜMEE	115
9 Fazit und Implikationen.....	116
Literatur... ..	119
Anhang.....	138

Tabellenverzeichnis

Tabelle 1. Arten des Reality und Source Monitorings (nach Roberts & Blades, 2000).	15
Tabelle 2. Antwortmatrix in der experimentellen Untersuchung von Quellenidentifikationsleistungen (nach Batchelder & Riefer, 1990).	16
Tabelle 3. Interpretation der Effektstärken (nach Cohen, 1988; Eid et al, 2013).	21
Tabelle 4. Systematik der Forschungsliteratur zu kindlichen Zeugenaussagen.	39
Tabelle 5. CBCA Kriterien (nach Steller & Köhnken, 1989).	50
Tabelle 6. Versuchsdesign.	74
Tabelle 7. Berechnungsformeln der verwendeten Indizes.	77
Tabelle 8. Mittlere Anzahl beantworteter Items in den vier Versuchsbedingungen...	86
Tabelle 9. Mittlere Qualität der Itemrekognitionsleistung in den vier Versuchsbedingungen.	86
Tabelle 10. Mittlere Qualität der Quellenidentifikationsleistung in den vier Versuchsbedingungen.	87
Tabelle 11. Verteilung der <i>IWN</i> -Antworten.	88
Tabelle 12. Übersicht zu den Ergebnissen der ANOVA für die Itemrekognition.	90
Tabelle 13. Übersicht zu den Ergebnissen der ANOVA für die Quellenidentifikation.	92
Tabelle 14. Mittelwerte und Standardabweichungen korrekter Erinnerungsleistungen über alle Items der Kategorien <i>QI</i> -Bogen, Zusatz und Rahmentext über die gesamte Stichprobe und nach Altersklassen im freien Bericht.	99

Abbildungsverzeichnis

Abbildung 1. Multinomiales Modell der Quellenidentifikation mit sieben Parametern (nach Batchelder & Riefer, 1990).	19
Abbildung 2. Das schematische Modell der strategischen Regulationsleistung der Erinnerungsgenauigkeit und –menge (nach Koriat & Goldsmith, 1996).	61
Abbildung 3. Simulierte Gedächtnisleistung hinsichtlich Erinnerungsmenge und - genauigkeit als Funktion des Antwortkriteriums <i>Prc</i> (nach Koriat & Goldsmith, 1996).	62
Abbildung 4. Alters- und Geschlechtsverteilung.	79
Abbildung 5. Auswertungsbogen 1 für die <i>free</i> -Bedingungen.	80
Abbildung 6. Auswertungsbogen 2 für die <i>forced</i> -Bedingungen.	81
Abbildung 7. Beispielillustrationen aus Geschichte A und B.	82
Abbildung 8. Vergleich der mittleren Qualität der Itemrekognitionsleistung (Q_{IR}) in den <i>free</i> - und <i>forced</i> -Bedingungen.	90
Abbildung 9. Vergleich der mittleren Qualität der Quellenidentifikationsleistung (Q_{QI}) und der Qualität der Itemrekognitionsleistung (Q_{IR}) in den <i>free</i> - und <i>forced</i> - Bedingungen.	93
Abbildung 10. Das Kategoriensystem der qualitativen Inhaltsanalyse zur Untersuchung der kindlichen Erinnerungsleistungen im freien Bericht.	96

A THEORIE UND GRUNDLAGEN

Einleitung

In der vorliegenden empirischen Untersuchung wird der Effekt einer Instruktion zur Zuverlässigkeitsmotivation auf die Erinnerungsmenge und Genauigkeit von kindlichen Augenzeugenaussagen untersucht. Als abhängige Variable wird die Quellenidentifikationsleistung von vorab präsentierten Material erhoben. Nach Koriat und Goldsmith (1994, 1996) wird von einem Qualitäts-Quantitäts Trade-Off ausgegangen, wenn das Antwortkriterium dahingehend beeinflusst wird, dass eine intensivere Überwachung geäußerter Erinnerungen angestoßen wird. Dies wird als Monitoringprozess bezeichnet, welcher in diversen Untersuchungen von Roebbers und Kollegen (2001; Roebbers & Fernandez, 2002; Roebbers & Schneider, 2005a) durch die Bereitstellung einer "Ich weiß nicht" (*IWN*)-Antwortoption realisiert wurde. Damit resultieren im Sinne des Trade-offs weniger aber akkuratere Antworten. Erwachsene Personen profitieren mehr als Kinder von der *IWN*-Option, was auf ausgeprägtere Monitoringkompetenzen zurückgeführt wird.

Im Folgenden werden zunächst gedächtnispsychologische Grundannahmen präsentiert, insbesondere Vorstellungen über das Quellengedächtnis und die Konzepte des Reality Monitoring und das Source Monitoring Framework (Johnson, Hashtroudi & Lindsay, 1993; Johnson & Raye, 1981). Im Zuge dessen werden Besonderheiten der Gedächtnisentwicklung bei Kindern und empirische Studien zum Source Monitoring vorgestellt. Im zweiten Kapitel werden kognitive Aspekte der Zeugenaussage besprochen, darauf folgen Besonderheiten von kindlichen Zeugen. Im dritten Kapitel wird die Methode der aussagepsychologischen Glaubhaftigkeitsbegutachtung vorgestellt, gleichzeitig wird die Methodik des kognitiven Interviews beschrieben, dessen freier Bericht einen Teil der methodischen Untersuchung der hier vorliegenden Arbeit ausmacht. Das vierte Theoriekapitel stellt das zentrale Thema der Untersuchung in den Vordergrund, strategische kognitive Prozesse, welche es ermöglichen Erinnerungen zu überwachen. Des Weiteren erfolgt hier die Besprechung empirischer Studien, welche die Zuverlässigkeitsmotivation mit Hilfe der Manipulation von Antwortoptionen beeinflussen. Nach der theoretischen Basis werden Methode, Ergebnisse und Diskussion der vorliegenden Studie vorgestellt.

1 Gedächtnisprozesse und Modellvorstellungen

Menschliche Erinnerungen sind rekonstruktiv; bei jedem Abruf können sie sich in Abhängigkeit von Stimmung, Umgebung und situativen Bedingungen verändern oder es werden fehlende Informationen durch logisch passende Elemente ergänzt bzw. Details weggelassen. Erinnerungen sind immer subjektiv und durch individuelle Interpretationen eines Sachverhaltes gekennzeichnet (Balloff, 2004).

Im Rahmen der psychologischen Grundlagenforschung zu kognitiven Prozessen konnten zunehmend Erkenntnisse über die Elemente und Vorgänge des Wissenserwerbs gewonnen werden. Das Gedächtnis ist ein kognitives System, welches die Enkodierung, Speicherung und den Abruf von Informationen ermöglicht (Greuel, 2001). Die Entwicklung gedächtnispsychologischer Modelle erfolgte auf Grundlage der Informationsverarbeitungstheorie, wobei Gedächtnisprozesse als Teilprozesse der Informationsverarbeitung verstanden werden, welche eng mit der Wahrnehmung verknüpft sind (Büttner, 2003). Zu den zentralen Prozessen zählen die Wahrnehmung bzw. Enkodierung, die Speicherung und der Abruf bzw. die Wiedererkennung (Tulving, 1972). Enkodierung beschreibt den Prozess der Verarbeitung und Übersetzung von Umweltreizen in mentale Repräsentationen. Speicherung wird im Sinne der Gedächtnispsychologie als Ablage von enkodierten Informationen in verschiedene Gedächtnissysteme verstanden, welche dann im Prozess des Abrufs wieder aufgefunden werden. Hier wird zwischen der Wiedererkennung bzw. Rekognition und der Reproduktion (Büttner, 2003) unterschieden. Die Reproduktionsleistung wird als komplexer und schwieriger als die reine Wiedererkennung angesehen und beinhaltet neben bloßem Wiedererkennen auch die aktive Wiedergabe von Informationen. Der Abruf kann durch Erinnerungshilfen oder Hinweisreize erleichtert werden (Ornstein, Naus & Liberty, 1975; Roebbers & Elischberger, 2002). So wird die Suche nach entsprechenden Informationen dann erleichtert, wenn während der Suche Kontextinformationen wie z.B. Ort, Zeit und Stimmung, mit der gesuchten Information assoziiert und aktiviert werden. Dieser Effekt der Erleichterung durch Hinweisreize wird durch das Modell der Kodierungsspezifität beschrieben (Tulving & Osler, 1968). Die Quelle von Informationen, die in der vorliegenden Arbeit im Fokus steht, stellt einen Teil der Kontextinformation dar.

In den Modellvorstellungen von Atkinson und Shiffrin (1968), Baddeley und Hitch (1974) und Anderson (1976) werden verschiedene Grundannahmen präsentiert, unter anderem, dass Informationen in temporal distinkten und seriell aufeinanderfolgenden mentalen Schritten verarbeitet werden. Dabei werden limitierte kognitive Ressourcen benötigt, deren Verfügbarkeit in Abhängigkeit vom Grad der Automatisierung variiert; Übung kann beispielsweise die Automatisierung von Prozessen bedingen. Aufgenommene Informationen können als Wissen gespeichert werden, wobei in Gedächtnismodellen die Art des Speichersystems und der Information unterschieden werden (Büttner, 2003). Das bekannteste Modell ist das Mehrspeichermodell von Atkinson und Shiffrin (1968), welches drei separate Speichereinheiten postuliert: das sensorische Register oder Ultrakurzzeitgedächtnis, das Arbeits- oder Kurzzeitgedächtnis und das Langzeitgedächtnis. Das Langzeitgedächtnis ist ein Speicher, der theoretisch in seiner Kapazität und Speicherdauer unbegrenzt ist (Büttner, 2003). Als relevant identifizierte Informationen werden vom Kurzzeitgedächtnis in das Langzeitgedächtnis übertragen, dies kann durch Prozesse der Elaboration und auch Konsolidierung erfolgen. Elaboration beschreibt die vertiefte Verarbeitung und Verknüpfung von Informationen zu bestehendem Wissen (Friedrich & Mandl, 1992), Konsolidierung ist die sog. Festigung von Wissensinhalten, welche Prozesse der Aufnahme, Abspeicherung und Ablage von Informationen beinhaltet (Markowitsch & Staniloiu, 2011).

Die Informationen bzw. Subsysteme des Langzeitgedächtnisses werden klassischerweise in deklarativ und prozedural unterteilt, wobei das deklarative Gedächtnis das episodische und semantische Gedächtnis umfasst (Anderson, 1976; Tulving, 1972; Tulving & Thomson, 1973). Das prozedurale Gedächtnis enthält Wissen über die Ausführung und den Ablauf motorischer Handlungen (Büttner, 2003) und kognitiver Prozesse. Semantisches Wissen bezieht sich auf Fakten und Wissen ohne kontextuelle, das heißt zeitliche oder räumliche, Bezüge (Tulving, 1972). Episodisches Wissen hingegen ist autobiografisch und bezieht sich auf persönliche Erlebnisse und Erfahrungen mit raum-zeitlicher Einbettung. Das episodische Langzeitgedächtnis wird Nyberg (1999) zufolge in das Item- und das Quellengedächtnis unterteilt. Dabei ist das Itemgedächtnis, welches die Inhalte eines Ereignisses speichert und für den Prozess der Alt-Neu-Rekognition verantwortlich und das Quellengedächtnis, welches die Kontextelemente einer Wahrnehmungssituation enthält, für die Erinnerung des Ursprungs

einer Information. Der Multi-Komponenten Theorie von Baddeley und Hitch (1974) zufolge besteht das Gedächtnis aus einem Netzwerk, welches durch assoziative Verknüpfungen gekennzeichnet ist (Collins & Loftus, 1975).

1.1 Quellengedächtnis

Im Rahmen der Mehrspeichermodelle (z.B. Atkinson & Shiffrin, 1968; Büttner, 2003) wird dem Langzeitgedächtnis die Speicherung von Quelleninformationen zugeschrieben. Die Erinnerung an die Quelle einer Information wird im noetischen Teil des episodischen Gedächtnisses (autonoetische Informationen sind rein autobiografische Fakten, wie z.B. der Name oder der Geburtsort), dem Quellengedächtnis, verortet (Greuel, 2001). *Quellenidentifikation* beschreibt den Prozess der Zuordnung einer Erinnerung oder Information zu ihrer Quelle (Johnson, Hashtroudi & Lindsay, 1993) und dient als Überbegriff für Reality Monitoring und Source Monitoring (siehe Kap. 1.1.1 & 1.1.2). Die Quelleninformation zählt zu den Kontextinformationen, wie z.B. Ort, Zeit oder Stimmung, und kann unter Umständen den Abruf- bzw. Erinnerungsprozess erleichtern (Bröder & Meiser, 2007; Meiser & Bröder, 2002). Diesen perzeptuellen Eigenschaften wird im Rahmen des Reality Monitorings (Johnson & Raye, 1981), welches im Folgenden vorgestellt wird, besondere Bedeutung beigemessen.

1.1.1 Reality Monitoring

Das Reality Monitoring Paradigma von Johnson und Raye (1981) basiert auf der Grundannahme, dass sowohl externale Ereignisse und Wahrnehmungsinhalte als auch internale Gedanken und Vorstellungen persistente Erinnerungspfade im Gedächtnis produzieren. Dabei beinhalten Erinnerungen an externale Ereignisse Informationen der Rezeptionssituation, im Gegensatz zu nur vorgestellten Ereignissen. Der Prozess des Reality Monitorings bezeichnet die Entscheidung, ob eine episodische Information einen internalen Ursprung hat und selbst generiert ist oder einem realen Wahrnehmungsprozess entstammt und einen externalen Ursprung hat. „Real“ wird in diesem Kontext als außerhalb der Person existierend verstanden. Die Autoren konstatieren, dass die meisten Gedächtnisfehler, wie etwa das Phänomen der *False Recognition* (siehe Kapitel 2.2.2 „Pseudoerinnerungen und Konfabulationen“), auf Defiziten bei der Diskrimination der Quelle eines Erinnerungspfades, also der Unterscheidung zwischen Vorstellung und tatsächlich Erlebtem, basieren (Johnson & Raye,

1981). Johnson und Raye (1981) beschreiben in ihrem Modell zwei zentrale Einflussfaktoren auf das Reality Monitoring. Dabei spielt die Natur der Gedächtnispfade eine Rolle, wobei Art und Charakteristika einer erinnerten Information die Entscheidung bestimmen. Dazu zählen zum Beispiel sensorische Eigenschaften, operationale Attribute, welche den involvierten kognitiven Prozess beschreiben, oder der semantische Inhalt sowie Kontextinformationen. Internal und external generierte Informationen lassen sich demzufolge hinsichtlich dimensionaler Eigenschaften abgrenzen: Externale Erinnerungen weisen naturgemäß mehr Kontextinformationen auf, wie räumliche, zeitliche, sensorische und semantische Bezüge und haben zudem semantisch detailliertere Erkennungszeichen. Internal generierte Erinnerungen werden vermehrt mit operationalen Attributen verbunden. Unabhängig vom Inhalt der Erinnerung wird für den Entscheidungsprozess die Ausprägung der o.g. Dimensionen genutzt, welche auch kombiniert und mit unterschiedlicher Gewichtung angewendet werden können. Die Autoren gehen weiterhin davon aus, dass die Prozesse der Zuordnung internal versus external nicht immer gleich ablaufen, sondern auf der Bewertung weiterer Informationen basieren (Johnson & Raye, 1981). So werden qualitative Charakteristika des evaluierten Pfades beispielsweise in Kombination mit Vorwissen, in Abhängigkeit von individuellen Entscheidungsstrategien oder antizipierten Kosten von Fehlurteilen evaluiert. Dabei ist der Inhalt der Erinnerung statt der dimensional Charakteristika von Bedeutung. Der Entscheidungsprozess verläuft hier systematischer und ist damit zeitaufwändiger. Zusammenfassend bezeichnen die Autoren den Prozess des Reality Monitorings als Vorgang, der dem Metagedächtnis zugeordnet werden kann und somit die Selbstreflektion eines Einzelnen über seine eigenen Erinnerungen repräsentiert (Johnson & Raye, 1981).

1.1.2 Source Monitoring

Johnson, Hashtroudi und Lindsay (1993) erweiterten die Theorie des Reality Monitorings (Johnson & Raye, 1981) um den Prozess des Source Monitorings. Gemäß den Autoren müssen drei Arten von Quellenidentifikation unterschieden werden: Reality Monitoring (siehe Kap. 1.1.1), externes und internes Source Monitoring. Der Prozess des externalen Source Monitorings differenziert externe Quellen, wie verschiedene Personen oder verschiedene Medien. Internales Source Monitoring unterscheidet interne Informationen (siehe Tabelle 1). Analog zum Arbeitsmodell von Johnson und Raye (1981) postulieren Johnson et al. (1993) zwei Entscheidungsprozesse der

Quellenidentifikation von. Der erste verläuft unbewusst und automatisch auf Basis qualitativer Charakteristika der aktivierten Erinnerung. Perzeptuelle und semantische Details, Kontextinformationen, affektive Eigenschaften und kognitive Operationen liegen bei internal und external generierten Informationen in verschieden starker Ausprägung vor. Mitunter werden Quellenidentifikationsentscheidungen strategisch, argumentativ und unter Hinzunahme von Vorwissen getroffen, womit diese langsamer als die unbewussten Entscheidungen ablaufen. Die genannten Prozesse bedingen und kontrollieren einander; wird eine Erinnerung fälschlicherweise aufgrund mangelnder Plausibilität als internal generiert (z.B. vorgestellt, geträumt) kategorisiert, so kann sie über den automatischen Entscheidungsweg aufgrund ihrer Lebhaftigkeit einem externalen Ursprung zugeordnet werden. Welche Art von Entscheidungsprozess angewandt wird, ist von persönlichen Zielen, Überzeugungen und allgemeinen Entscheidungskriterien abhängig. So ist die Gewichtung und Beimessung von Bedeutsamkeit verschiedener dimensionaler Ausprägungen, wie perzeptuelle oder affektive Informationen, individuell verschieden. Abhängig vom Kontext kann eine strenge Beurteilung unter Hinzunahme des heuristischen als auch des systematischen Prozesses erfolgen. Zu beachten bleibt, dass die Charakteristika einer Information nicht immer vollständig zugänglich oder unter Umständen nur teilweise verfügbar sind. So kann beispielsweise der Ort der Rezeption ohne Kontextmerkmale erinnert werden, wobei alle Merkmale, die zu einer vollständigen Kontextualisierung der Information beitragen, die Quellenidentifikationsleistung verbessern und die Wahrscheinlichkeit einer korrekten Zuordnung erhöhen (Johnson et al., 1993).

Neben den Quellencharakteristika internal und external postulieren Roberts und Blades (2000) eine weitere Dimension, welche für den Prozess der Quellenidentifikation von Bedeutung ist. Hinsichtlich des Status einer Information bzw. Erinnerung muss zwischen öffentlichen und privaten Informationen unterschieden werden. Somit ergibt sich für das Reality Monitoring und das internale Source Monitoring eine weitere Unterscheidung (siehe Tabelle 1).

Für das externale Source Monitoring resultieren keine weiteren Differenzierungen, da ausschließlich fremd generierte Informationen betrachtet werden, deren Status konstant öffentlich ist (Roberts & Blades, 2000). Im Verlauf der vorliegenden Arbeit ist nur das externale Source Monitoring von Bedeutung.

Tabelle 1. Arten des Reality und Source Monitorings (nach Roberts & Blades, 2000).

Art	Status	Quelle
Reality Monitoring • Staturebene	öffentlich vs. privat	external vs. internal
Reality Monitoring • Quellenebene	öffentlich vs. öffentlich	external vs. internal
Reality Monitoring • Quellenebene	öffentlich vs. privat	external vs. internal
Internales Source Monitoring	öffentlich vs. öffentlich	internal vs. internal
Internales Source Monitoring	öffentlich vs. privat	internal vs. internal
Internales Source Monitoring	privat vs. privat	internal vs. internal
Externales Source Monitoring	öffentlich vs. öffentlich	external vs. external

Aschermann (1998) postuliert bei der Untersuchung imaginaler Erinnerungsstrategien eine zweidimensionale Differenzierung der Quellenidentifikation. Die Prozesse des Source Monitorings werden hier über die Dimensionen Modalität und Akteur definiert. Dabei werden tatsächliche und vorgestellte Handlungen (Modalität) über selbst oder fremd ausgeführte Handlungen (Akteur) unterschieden.

Indikatoren zur Messung des Source Monitorings

In experimentellen Untersuchungen zur Quantifizierung der Itemrekognitions- und Quellenidentifikationsleistung werden Items aus verschiedenen Quellen präsentiert, welche anschließend von Distraktoren differenziert und den jeweiligen Quellen zugeordnet werden sollen. In Anlehnung an die *Signal-Detection-Theory* (SDT) (Green & Swets, 1966) werden vier Arten von Antworten unterschieden: *Hits*, *Errors*, *Missings* und *False Alarms*. In Tabelle 2 sind diese schematisch unter Berücksichtigung der Quelle und der gegebenen Antwort dargestellt. An erster Stelle ist die tatsächliche Quelle, an zweiter Stelle die gegebene Antwort indiziert.

Tabelle 2. Antwortmatrix in der experimentellen Untersuchung von Quellenidentifikationsleistungen (nach Batchelder & Riefer, 1990).

		Antwort		
		„A“	„B“	„N“
Quelle	A	Hit X_{AA}	Error X_{AB}	False Alarm X_{AN}
	B	Error X_{BA}	Hit X_{BB}	False Alarm X_{BN}
	N	Miss X_{NA}	Miss X_{NB}	Hit/Correct rejection X_{NN}

Anmerkung: A=Quelle A; B=Quelle B; N=Neues Distraktoritem.

(Average) Conditional Source Monitoring Identification Measure

Unter Einbezug dieser Antworttypen lassen sich drei Indizes berechnen, welche üblicherweise in Untersuchungen zum Item- und Quellengedächtnis Anwendung finden. Für das Itemgedächtnis handelt es sich um die *Hit Rate (HR)*, *False Alarm Rate (FAR)* und *Identification-of-origin (I)*. Bei diesen Maßen wird eine Relativierung an der jeweils zugrundeliegenden Gesamtzahl an Items vorgenommen. Im Folgenden sind die jeweiligen Formeln zur Berechnung der Indizes aufgeführt, welche sich der Notation aus Tabelle 2 bedienen.

$$HR = \frac{X_{AA} + X_{AB} + X_{BA} + X_{BB}}{X_{AA} + X_{AB} + X_{AN} + X_{BN} + X_{BA} + X_{BB}} \quad (1) \text{ Hit Rate}$$

$$FAR = \frac{X_{NA} + X_{NB}}{X_{NA} + X_{NB} + X_{NN}} \quad (2) \text{ False Alarm Rate}$$

$$I = \frac{X_{AA} + X_{BB}}{(X_{AA} + X_{BB}) + (X_{AB} + X_{BA})} \quad (3) \text{ Identification of Origin}$$

Batchelder und Riefer (1990) merken einige methodische Mängel an der Berechnungsgrundlage der Maße HR , FAR und I an. Zum einen werden Daten von verschiedenen Quellen kombiniert, was dazu führen kann, dass eine unterschiedliche Entdeck- und Unterscheidbarkeit verschiedener Quellen maskiert wird. Des Weiteren repräsentieren HR und FAR in den oben angegebenen Formeln die übliche Messung der Hits und False Alarms in einer Ja-/Nein-Rekognitionsaufgabe, wenn die Quellenvariable ignoriert wird. Anhand von Gruppenvergleichen wurde jedoch deutlich, dass eine getrennte Berechnung für HR und FAR zu einem falschen Bild von Rekognitionsdaten führen kann (Batchelder & Riefer, 1990). Lange Zeit wurden diese Maße kombiniert, um das Sensitivitätsmaß d' aus der Signalentdeckungstheorie für die Quellenidentifikationsleistung zu ermitteln. d' wird aus der Differenz der z-standardisierten Treffer und Falschen Alarmen berechnet.

$$d' = Z_{HR} - Z_{FAR} \quad (4) \text{ Sensitivitätsmaß}$$

Murnane und Bayen (1996) nutzen d' zur Bestimmung der Itemwiedererkennung und begegnen der methodischen Kritik von Batchelder und Riefer (1990) mit dem Argument, dass d' gegenüber multinomialen Modellierungen für bestimmte Datenmuster passender ist. Murnane und Bayen (1996) verwenden weiterhin zur Bestimmung der Quellenidentifikationsleistung den (*Average-Conditional-Source-Identification-Measure* Index (*(A)CSIM*), welcher weitestgehend unabhängig von Quelleneigenschaften, respektive Itemrekognitionsleistungen ist. Dabei wird *CSIM* für den Vergleich einzelner Quellen und *ACSIM* für den Vergleich von zwei Paaren von Quellen verwendet. Für jede Quelle ergibt sich ein separater Identifikationskoeffizient, aus deren Summation anschließend ein Mittel gebildet wird (Bayen, Murnane & Erdfelder, 1996; Murnane & Bayen, 1996).

$$ACSIM = \frac{\frac{X_{AA}}{X_{AA}+X_{AB}} + \frac{X_{BB}}{X_{BB}+X_{BA}}}{2} \quad (5) \text{ Average Conditional Source Identification Measure}$$

Exkurs: Multinomiale Modelle

Multinomiale Modelle stellen eine Variante der Schwellen bzw. Threshold-Modelle dar, welche der Erfassung der Signalentdeckung und Rekognitionsleistung dienen (Schütz, 2011; Velden, 1982). Batchelder und Riefer (1990) empfehlen zur Quantifizierung kognitiver Leistungen die multinomiale Modellierung kognitiver Prozesse und damit einhergehend die Schätzung von entsprechenden Parametern. Dabei wird die Wahrscheinlichkeit der Identifikation für jedes Item respektive jede Quelle separat betrachtet. Im Gegensatz zu den Modellvorstellungen der Signalentdeckungstheorie gehen die Schwellenmodelle nicht von einem Kontinuum der Empfindungsstärke aus, sondern von Schwellen, welche die Wahrnehmung der Reizempfindung bestimmen. Die Modelle integrieren die kognitiven Prozesse der Reizeentdeckung, Quellendiskrimination und der Antworttendenz. In Abbildung 1 sind sieben Parameter im multinomialen *Two-High-Threshold* Modell von Bayen et al. (1996) der Itemerkennung und Quellenidentifikation der Quellen A, B und neuer Items N dargestellt. In dem Modell lassen sich zwei Schwellen ablesen, je eine Hochschwelle für alte und neue Items (Bayen et al., 1996). Dies modelliert Rekognition und Quellenidentifikation, wobei für das Quellengedächtnis eine Schwelle überschritten werden muss, um sicher zu sein, dass ein Item aus Quelle A stammt (Hochschwelle 1), dies gilt in gleicher Form für Items der Quelle B (Hochschwelle 2). Hinsichtlich der Alt-Rekognition wird ebenfalls angenommen, dass eine Schwelle bzgl. des Vorkommens des Items überschritten werden muss und zusätzlich eine Schwelle für das Nichtvorkommen, wonach es sich bei der Identifikation neuer Items um mehr handelt als um den Ausschluss der Alt-Rekognition.

Ableitungen für die Methodologie der vorliegenden Untersuchung

Für die vorliegende Untersuchung gilt vor der Auswertung zu prüfen, ob Unterschiede in der Stärke der Repräsentation in den experimentellen Bedingungen angenommen werden müssen. Modelle sind hilfreich, wenn von unterschiedlicher Repräsentationsstärke ausgegangen wird; die multinomiale Modellierung macht eine Vielzahl an Prämissen und nimmt den Rateparameter a für verschiedene Quellen als gleich an, womit

das Modell als überdeterminiert angesehen werden kann. Für die vorliegende Untersuchung kann jedoch nicht davon ausgegangen werden, dass unterschiedliche Quellen in unterschiedlicher Gedächtnisstärke repräsentiert sind. Es wird von einer konstanten Gedächtnisstärke der Teilnehmer innerhalb der experimentellen Untersuchung ausgegangen. Es können keine Ursachen dafür identifiziert werden, dass sich die Gedächtnisstärke zwischen den Bedingungen systematisch unterscheidet, da kontrollierte und standardisierte Darstellungen, Inhalte und Befragungsbedingungen realisiert wurden. Die Speicherungsbedingungen können demnach für alle als identisch angenommen werden. Damit wirkt sich die Bedingungsvariation nur auf die Quellenidentifikationsleistung aus. Multinomiale Modelle (siehe Abbildung 1) modellieren den Erinnerungsprozess für Quellenidentifikationsaufgaben zweier abhängiger Variablen, Gedächtnisstärke und Ratetendenz. Bei beiden Arten der Messung des Source Monitorings wird von Vorannahmen sowie einer Fehlerbehaftung in den Randbereichen ausgegangen. (A)CSIM integriert Messungen verschiedener Quellen in die Grundgesamtheit und nimmt Gedächtnisstärke und Ratetendenz nicht getrennt in die Messung auf.

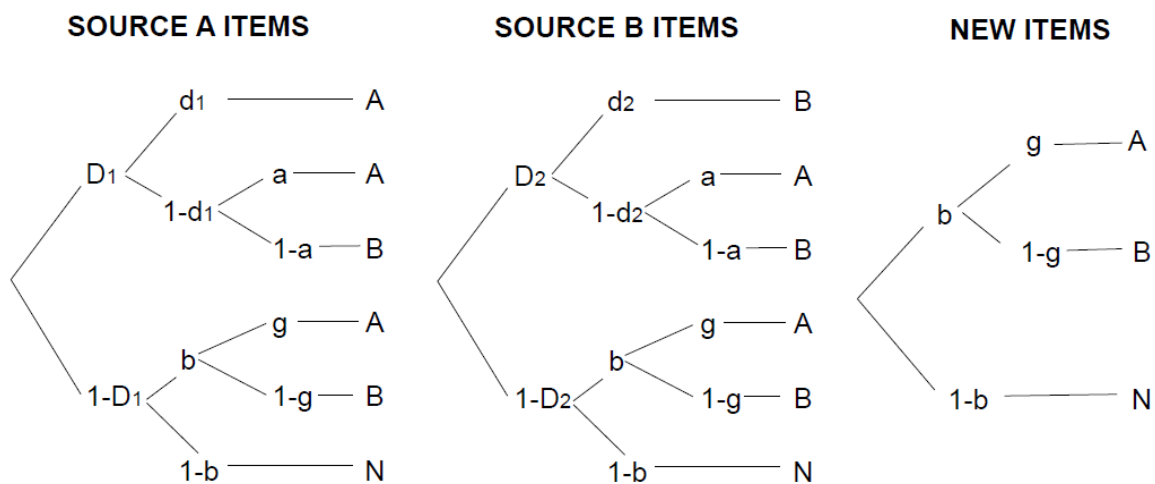


Abbildung 1. Multinomiales Modell der Quellenidentifikation mit sieben Parametern: D1 = Erinnerungsstärke für Items Quelle A; D2 = Erinnerungsstärke Items Quelle B; d1 = Quellendiskrimination für Items der Quelle A; d2 = Quellendiskrimination für Items der Quelle B; a = Ratewahrscheinlichkeit, dass ein entdecktes, aber nicht diskriminiertes Item zu Quelle A gehört; b = Tendenz ein nichtentdecktes Item als „alt“ zu identifizieren; g = Ratewahrscheinlichkeit ein nicht entdecktes Item als Item der Quelle A zu identifizieren (nach Batchelder & Riefer, 1990).

1.2 Gedächtnisentwicklung und Source Monitoring bei Kindern

Aufgrund des Mangels an eigenen Erfahrungen sind Kinder in besonderem Maße auf die Berichte, Erlebnisse und Erinnerungen anderer Menschen als Grundlage ihrer Überzeugungen und ihres Wissens angewiesen. Gleichzeitig sind die Gedächtnisfunktionen von Kindern im Vergleich zu Erwachsenen zunächst noch defizitär (Büttner, 2003). In diesem Sinne ist es nicht verwunderlich, dass Kinder auch hinsichtlich ihrer Quellenidentifikationsleistungen Einschränkungen gegenüber Erwachsenen aufweisen (Roberts & Blades, 2000).

Im Rahmen der Erforschung von Gedächtnisprozessen wurden vier Faktoren identifiziert, welche die Entwicklung der Gedächtnis- und Erinnerungsleistung beeinflussen: Kapazität, Wissen, Metakognition und Strategiegebrauch (Büttner, 2003). Bevor also die Gedächtnisleistung im Erwachsenenalter relativ stabil bleibt, sind relevante Entwicklungsschritte zu verzeichnen, welche insbesondere in der frühen Kindheit stattfinden. Dem Vorwissen wird dabei eine besondere Funktion zugeschrieben; dieses stellt eine Voraussetzung für den Erwerb neuen Wissens dar. Hervorzuheben ist die steigende Verfügbarkeit gedächtnisgebundener Strategien, die Herausbildung des Metagedächtnisses sowie die zunehmende Arbeitsgedächtniskapazität. Besondere Fortschritte erfolgen in den ersten zwei Lebensjahren und dem Jugendalter, diese sind jedoch interindividuell verschieden und nicht verallgemeinerbar (Büttner, 2003).

Mit etwa zwei Jahren können Kinder Fragen zur Vergangenheit beantworten (Fivush & Hammond, 1990). Erst ab einem Alter von etwa sechs Jahren erzielen Kinder vergleichbare Leistungen wie Erwachsene in einer Quellenidentifikationsaufgabe (Lindsay, Johnson & Kwon, 1991). Kinder sind mehr als Erwachsene von der Abfragebedingung abhängig und tendenziell leichter zu beeinflussen (Roebbers, 2010). Laut Foley und Johnson (1985) lassen sich Unterschiede in den Kompetenzen von Kindern hinsichtlich der verschiedenen Arten der Quellenidentifikationsleistung verzeichnen. Dabei erreicht das Reality Monitoring als erstes ein mit den Fähigkeiten Erwachsener vergleichbares Niveau. Hervorzuheben ist dabei, dass das Alter nicht als Prädiktor für die Fähigkeiten der Quellenidentifikation zu interpretieren ist, sondern lediglich einen möglichen Indikator darstellt (Roberts & Blades, 2000).

Die internationale Forschung der vergangenen Jahre zum Source Monitoring bei Kindern lässt sich grob in drei Kategorien unterteilen. Dazu zählen erstens Personenmerkmale, zweitens Charakteristika der Lernphase bzw. Rezeptionssituation und drittens Umstände der Abruf- bzw. Befragungssituation. Die im Folgenden vorgestellten Studien wurden durch eine systematische Literaturrecherche in den einschlägigen Datenbanken, PsycInfo und Psynindex ausgewählt, wobei entsprechende Schlagworte, wie Kinder, Children und Source Monitoring verwendet wurden. Die Ergebnisse der im Folgenden dargestellten Studien werden nach den o.g. Kategorien strukturiert und im Hinblick auf die Effektstärken Cohens d (Cohen, 1988; Riehle & Zwingmann, 2000) und/oder das partielle Eta-Quadrat η^2 (Eid, Gollwitzer & Schmitt, 2013; siehe Tabelle 3) berichtet. Dabei wird zwischen kleinen, mittleren und großen Effekten unterschieden. Hierbei wurden d und das partielle η^2 für fehlende Angaben von der Autorin selbst berechnet, soweit nicht anders angegeben.

Tabelle 3. Interpretation der Effektstärken (nach Cohen, 1988; Eid et al, 2013).

Interpretation	d	η^2
kleiner Effekt	> 0,2	> 0,01
mittlerer Effekt	> 0,5	> 0,06
großer Effekt	> 0,8	> 0,14

Anmerkung: d = Cohens d ; η^2 = partielles Eta-Quadrat.

Aus differentialpsychologischer und entwicklungspsychologischer Perspektive werden unter Personenmerkmalen und Persönlichkeitsmerkmalen alle Eigenschaften einer Person verstanden. Personenmerkmale sind Eigenschaften, wie Geschlecht oder sozioökonomischer Status. *Persönlichkeitsmerkmale* umfassen latente, nicht direkt beobachtbare Eigenschaften. Diese werden als hypothetische Konstrukte beschrieben, durch Indikatoren operationalisiert und damit messbar gemacht. Dazu zählen beispielsweise Intelligenz, Interesse und Motivation. Gemäß Deci und Ryan (1993) sind das Streben nach Autonomie und Selbstbestimmung, das Kompetenzerleben und das Streben nach sozialer Eingebundenheit im Rahmen der Theorie der Selbstbestimmung zentrale Komponenten der Lern- und Leistungsmotivation.

Eine Gruppe um Haigh und Robinson (Haigh & Robinson, 2009; Robinson, Haigh & Nurmsoo, 2008) untersuchte den Bewusstseitsgrad bei der Verarbeitung von Quelleninformationen vor dem Hintergrund des individuellen Wissensstands bei drei- bis fünfjährigen Kindern ($N = 117$). Dabei stellte sich heraus, dass ein Vorteil bei der Beantwortung bestand, wenn sich die Kinder nicht auf Hinweise eines Versuchsleiters verlassen mussten ($d = 0,39$ bis $d = 2,5$) und auf ihre eigene Wissensgrundlage zurückgreifen konnten. Auf Basis der Ergebnisse schlussfolgerten die Autoren, es gebe erstens einen automatischen Prozess, bei dem indirekt erlangten Überzeugungen so wie selbst erworbenem Wissen vertraut werde, bis es offensichtliche Gründe zum Zweifel gebe. Zweitens betrachteten Kinder die spezifischen Implikationen von geäußerten Zweifeln des Versuchsleiters als Bewertungsgrundlage zur Einschätzung der Sicherheit ihrer Aussagen.

Hala, Brown, McKay und San Juan (2013) untersuchten Reality und Source Monitoring Leistungen anhand des *I did it* Bias. Dieser beschreibt, dass Kinder beobachtete Handlungen eines Versuchsleiters im Nachgang eher als selbst ausgeführt identifizieren. Dabei konnten die Autoren feststellen, dass bereits zweieinhalb jährige Kinder ($N = 50$) zum Reality Monitoring in der Lage sind und demnach auch kognitive Bindungsprozesse für Handlungen und Akteure stattfinden ($d = 0,72$). Im Vergleich zu älteren Kindern ordneten die jüngeren trotz kontextuellen Hinweisreizen, Handlungen des Versuchsleiters häufiger fälschlicherweise sich selbst zu, waren also dem *I di it* Bias unterworfen (Hala et al., 2013).

Zu den Personenmerkmalen im Rahmen des Source Monitorings untersuchten Hala, Rasmussen und Henderson (2005) den Zusammenhang von Exekutivfunktionen, hier insbesondere der Arbeitsgedächtniskapazität, und der Rekognitionsleistung sowie der Quellenidentifikationsleistung bei autistischen im Vergleich zu normal entwickelten Vorschulkindern ($N = 26$). Die Autoren fanden Belege dafür, dass keine generellen Einschränkungen der Arbeitsgedächtniskapazität autistischer Kinder im Vergleich zur Kontrollgruppe bestehen, sondern ein Nachteil bei der Kombination von Exekutivfunktionen, wie es bei der Quellenidentifikation notwendig ist, vorliegt (Reality Monitoring: $d = 0,05$; externes Source Monitoring $d = 0,04$; internes Source Monitoring $d = 0,21$).

Sluzenski, Newcombe und Ottinger (2004) beschäftigten sich in drei Teilstudien mit dem Zusammenhang zwischen der Entwicklung des episodischen Gedächtnisses und Reality Monitoring bei Kindern im Rahmen komplexer Ereignisse in einem feldexperimentellen Design. Zudem sollte die Verbindung mit den exekutiven Funktionen eruiert werden. Dazu wurden vier-, sechs- und achtjährigen Kindern ($N = 60$) Tonbandaufnahmen vorgespielt, in denen sie dazu aufgefordert wurden, bestimmte Handlungen, wie beispielsweise Pudding kochen, auszuführen oder sich diese nur vorzustellen. Nach einer Latenz von einer Woche wurden sie dazu befragt, welche Handlungen sie sich vorgestellt und welche sie tatsächlich ausgeführt hatten. Zudem sollten sie alles berichten, woran sie sich bezüglich der Handlung erinnern konnten, bevor ihnen Fragen zu spezifischen perzeptuellen, räumlichen, zeitlichen und semantischen Aspekten zur Bewertung der Erinnerungsleistung gestellt wurden. Die Exekutivfunktionen wurden anhand einer *Category Fluency Task* erhoben, bei der in begrenzter Zeit so viele Worte bzw. Repräsentanten wie möglich zu einer vorgegebenen Kategorie erdacht werden müssen (Nolde, Johnson & D'Esposito, 1998; Sluzenski et al., 2004). Nolde et al. (1998) konnten zeigen, dass Category fluency tasks zu den Gedächtnisaufgaben zählen, welche hohe Anforderungen an die Exekutivfunktionen stellen. Die vierjährigen Teilnehmer zeigten in allen Studien Defizite im Reality Monitoring gegenüber den älteren Versuchsteilnehmern, wobei signifikant mehr Erinnerungen an reale als an vorgestellte Ereignisse berichtet wurden ($d = 1,44$). Die Auswertung erfolgte varianzanalytisch, dabei wurde das Alter als Intersubjektfaktor verwendet (Sluzenski et al., 2004). Der sehr große Alterseffekt wurde durch die Latenz zwischen Lernphase und Erhebung mediiert und nahm bei direkter Testung im Anschluss an die Lernphase ab. Die Autoren schlussfolgerten, dass das Defizit durch die geringere Gedächtniskapazität der jüngeren Versuchsteilnehmer bedingt wurde. Gleichzeitig hatten die achtjährigen Kinder im Vergleich über alle Messzeitpunkte hinweg einen Vorteil gegenüber den anderen Altersklassen, wobei sich die vier- und sechsjährigen hier nicht signifikant unterschieden. In einem zweiten Experiment sollte untersucht werden, ob die Unterschiede im Reality Monitoring durch schlechte Gedächtnisleistungen oder geringe Strategieverwendung zustande kommen. Dafür wurde ein identischer Versuchsaufbau wie im ersten Experiment verwendet, nur wurden sämtliche Maße in derselben Phase erhoben, so dass keine Latenz zwischen Lern- und Testphase vorlag. Dabei zeigte sich, dass in Experiment 2 nur 25 % der Teilnehmer Fehler in der Reality Monitoring Aufgabe begingen im Vergleich zu 75 % nach einer Woche Delay, was die Autoren mit

der Attenuationstheorie des Gedächtnisses von Treisman (1964) erklärten. Die Attenuationstheorie ist eine Aufmerksamkeitstheorie und kann die attenuierte, d.h. abgeschwächte Weiterleitung von Informationen erklären, welche nicht direkt beachtet wurden, aber erinnert werden. Die Befunde sprechen laut Sluzenski et al. (2004) dafür, dass Defizite in der *Strategieverwendung* ursächlich für die fehlerhafte Quellenzuschreibung der Vier- und Sechsjährigen sind. Zusätzlich konnten Zusammenhänge zwischen den Reality Monitoring Leistungen und den Kapazitäten des autobiographischen Gedächtnisses gefunden werden. Es konnte keine Korrelation zwischen der Category Fluency und dem Reality Monitoring, jedoch ein Zusammenhang zwischen Category Fluency und den freien Erinnerungsleistungen nachgewiesen werden.

Im Hinblick auf die Personenmerkmale beim Source Monitoring kann zusammenfassend festgehalten werden, dass Kinder spätestens ab einem Alter von zweieinhalb Jahren zur Quellendiskrimination in der Lage sind, dabei aber dem *I did it*-Bias unterliegen. Dabei wurden die Ergebnisse auch durch sprachliche Kompetenzen und kognitive Bindungsprozesse beeinflusst (Hala et al., 2013). Eindeutige Entwicklungsschritte im Reality und Source Monitoring sind in der Kindheit zu verorten (Haigh & Robinson, 2009; Hala et al., 2005; Sluzenski et al., 2004); diese spiegeln sich auch in den zugrundeliegenden Entscheidungsprozessen wieder (Robinson et al., 2008). Übereinstimmend konnte die Bedeutung der Exekutivfunktionen bei der Quellenidentifikation identifiziert werden. Ein genereller Nachteil jüngerer Kinder ist nicht allein durch die geringere Gedächtniskapazität zu begründen, sondern vielmehr durch die fehlende Strategieverwendung, welche über das Arbeitsgedächtnis und die inhibitorische Kontrolle die indirekte Verknüpfung zum episodischen Gedächtnis repräsentiert wird. Als relevant erwiesen sich auch verbale Fähigkeiten, kognitive Bindungsprozesse (Hala et al., 2013) sowie die zeitliche Latenz zwischen Lern- und Testphase (Hala et al., 2005; Sluzenski et al., 2004). Bezogen auf die interne Validität der Ergebnisse muss jedoch angemerkt werden, dass insbesondere individuelle Merkmale, wie die Strategieverwendung (Johnson et al., 1993) als auch Vorwissen (Büttner, 2003) wichtige Determinanten der Gedächtnisleistung und damit auch der Quellenidentifikation sind. Nicht in allen Studien wurde bereichsspezifisches Vorwissen kontrolliert, zudem wurde vornehmlich kindgerechtes, relativ alltägliches Testmaterial verwendet (z.B. Pudding kochen), so dass die gelernten Inhalte möglicherweise mit eigenen Erfahrungen angereichert bzw. verknüpft wurden und damit den interindividuellen Vergleich verzerrt

haben. Um valide Aussagen über den Entwicklungsverlauf der Fähigkeit zum Reality und Source Monitoring treffen zu können, ist grundsätzlich der Einsatz von Studien mit längsschnittlichem Design vorzuziehen, dem wurde jedoch nicht nachgekommen.

Die Einflussvariablen auf Seiten der Lernphase bzw. Rezeptionssituation beziehen sich unter anderem auf die Aufmerksamkeitslenkung oder Kontextgestaltung, darunter fällt aber auch die Verwendung bestimmter Lern- und Versuchsmaterialien. Kovacs und Newcombe (2006) sowie Crawley, Newcombe und Bingman (2010) untersuchten die emotionale Fokuslenkung und deren Auswirkung auf die kindliche Quellenidentifikationsleistung ($N = 190$ vier- und fünfjährige Kinder), indem eine Beschreibung der eigenen Emotionen (*self focus*) bzw. der externen Emotionen der Quelle (*other focus*) in der Lernphase gefordert wurde. Nur die älteren Kinder hatten bei Nutzung des *other focus* einen Vorteil beim Erinnern der Quelle (Kovacs & Newcombe, 2006: $d = 2,54$; Crawley et al., 2010: $d = 1,74$). Bei äußerlich gut diskriminierbaren Quellen ergab sich eine Verbesserung der Quellenidentifikationsleistung der Vierjährigen durch die Fokuslenkung ($d = 0,83$). Crawley et al. (2010) erweiterten die Studien um zwei Bedingungen, den *perceptual self* und *perceptual other focus* mit einer Stichprobe von $N = 240$ vier- und sechsjährigen Kindern. Es konnte gezeigt werden, dass die Fokussierung auf Merkmale, wie z.B. Haarfarbe, in Bezug auf die Quellenidentifikation hinsichtlich altersbezogener (z.B. emotional-self: $d = 1,41$) und emotionaler Fokuseffekte ($d = 1,8$; sehr große Effektstärke) vergleichbare Befunde zu Kovacs und Newcombe (2006) erzielt. Die Quellenidentifikationsleistungen in der *perceptual other focus*-Bedingung waren besser als in der *perceptual self focus*-Bedingung ($d = 0,63$). Die Quellenidentifikation in den Bedingungen mit einem perzeptuellen Fokus war insgesamt schlechter als bei einem emotionalen Fokus ($d = 1,2$).

Sugimura (2008) interessierte sich für den Einfluss der Berücksichtigung von Quellenmerkmalen in komplexen und alltagsnahen Situationen und deren Einfluss auf die Quellendiskrimination sowie den Zusammenhang mit der generellen Identifikationsleistung. In der Studie wurden Vorschulkindern und Erwachsenen ($N = 97$) Zaubertricks von ähnlich aussehenden, aber mit verschiedenen Attributen (beispielsweise farbiges Shirt, Kappe, Shorts) ausgestatteten Frauen präsentiert. Einen Monat später sollten die Probanden erstens die Tricks, die sie gesehen hatten, wiedererkennen (Rekognition), die Zauberinnen unabhängig von ihren Tricks erkennen (Identifikation), die Zauberinnen den Tricks zuordnen (Source Monitoring) und die Attribute den Frauen

zuordnen (Source Monitoring: Attributzuordnung). Dabei zeigten die Erwachsenen eine signifikant bessere Rekognition und Quellendiskrimination als die Kinder ($\eta^2 = 0,18$), die Rekognitionsleistung war über beide Altersgruppen hinweg höher als die Quellendiskriminationsleistung ($\eta^2 = 0,52$).

Auch Thierry und Pipe (2009) untersuchten die Quellenidentifikationsleistung von Kindern im Rahmen realer Erlebnisse und variierten dabei die Ähnlichkeit verschiedener Quellen. Dazu wurden drei- bis vierjährigen und fünf- bis sechsjährigen Kindern ($N = 96$) live naturwissenschaftliche Experimente vorgeführt. Im Anschluss daran sah die eine Hälfte der Kinder ein Video von einem vergleichbaren Experiment (hohe Quellenähnlichkeit), der anderen Hälfte wurde mithilfe eines Bilderbuchs von einem ähnlichen Experiment erzählt (geringe Quellenähnlichkeit). Die älteren Kinder erbrachten bessere Quellenidentifikationsleistungen als die jüngeren ($\eta^2 = 0,22$), zudem konnten die älteren Kinder besser zwischen unähnlichen Quellen unterscheiden ($\eta^2 = 0,99$) wenn die Testphase erst nach vier Tagen stattfand (Exp. 2), konnten beide Effekte repliziert werden ($\eta^2 = 0,46$; $\eta^2 = 0,21$), jedoch wurden Interaktionen zwischen Alter, Quellenähnlichkeit und Ereignistyp hier signifikant ($\eta^2 = 0,04$). Der Effekt der Quellenähnlichkeit zeigte sich bezüglich der realen Erlebnisse nur bei den jüngeren Kindern. Die älteren Kinder konnten, sowohl wenn die alternative Quelle die Geschichte war als auch wenn die alternative Quelle das Video war, die Items des realen Erlebnisses korrekt zuordnen. Die Zuordnung der alternativen Quelle verbesserte sich hingegen für beide Altersklassen, wenn sie dem realen Erlebnis unähnlich war.

In Anbetracht der dargestellten Befunde zur Wirkung von Merkmalen der Lernphase lässt sich festhalten, dass der Aufmerksamkeitsfokus während der Enkodierung eine wichtige Rolle für die Quellenidentifikationsleistung spielt. Vor dem Hintergrund der Erkenntnisse von Johnson et al. (1993), dass zwei Prozesse des Reality und Source Monitorings unterschieden werden, von denen einer unbewusst auf der Basis von Vorwissen verläuft, ist der Leistungsvorteil bei der Beachtung von spezifischen Quellenmerkmalen nicht überraschend. Johnson et al. (1993) zufolge nutzen Kinder vornehmlich den automatischen, unbewussten Identifikationsprozess, womit dem Bewusstheitsgrad eine stärkere Beachtung bei der weiteren Untersuchung von Monitoring Prozessen zukommen sollte. Durch die Studien von Crawley et al. (2010) wird einerseits deutlich, dass ein emotionaler Fokus die positive Wirkung einer externalen

Aufmerksamkeitslenkung auf die Quelle besonders verstärkt. Die Ergebnisse von Bemis, Leichtmann und Pillemer (2013) und Sugimura (2008) weisen dennoch auch auf die Bedeutung perzeptueller Quellenmerkmale hin. Bezüglich dieser Merkmale verdeutlichen die Studien von Thierry und Pipe (2009), dass eine große Unterscheidbarkeit der Quellen deren Identifikation begünstigt. Zudem sprechen die vorliegenden Befunde für eine Entwicklung der Quellenidentifikationsfähigkeiten von Kindern während des Vorschulalters.

Sowohl die Ergebnisse von Thierry und Pipe (2009) als auch die Befunde von Kovacs und Newcombe (2006) zeigen eine besondere Bedeutung der guten Unterscheidbarkeit verschiedener Quellen für sehr junge Kinder im Vergleich zu Grundschulkindern. Bezüglich der hier dargestellten Zusammenhänge sind mit $d = 1,8$ (Crawley et al., 2010) oder $\eta^2 = 0,46$ (Sugimura, 2008) mittlere bis große Effekte berichtet worden. Dabei ist besonders auffällig, dass die Befunde aus den Studien der Forschergruppe um Newcombe (2008, 2010) und dem Experiment von Sugimura (2008), die sich alle mit der Aufmerksamkeitssetzung beschäftigten, durchweg große Effekte erzielten. In dieser Hinsicht bekommen die Aussagen hinsichtlich der Wirkung einer spezifischen Fokussierung während der Lernphase eine besondere Relevanz. Es muss jedoch in Bezug auf die externe Validität der Ergebnisse beachtet werden, dass die Intervalle zwischen Lern- und Testphase in den Untersuchungen zwischen zwei Minuten (Robinson et al., 2008) und einem Monat (Sugimura, 2008) variierten und demnach nur bedingt auf reale Ereignisse übertragbar sind. Die Präsentation des Wissens fand teils im Einzel- (Hala et al., 2005), teils im Gruppensetting (Kovacs & Newcombe, 2006) statt, was einen Einfluss auf die Rezeption und Speicherung der Inhalte ausüben kann. Die Modalität, über welche das Wissen erlangt wurde, variierte zwischen Tonband, Video, Vorlesen (Bright-Paul, Jarrold & Wright, 2005) oder live bzw. direkte Interaktion (Hala et al., 2013) und bestimmt somit die Komplexität der Lernsituation und die Realitätsnähe, wobei bei einer Steigerung dieser auch die externe Validität zunimmt. So bestehen gemäß Foley und Johnson (1985) intraindividuelle Unterschiede im internalen/externalen Source Monitoring, deren Zusammenhang mit der Entwicklung exekutiver Funktionen weitere Beachtung finden sollte.

Besonderer Bedeutung im gedächtnispsychologischen und vor allem forensischen Kontext kommt Merkmalen der Befragungssituation zu. Merkmale der Befragungssituation beziehen sich auf die konkrete Befragung, deren Kontext, Fragenformulierungen oder auch auf den Interviewer.

Im Jahr 2005 untersuchten Shapiro und Purdy neben anderen Variablen die Auswirkungen von Interviewstil bzw. Fragetechnik und Interviewerkonstanz auf die Quellenidentifikationsleistung und Suggestibilität von fünf- bis achtjährigen Kindern. Sechzig Teilnehmer sahen Videos von Diebstahlszenen und wurden direkt im Anschluss zu Details der Situation (*true*) sowie zu Fehlinformationen (*false*) interviewt, wobei sie vorab instruiert wurden, dass bei Befragungen im polizeilichen Kontext alles genannt werden soll, woran man sich erinnert. Im Rahmen der Befragung wurden die Kinder gebeten, sich falsche Details auszudenken (*forced confabulation*), wenn sie die Antwort nicht wussten, oder ihnen wurden falsche Details suggeriert (*suggestive Bedingung*). Nach sieben Tagen Latenz wurde die Quellenidentifikationsleistung mit vier verschiedenen Itemtypen, *true*, *false*, *control/new* und *confabulated*, getestet. In der MANOVA für wiederholte Messungen erwies sich die Suggestibilität in der forced confabulation-Bedingung höher als in den suggestiven Bedingungen ($\eta^2 = 0,75$). Für die Quellenidentifikation zeigte sich für die *false* Items eine Zwei-Weg-Interaktion zwischen Itemtyp und Interviewstil ($\eta^2 = 0,10$) und eine Drei-Weg-Interaktion zwischen Itemtyp, Interviewstil und Interviewerkonstanz ($\eta^2 = 0,05$). Die Teilnehmer konnten *true* und *false* Items nicht zuverlässig ihren Quellen zuordnen; in der Fehleranalyse wurde deutlich, dass Items einer Quelle beiden Quellen zugeordnet wurden, wobei diese Misattributionen auch dadurch zustande gekommen sein könnten, dass die Quellen sich nicht gegenseitig ausgeschlossen haben. In der Bedingung forced confabulation wurden bei den *confabulated* Items im Interview mehr Quellenfehler gemacht wurden, wenn Interviewerkonstanz gegeben war; der Effekt zeigte sich für die *control* Items nur bei der Befragung durch verschiedene Interviewer. Für die suggestive Bedingung ergaben sich gegenteilige Effekte: Es zeigten sich mehr Fehler für *confabulated* Items bei verschiedenen Interviewern und mehr Fehler bei *control* Items und Interviewerkonstanz. In der Gesamtschau wird deutlich, dass erzwungene Konfabulation in Interviewsituationen zu höherer Suggestibilität führt als die Präsentation von Falschinformationen (Shapiro & Purdy, 2005). Des Weiteren konnte die Bedeutung des Interviewstils und der Interviewerkonstanz auf die Quellenidentifikationsleistung in der

Befragung herausgestellt werden, nämlich ob die Quelle das Video oder das direkt angeschlossene Interview war. Die Personen in der forced confabulation-Bedingung ordneten die Items eher dem Video als Quelle zu, als die Personen in der Suggestionsbedingung.

In einer methodisch vergleichbaren Studie von Bright-Paul, Jarrold und Wright (2005) wurde der kausale Zusammenhang zwischen der Quellenidentifikation und der Befragungsmethode genauer betrachtet ($N = 116$; drei bis sieben Jahre). Erhoben wurde die Rekognition (Ja/Nein) und Quellenidentifikation (Film/Geschichte). Zusätzlich wurde die simultane oder sequentielle Darstellung der Quellen variiert. Es zeigte sich, dass jüngere Kinder häufiger Fehlattributionen bei der Quelle als ältere ($\eta^2 = 0,37$) machen. Bei zeitgleicher Darbietung aller Quellen zur Identifikation der Items, wurden weniger Fehler gemacht als in der reinen Rekognitionsbedingung (jüngere Kinder: $d = 1,35$; ältere Kinder: $d = 0,86$). Thierry, Lamb, Pipe und Spence (2010) wiesen nach, dass bereits kleine Kinder ($N = 105$; drei bis vier Jahre) von einem Training zur Verbesserung der Quellenidentifikationsleistung profitieren können und durch die Aufmerksamkeitslenkung auf Quellenmerkmale besser zwischen selbst erlebten und berichteten Ereignissen ($\eta^2 = 0,11$) unterschieden. Vergleichbar mit Bright-Paul et al. (2005) beschäftigten sich auch Thierry, Goh, Pipe und Murray (2005) mit der Wirkung einer quellenorientierten Befragung auf die Quellenidentifikationsleistung von Kindern. Zusätzlich untersuchten sie die Effekte der Test-Modalität (verbal vs. Bild). Sieben- und achtjährigen Kindern ($N = 30$) wurden zwei verschiedene Geschichten eines Piraten via Bild oder ausschließlich verbal präsentiert und zwei Tage später dazu befragt. In der *Source Recall*-Bedingung sollten die Ereignisse getrennt für die beiden Quellen Bild und Geschichte berichtet werden - in der *No Source Recall*-Bedingung wurde frei befragt. Nach einer Woche sollten die Kinder entweder via Bild oder durch verbale Hinweise die Quellenidentifikation der Geschichtselemente vornehmen. Es zeigte sich, dass der *Source Recall* die spätere Quellendiskrimination im Vergleich zum *No Source Recall* ($d = 2,14$) und einem ausbleibenden Recall bzw. nicht Erinnern ($d = 0,98$) begünstigte. Dabei wurde jedoch nur die Rekognition, nicht aber die Quellenidentifikation positiv beeinflusst. In einem zweiten Experiment wurde die Diskriminierbarkeit zwischen den Bild- und Geschichtselementen erhöht und das Intervall zwischen Recall und Quellenidentifikationstest verringert, was insgesamt zu besseren Ergebnissen

führte. Somit konnte analog zu Thierry, Spence und Memon (2001) ein aufgabenbezogener Effekt zur Vereinfachung des Source Monitorings nachgewiesen werden (Thierry et al., 2005). Die Ergebnisse legen eine explizite Aufforderung zur Unterscheidung zwischen vorgestellten, erlebten und gehörten Inhalten nahe.

Insgesamt kann die Quellenidentifikationsleistung durch die Gestaltung der Befragungssituation beeinflusst und damit verbessert oder erleichtert werden. So kann die Quellenidentifikationsleistung u.a. durch eine simultane Darstellung der Quellen gesteigert werden (Bright-Paul et al., 2005). Ein Training zur Lenkung der Aufmerksamkeit auf Quellenmerkmale in der Rezeptionsphase kann die spätere Identifikation und Zuordnung verbessern und erweist sich bereits bei sehr jungen Kindern als wirksam (Thierry et al., 2010). Eine Verbesserung der Quellenidentifikation kann ebenfalls durch einen quellenfokussierten Recall erzielt werden (Thierry et al., 2005). Eine forcierte Konfabulation basierend auf Unwissen führt bei Interviewerkonstanz zu einer tieferen Verarbeitung und Assoziation der Fehlinformation mit dem Interviewer und somit zu schlechteren Leistungen (Shapiro & Purdy, 2005). In Anbetracht der Suggestibilität lässt sich festhalten, dass diese bei forcierten Antworten und dem Fehlen einer *IWN*-Antwortoption zunimmt (Shapiro & Purdy, 2005). Im Kontext der Studien, im Rahmen derer Merkmale der Befragungssituation untersucht wurden, ist zweifelsohne zu diskutieren, inwieweit eine fehlende Differenzierung zwischen Quelle und Versuchsleiter (Crawley et al., 2010; Robinson et al., 2008) dazu führen kann, dass dieser als Hinweisreiz wahrgenommen und die Quellenidentifikationsleistung erleichtert wird. Die Effektgrößen der Studien zur Verbesserung der Quellenidentifikationsleistung durch Manipulation der Befragungssituation erweisen sich als relativ groß, so dass eben diese in besonderem Maße Beachtung finden sollten.

In der Gesamtschau lassen sich sowohl Personenmerkmale als auch Faktoren der Lern- und Befragungssituation feststellen, welche die Quellenidentifikationsleistung von Kindern beeinflussen. Die empirischen Studien weisen eine große Vielfalt an methodischen Zugängen auf, welche dem forensisch relevanten Kontext entsprechen, zeitgleich aber auch die Übertragbarkeit der Ergebnisse einschränken. Die unterschiedlichen kognitiven Anforderungen des Reality Monitoring im Vergleich zum externalen und internalen Source Monitoring zeigen sich sowohl bei normal entwickelten als auch bei autistischen Kindern (Hala et al., 2005). Besonders Defizite in den exekutiven Funktionen spielen eine Rolle für die Probleme von Kindern, die Quellen ihres

Wissens zu identifizieren. Durch eine Manipulation der Aufmerksamkeitslenkung auf perzeptuelle und emotionale Merkmale der Quelle während der Lernphase (z.B. Kovacs & Newcombe, 2006) oder der Befragungsphase (Bright-Paul et al., 2005) kann die Identifikationsleistung von Kindern verbessert werden. Für Kinder ist ein Benennen der Quelle ihrer Information nicht nötig, um die Fehlbarkeit dieser auf die Evaluation des eigenen Wissensstands anzuwenden (Haigh & Robinson, 2009).

2 Kognitive Aspekte der Zeugenaussage

Zeugenaussagen zählen zu den wichtigsten und häufigsten Beweismitteln zur Sachaufklärung in rechtlichen Fällen (Loohs, 1996). Die Zuverlässigkeit von Augenzeugenaussagen ist jedoch stark eingeschränkt, infolgedessen Fehlurteile entstehen (Arntzen, 1993; Arntzen, 2011; Innocence Project, o.D.).

2.1 Validität von Augenzeugenaussagen

Die Validität von Augenzeugenaussagen erweist sich in einer Vielzahl von Untersuchungen als gering (Oberlader, Naefgen, Koppehele-Goseel, Quinten, Banse & Schmidt, 2016; Volbert, Steller & Galow, 2010). Grundsätzlich lassen sich die methodischen Ansätze dahingehend differenzieren, dass ein freier oder unterstützter Abruf forciert wird oder eine Wiedererkennungslleistung erfolgen soll. Im Folgenden werden die kognitionspsychologischen Grundkonzepte übersichtsartig dargestellt, die einen Einfluss auf die Qualität und den Umfang von Zeugenaussagen haben.

2.2 Augenzeugengedächtnis und autobiografisches Gedächtnis

Eine wichtige Kompetenz, um vollständige und genaue Aussagen tätigen zu können, ist die sog. Aussagetüchtigkeit. Sie beschreibt die Fähigkeit zur adäquaten Situationswahrnehmung, zum Erinnern, d.h. Speichern des Wahrgenommenen über einen längeren Zeitraum sowie zur Reproduktion dieses Inhalts speziell im forensischen Kontext. Letzteres umfasst Quellenmonitoring, die Identifizierung des Ursprungs der jeweiligen Erinnerung, um sie zu verstehen und zu validieren (Ceci & Bruck, 1993), den weitgehend selbstständigen Informationsabruf, das sprachliche Ausdrucksvermögen und andere kommunikative Fähigkeiten, wie z.B. eine nachvollziehbare, valide Schilderung abzugeben. Zu beachten ist daher, dass es sich bei der Aussagetüchtig-

keit nicht etwa um ein zeitlich überdauerndes, feststehendes Konstrukt handelt, sondern um ein Können das inter- und intraindividuell variabel sein kann und nicht immer in tatsächliches Handeln umgesetzt werden muss (Volbert, 2005). Zunächst wird von einer grundsätzlich vorhandenen Aussagetüchtigkeit ausgegangen, diese wird bei Hinweisen auf psychische (z.B. Intelligenzminderung) oder entwicklungsbedingte Besonderheiten (z.B. junges Alter) genauer hinsichtlich möglicher Einschränkungen geprüft.

Persönliche Erlebnisse, mit atypischem und einmaligem, vom Alltag und der Normalität abweichendem Charakter, zeichnen sich durch eine hervortretende Gedächtnisposition aus. Details an das Geschehen werden besser erinnert, sind leichter abrufbar und zuverlässiger als typische Ereignisse ohne persönliche Bedeutsamkeit. Wiederholte Ereignisse oder solche, die lange zurückliegen, unterliegen zahlreichen möglichen Einflüssen, die zu Fehlern in der Aussage führen können. Durch das mehrfache Erleben eines Ereignisses können sich Wissensstrukturen zum Ablauf und Elementen im Gedächtnis als Schemata oder Skripte ausbilden, so dass spezifische Ereignisdetails nicht mehr einem Einzelereignis zugeordnet werden können (Erdfelder & Bredenkamp, 1998). Im Vernehmungsalltag kann es unter Umständen auch zu Überlappungen mit aktuellen Geschehnissen, suggestiver Beeinflussung oder auch Überlagerungen kommen (Greuel, 2001; Volbert, 2005).

2.2.1 Schemawissen

Schemata sind organisierte Informationseinheiten des Erfahrungswissens über Erlebnisse oder Begebenheiten, welche unter anderem durch wiederholtes Erleben entstehen (Brewer & Treyens, 1981). Schemata sind im Langzeitgedächtnis verortet und sowohl mit dem episodischen als auch mit dem semantischen Gedächtnis verknüpft, da zum einen deklarative Begriffe zur Benennung von Situationselementen notwendig sind und zugleich spezifische Situationen gespeichert werden (Mantwill, Köhnken & Aschermann, 1995). Sowohl verbalisierbares, deklaratives Wissen als auch prozedurale Abläufe können zu thematisch organisierten Einheiten zusammengefasst werden, welche die Informationsverarbeitung erleichtern (Büttner, 2003). Das Vorwissen, welches im Rahmen der Bildung von Schemata genutzt wird, kann die Erwartungen an eine Situation sowie die Wahrnehmung, Enkodierung und Speicherung situativer Elemente oder Details beeinflussen. So wird zum Beispiel die Wahrnehmung situationsspezifischer Details reduziert, wenn eine schematisierte Situation auftritt (Brewer & Treyens, 1981).

Nelson (1986) hat argumentiert, dass Erinnerungen dazu dienen, Vorhersagen über die Zukunft zu erleichtern und dass wiederholte Erfahrungen in der Regel bessere Vorhersagen erlauben. Weiterhin sind Erinnerungen an Elemente die wiederholt auftreten weniger anfällig für Suggestionen und Falschinformationen (Connolly & Lindsay, 2001). Im Rahmen von Zeugenaussagen kann es zu massiven Einschränkungen in der Aussagegenauigkeit kommen, wenn es sich um ein Erlebnis handelt, welches in Form schematischer Wissensstrukturen abgespeichert ist (Steffens & Mecklenbräuer, 2007). Fehlen bestimmte Detailerinnerungen, werden diese möglicherweise automatisch aus dem vorhandenen Schema ergänzt, was als Konfabulation bezeichnet wird. Erinnerungen, die in Skripten abgespeichert sind, werden oftmals nur gerüstartig wiedergegeben, ohne dass Details, welche sich für jedes individuelle Ereignis unterscheiden, wiedergegeben werden können, was in forensischen Kontexten jedoch von besonderer Bedeutung sein kann (Volbert & Dahle, 2010).

2.2.2 Pseudoerinnerungen und Konfabulationen

Bei Pseudoerinnerungen oder false memories handelt es sich um subjektiv für wahr gehaltene Erinnerungen an ein Ereignis, welches in der Realität nicht stattgefunden hat (Steffens & Mecklenbräuer, 2007). Hiervon sind bewusste Falschaussagen klar abzugrenzen. Grundsätzlich können Pseudoerinnerungen durch verschiedene Formen der Suggestion hervorgerufen werden. Falschinformationseffekte zählen ebenfalls zu den Suggestionen, sind jedoch durch eine dem Ereignis nachgelagerte Beeinflussung der Ereigniserinnerung durch neue, hinzugefügte Informationen gekennzeichnet. Falsche Erinnerungen werden laut Schacter, Norman und Koutstaal (1998) in Intrusionen und Konfabulationen sowie falsches Wiedererkennen unterteilt. Dabei handelt es sich bei Intrusionen um neue, erfundene Abschnitte einer erlebten Episode, bei Konfabulationen um Narrationen über Dinge, die nicht stattgefunden haben. Falsches Wiedererkennen hingegen ist die fehlerhafte Identifikation von Elementen als bereits bekannt. Johnson und Raye (1998) führen Konfabulationen auf Quellenidentifikationsfehler zurück, was Heidler (2010) zufolge nur bedingt zutrifft, da keine sog. „Fantastischen Konfabulationen“ erklärt werden können. Principe und Schindewolf (2012) zufolge können auch „natürliche“ Gespräche zwischen Kindern und ihren Eltern oder Gleichaltrigen zu der Bildung falscher Erinnerungen führen. Dabei kann nicht nur der Inhalt von Gesprächen über die Vergangenheit in das spätere Gedächtnis

eindringen, sondern ein solcher Austausch kann auch die Generierung von völlig falschen Erzählungen bewirken, die detaillierter sind als wahre Berichte von erlebnisbasierten Ereignissen.

Diesen Effekten fehlerhafter Erinnerungen kann durch eine standardisierte und gesicherte Befragungspraxis entgegengewirkt werden. In Kapitel 3.4 „Aktuelle Interviewtechniken und -verfahren“ werden entsprechende Interviewverfahren aufgeführt, welche auf der Grundlage empirischer Forschungsergebnisse Hinweise zur suggestionfreien Befragung liefern.

2.3 Kinder als Zeugen

Hinsichtlich der Vernehmung von kindlichen Augenzeugen sind zahlreiche strukturelle, kontextuelle und individuelle Voraussetzungen sowie Bedingungen der Befragung zu beachten. Die Funktionsweise des menschlichen Gedächtnisses entwickelt und verändert sich über die Lebensspanne und beeinflusst dementsprechend die Aussagefähigkeit. Obgleich die Beziehung zwischen Alter und Gedächtnis komplex ist, können Kinder durchaus valide Informationen als Zeugen bereitstellen. Dabei ist festzuhalten, dass Alter nur als ein Indikator, nicht aber als Prädiktor für die Gedächtnisleistung angesehen werden darf (Roberts & Blades, 2000; siehe Kapitel 1.2 „Gedächtnisentwicklung und Source Monitoring bei Kindern“).

Die Qualität bzw. Quantität der Erinnerungsleistung wird von einer Vielzahl an Faktoren beeinflusst, welche im Befragungskontext Beachtung finden müssen (Lamb, Orbach, Hershkowitz, Esplin & Horowitz, 2007). Kinder sind in besonderem Maße abhängig von den Befragungsbedingungen (Roebbers, 2010) und beispielsweise anfälliger für Suggestionen (Volbert & Dahle, 2010). Dementsprechend sollten Wahl- oder Vorhaltfragen vermieden werden. Im Rahmen der Forschung zur Durchführung forensischer und investigativer Interviews wurden zahlreiche Standards entwickelt, die Praktikern helfen, Kinder zu befragen, wie das Protokoll des National Institute of Child Health and Disease (NICHD; Lamb et al., 2007), welches seit 2016 auch in einer deutschen Übersetzung der Westfälischen Wilhelms-Universität Münster vorliegt. Neben halbstandardisierten Protokollen stehen Listen der Best Practice zur Verfügung (siehe Home Office, 1992, 2002; Litchford, 2017 zit. n. APSAC Taskforce, 2017; Lamb, 1994; Lamb, Sternberg & Esplin, 1998; Orbach, Hershkowitz, Lamb, Sternberg, Esplin & Horowitz, 2000; Poole & Lamb, 1998; Warren & McGough, 1996).

Kinder haben im Vergleich zu Erwachsenen eine geringere Gedächtniskapazität und eingeschränkte Regulations- und Monitoringfähigkeiten (Balloff, 2004; Büttner, 2003; Nelson & Narens, 1994). Nicht zuletzt aufgrund der Ergebnisse der anfänglichen Aussageforschung wurde zu Beginn des 20. Jahrhunderts von einer nicht vorhandenen Aussagetüchtigkeit bei Kindern ausgegangen (Regber, 2007). Erst als in den 70er Jahren vermehrt Fälle von sexuellem Kindesmissbrauch an die Öffentlichkeit gelangen, gewannen kindliche Zeugenaussagen an Bedeutung (Goodman, 1984). Im Zuge der Ermittlungen an einer kalifornischen Vorschule trat das Phänomen der suggestiven Beeinflussung in den Fokus: Durch wiederholte suggestive Befragungen wurden Angaben zu vermeintlichen Missbrauchsfällen gemacht, welche jedoch keine Erlebnisbasis hatten (Garven, Wood & Malpass, 2000). Aus dem deutschen Sprachraum können in diesem Zusammenhang die sog. Wormser Prozesse aufgeführt werden, welche weitreichende Konsequenzen für die Beteiligten, die Justiz und die psychologische Gutachtertätigkeit hatten, u.a. ein Urteil des Bundesgerichtshof (BGH; Schade & Harschneck, 2000). Dies führte zu einer Anpassung der Standards in der Vorgehensweise aussagepsychologischer Begutachtungen. In Folge dieser Ereignisse wurden Umstände und Techniken von Interviews mit Zeugen im Rahmen der Forschung näher betrachtet, um Falschaussagen von Kindern zu vermeiden (Ceci & Bruck, 1993). Im Rahmen der deutschen Strafprozessordnung (StPO) ist dies von besonderer Relevanz, da keine Altersbeschränkung für Zeugenaussagen vorgesehen ist (Volbert & Dahle, 2010).

Kinder sind aus entwicklungs- und gedächtnispsychologischer Hinsicht (siehe Kapitel 1.2 „Gedächtnisentwicklung und Source Monitoring bei Kindern“) aussagetüchtig. Die Aussagetüchtigkeit ist jedoch in besonderem Maße von der Qualität der Erinnerung, den verwendeten Abrufmechanismen und der Interaktion mit dem Interviewer beeinflusst (Pipe, Lamb, Orbach & Esplin, 2004). Kinder sind ab etwa drei Jahren in der Lage, selbstständig kohärente Darstellungen eines zurückliegenden Ereignisses abzugeben. Dabei benötigen sie viele Hinweisreize und ihre Angaben sind fehlerbehaftet. Bis ins Schulalter verbessern sich diese Fähigkeiten stetig (Nelson & Fivush, 2004). Hinsichtlich der ersten drei Lebensjahre wird von frühkindlicher Amnesie gesprochen, Ereignisse ab dem vierten Lebensjahr können sodann über einen Zeitraum von mehreren Monaten oder Jahren erinnert werden und auch raum-zeitlich eingeord-

net werden. Ab dem Grundschulalter sind Kinder in der Lage, Hintergrundinformationen in ihre Erfahrungen zu integrieren sowie zeitliche und kausale Verknüpfungen herzustellen (Haden, Haine & Fivush, 1997). Es besteht Konsens darüber, dass Kinder ab einem Alter von drei oder vier Jahren als Zeugen befragt werden können, dies geht auch einher mit der *Theory of Mind* (ToM; (siehe Exkurs: Theory of Mind; Cohen, Leslie & Frith, 1985; Leslie, 2000; Premack & Woodruff, 1978), welche die Fähigkeit zur Perspektivübernahme und Wiedergabe eigener Erlebnisse in narrativer Form umschreibt (Nelson, 2000; Volbert & Dahle, 2010). Die Aufgaben zur Überprüfung der ToM, ähneln jenen zur Überprüfung des Source Monitorings, welche in der vorliegenden Arbeit benutzt wurden. In früheren Studien konnte festgestellt werden, dass die Ergebnisse von Aufgaben der ToM und des Source Monitorings miteinander korrelieren (Leichtman, Morse, Dixon & Spiegel, 2000).

Etwa zeitgleich mit der Fähigkeit zur Perspektivübernahme im dritten bis vierten Lebensjahr entwickeln sich die Fähigkeiten zur Täuschung: Kinder verstehen bis dahin aber noch nicht, dass sie neben der verbal geäußerten Lüge, auch ihr Verhalten an die falsche Aussage anpassen müssen, um einen kohärenten Eindruck zu bewirken. Kinder betreiben noch kein Impression-Management (Volbert & Dahle, 2010), das möglichst glaubwürdige Auftreten und Vermeiden von kompromittierenden Verhaltensweisen, so dass ihre Täuschungsversuche oftmals noch leicht zu enttarnen sind (Talwar & Lee, 2008). Im Fokus der meisten Untersuchungen mit Kindern im aussagepsychologischen Kontext stehen Einflüsse durch Suggestion, was definiert ist als „degree to which children’s encoding, storage, retrieval, and reporting of events can be influenced by a range of social and psychological factors“ (Ceci & Bruck, 1993, S. 404). Als Formen der Suggestion werden falsche Informationen, leitende oder forced-choice Fragen, Druck oder Erwartungen des Interviewers unterschieden (Krähenbühl, Blades & Eiser, 2009).

Exkurs: Theory of Mind

Unter Theory of Mind wird aus entwicklungspsychologischer Perspektive die modellhafte Metarepräsentation von bzw. ein Bewusstsein für mentale Zustände verstanden, welche soziale Fähigkeiten und Interaktion ermöglichen (Premack & Woodruff, 1978). Im Detail geht es um das Verstehen von mentalen Zuständen, also zum Beispiel zu wissen was andere Personen wissen, wollen, fühlen oder glauben (Cohen et al., 1985; Leslie, 2000). Die Bedeutung dieser Fähigkeiten wird insbesondere im Kontext der Zeugenbefragung deutlich, da die soziale Situation eines Interviews bzw. einer Zeugenbefragung eben diese Anforderungen an den Zeugen stellt. Suggestibilität ist Roebers und Schneider (2005b) zufolge mit der Theory of Mind verbunden, da sich Kinder in Interviews darüber bewusst sein müssen, dass sie Wissen haben, welches das Gegenüber nicht kennt und dieses Bewusstsein bei der Beantwortung von Fragen berücksichtigen müssen. Scullin und Bonner (2006) haben den Zusammenhang zwischen Suggestibilität, inhibitorischer Kontrolle und Theory of Mind untersucht und inkonsistente Ergebnisse gefunden. Sie stellten fest, dass die Fähigkeit zur Perspektivübernahme bzw. Theory of Mind und die inhibitorische Kontrolle die Suggestibilität von drei- bis fünfjährigen Kindern voraussagen kann. Theory of Mind war in allen Fällen mit der Aussagegenauigkeit bzw. Fehlern bei unbeantwortbaren oder irreführenden Fragen verknüpft, jedoch konnten keine signifikanten Zusammenhänge zu anderen Messungen der Suggestibilität gefunden werden.

Die Forschungsfragen zu Zeugenaussagen von Kindern wurden für die vorliegende Arbeit den zwei übergeordneten Themenbereichen, personale und situative Merkmale, zugeordnet (siehe Tabelle 4). Die Literaturrecherche erfolgte anhand der Pearl-Finding-Growing-Methode (Ramer, 2005) über rechtspsychologische Studien mit dem Fokus der Zeugenbefragung von Kindern (Schlosser, Wendt, Bhavnani & Nail-Chiwetalu, 2006). Zu den personalen Merkmalen zählen 1. Eigenschaften des kindlichen Zeugen, wie Persönlichkeit und Kognition, sowie 2. Eigenschaften des Interviewers, wie beispielsweise Erwartungseffekte. Zu den situativen Merkmalen zählen zum Beispiel 3. Ereignismerkmale, wie der Grad des Involvements bzw. die Tatsache, ob man Augen-

oder Opferzeuge ist, der Umstand, ob es sich um ein einmaliges oder wiederholtes Ereignis handelt, und Merkmale bzw. Verhaltensweisen des Täters. Ein weiterer Forschungsbereich, der sich den situativen Merkmalen zuordnen lässt, ist 4. das Intervall zwischen Lernphase und Befragung, wobei Faktoren relevant sind, welche sich auf die reine Länge beziehen, aber auch (falsche) Post-Event Informationen. Ebenfalls hier zugeordnet wurden dem Ereignis nachgelagerte Gespräche über eben dieses mit den Eltern oder Peers sowie Quellenkonfusionen. Als letzte Kategorie ist 5. Die Befragungssituation aufgeführt, bei der situative Aspekte betrachtet werden. Dabei sind die zugehörigen Studien für die vorliegende Arbeit von besonderer Relevanz, da es sich um von außen beeinflussbare Umstände handelt.

2.3.1 Personale Aspekte

In Bezug auf die personalen Merkmale haben Bruck und Melnyk (2004) in einem Review demografische, kognitive und psychosoziale Einflussfaktoren auf Seiten der kindlichen Zeugen als relevante Eigenschaften hinsichtlich der interindividuellen Suggestibilität identifiziert. Kreativität erwies sich als relevanter Prädiktor mit positiver Vorhersageleistung für die Suggestibilität; semantisches Sprachverstehen hingegen korrelierte negativ mit Suggestibilität, keine Zusammenhänge konnten zur Intelligenz im Normalbereich, Erinnerungsleistung und Ablenkbarkeit gefunden werden (hier sei angemerkt, dass die Höhe der Korrelationen und die Effektstärken in der Metastudie von Bruck und Melnyk aus dem Jahr 2004 nicht angegeben sind).

Bezogen auf die Theory of Mind und exekutive Funktionen zeigen sich insgesamt inkonsistente Ergebnisse (Bruck & Melnyk, 2004), dabei konnte in vier von elf Untersuchungen kein Zusammenhang nachgewiesen werden (Principe, 1997, zit. n. Bruck & Melnyk, 2004; Crossman, 2001, zit. n. Bruck & Melnyk, 2004; Quas & Schaaf, 2002; Roebbers & Schneider, 2004, zit. n. Bruck & Melnyk, 2004).

Tabelle 4. Systematik der Forschungsliteratur zu kindlichen Zeugenaussagen.

PERSONALE MERKMALE	SITUATIVE MERKMALE
1. Eigenschaften des kindlichen Zeugen	3. Merkmale des Ereignisses
Persönlichkeit	Augenzeuge vs. Opferzeuge (bzw. Involvement)
Kognition	Wiederholtes vs. einmaliges Ereignis
Stress	Eigenschaften und Verhalten des Täters
Täuschungskompetenz	Erzählverbot
2. Eigenschaften des Interviewers	4. Intervall zwischen Ereignis und Befragung
Erwartungseffekte	Länge des Zeitintervalls
Kenntnis über den tatsächlichen Tathergang	Post-Event-Information
	Falsche Post-Event-Informationen
	Gespräche, z.B. mit den Eltern oder Peers
	Quellenkonfusion
	5. Befragungssituation
	Abruf und Regulation
	Soziale Aspekte
	Motivationale Aspekte

Die positiven Zusammenhänge, welche bei Templeton und Wilcox (2000) gefunden wurden, führen die Autoren des Reviews darauf zurück, dass die Stichprobe aus vergleichsweise jungen drei- bis sechsjährigen Versuchspersonen bestand, wobei Kinder im Alter von drei und vier Jahren grundsätzlich Schwierigkeiten bei der Bearbeitung von Theory of Mind Aufgaben haben und bei genauerer Betrachtung die sechsjährigen Kinder im Vergleich zu den anderen Altersgruppen am meisten suggestibel waren (Bruck & Melnyk, 2004). Insgesamt konnten signifikante Zusammenhänge zwischen

Missattributionen und Theory of Mind gefunden werden (Bright-Paul, 2004, zit. n. Bruck & Melnyk, 2004; Welch-Ross, 1999, zit. n. Bruck & Melnyk, 2004).

Hinsichtlich der Täuschungskompetenz können aus entwicklungspsychologischer Perspektive laut Robinson (1996) drei Phasen unterschieden werden, welche bei der Durchführung und Bewertung von Zeugenbefragungen mit Kindern beachtet werden sollten. Zunächst nutzen sie die schlichte Täuschung, das ledigliche Abstreiten, mit dem Erlernen entsprechender Kontingenzregeln sind sie in der Lage, Emotionen vorzutäuschen, bis später das komplexe Zusammenspiel verbal-sprachlicher und emotionaler Täuschung erlernt wird. Die Täuschungskompetenz zählt zu den personalen Merkmalen des Zeugen, findet aber vornehmlich Anwendung während der Befragung.

Die Eigenschaften des Interviewers untersuchten die Autoren Bederian-Gardner und Goldfarb (2014) vor dem Hintergrund der *Expectancy Violation Theory*. Dabei handelt es sich um eine Kommunikationstheorie, welche analysiert, wie Personen auf Grundlage ihrer Erwartung auf non-konformes, nicht erwartetes Verhalten des Gegenübers reagieren und dieses bewerten. In der Untersuchung sollten die erwarteten Emotionen kindlicher Opferzeugen hinsichtlich ihrer Intensität eingeschätzt werden, wobei Alter und Geschlecht der zu bewertenden Zeugen variiert wurden. Dabei konnte die empfundene Empathie der Erwachsenen ($N = 191$) für kindliche Opfer vorhersagen, wie intensiv die Einschätzung der gezeigten Emotionen von Trauer und Angst war. Des Weiteren konnte ein Alterseffekt nachgewiesen werden: so wurden von älteren Kindern (13 Jahre) stärkere Emotionsäußerungen erwartet, als von jüngeren (fünf Jahre).

2.3.2 Situative Aspekte

In diesem Abschnitt werden Studien dargestellt, welche sich mit beeinflussbaren Merkmalen auf Seiten der Befragungssituation im forensischen Kontext befassen. Dabei werden primär rechtspsychologische Untersuchungen mit Kindern zwischen drei und elf Jahren berichtet. Es stehen Prozesse im Fokus, die mögliche Auswirkungen auf die Aussagegenauigkeit haben.

Aussagepsychologische Studien zu situativen Merkmalen beziehen sich unter anderem auf Eigenschaften des Ereignisses bzw. des Lernmaterials, beispielsweise ob es sich um ein neues oder wiederholtes Ereignis handelt. Connolly und Lindsay (2001) konnten bereits bei vier- bis sechsjährigen Kindern Bestätigung für die Skripttheorie in Verbindung mit einer erhöhten Resistenz gegen suggestive Einflüsse bei wiederholten

Ereignissen nachweisen. Das Intervall zwischen Ereignis und Befragung stellt ein besonders relevantes situatives Merkmal dar, welches in der vorliegenden Arbeit insbesondere durch Quellenidentifikationsfehler Beachtung findet, so stehen diese in direktem Zusammenhang mit der Latenz zwischen Ereignis und Abruf bzw. Befragung. Im Folgenden werden in Anlehnung an das experimentelle Vorgehen dieser Arbeit, ausschließlich die veränderbaren Merkmale der Befragung, wie Abrufbedingungen und Effekte der Fragenformulierung (siehe Tabelle 4: 5. Befragungssituation), besondere Beachtung finden.

Im Review von Bruck und Melnyk (2004) werden in den Studien die Zielereignisse der Lernphase in Abhängigkeit vom Ausmaß der Interaktion unterschieden, dabei kann es sich um verbales, bildhaftes oder videografisches Material bis hin zu simulierten medizinischen Untersuchungen handeln. Der Grad der Partizipation kann sich auf die Erinnerungsleistung auswirken (Tobey & Goodman, 1992). Die Studien lassen sich wiederum Subkategorien zuordnen: Dabei handelt es sich um Abruf aus dem Langzeitgedächtnis, Regulation, sowie soziale und motivationale Aspekte der Befragungssituation. Im Folgenden werden darüber hinaus fokussiert Studien vorgestellt, welche sich mit den kognitiven Aspekten des Abrufs und der Regulation befassen, die sozialen Aspekte werden nur exemplarisch erläutert. Die motivationalen Aspekte werden in Kapitel 4 „Zuverlässigkeitsmotivation“ dargelegt, da es sich dabei um einen der experimentell variierten Faktoren dieser Forschungsarbeit handelt.

Die nachfolgenden Studien (Cole & Loftus, 1987; Garven et al., 2000; Howie, Sheehan, Mojarrad & Wrzesinska, 2004; Koriat, Goldsmith, Schneider & Nakash-Dura, 2001; Krähenbühl et al., 2009; Memon & Vartoukian, 1996; Nesbitt & Markham, 1999; Ricci, Beal & Dekle, 1996) beschäftigen sich mit Gedächtnisprozessen und metakognitiven Aspekten, die für den Abruf bzw. das Monitoring von Erinnerungen relevant sind. Des Weiteren wird untersucht, wie durch Manipulation eben dieser Aspekte Fehler vermieden und eine Verbesserung der Aussagegenauigkeit bewirkt werden kann. Wie in Kapitel 1.1.2 „Source Monitoring“ werden die Ergebnisse unter Berücksichtigung der Effektstärken d und η^2 berichtet, wobei diese wiederum selbst berechnet wurden, soweit nicht in den Studien angegeben.

In zahlreichen Untersuchungen konnte festgestellt werden, dass sich Zeugenaussagen von Kindern vorwiegend hinsichtlich der Quantität von den Aussagen Erwachsener unterscheiden, nicht aber in der Qualität. Kinder äußern sich weniger häufig spontan als Erwachsene zu einem Ereignis. Dies führt unter anderem dazu, dass kindlichen Zeugen vermehrt nachhakende oder wiederholte Fragen gestellt werden (Cole & Loftus, 1987).

Memon und Vartoukian (1996) wiesen nach, dass wiederholte Befragungen negative Auswirkungen auf die Aussagegenauigkeit von Kindern haben, wenn es sich um geschlossene Frageformate handelt. Zur Überprüfung der Fragestellung wurden $N = 44$ Kinder zwischen fünf und acht Jahren wiederholt zu einem inszenierten Ereignis befragt, wobei offene und geschlossene Fragen zu zentralen und peripheren Details der Geschichte gestellt wurden. Von insgesamt 30 Fragen wurden in derselben Befragung zehn Fragen wiederholt gestellt. Vor einer systematischen Befragung sollten die Kinder zunächst im freien Bericht (siehe Kapitel 3.4.2 „Der freie Bericht“) ihre Erinnerungen darlegen. Des Weiteren wurde eine Teilstichprobe darüber informiert bzw. gewarnt, dass wiederholte Fragen vorkommen können. Neben einem signifikanten Alterseffekt, ältere Kinder erinnern mehr und genauer periphere Details, wurde ein Haupteffekt für den Informationstyp, zentral versus peripher, bei geschlossenen Fragen beobachtet (aufgrund fehlender Angaben zu den Standardabweichungen, konnten keine Effektstärken berechnet werden).

Howie et al. (2004) untersuchten den Einfluss wiederholter Befragungen auf den Response Shift. Dabei wurden $N = 89$ Kinder zwischen vier und acht Jahren wiederholt von denselben und auch unterschiedlichen Interviewern zu einem Ereignis befragt. Die Veränderung der Aussage wurde als Maß des Response Shift erhoben. Als relevanter Faktor für einen möglichen negativen Effekt der Inkonsistenz wurde von den Autoren eine Aufklärung über den Sinn und Zweck einer wiederholten Befragung identifiziert. Die jüngeren Kinder im Alter von vier und fünf Jahren nahmen trotz Erhalt einer Aufklärung bzw. Begründung für eine wiederholte Befragung mehr Änderungen an ihren Aussagen vor ($d = -0,522$), wenn sie wiederholt von derselben Person mit Fragen ohne Bias befragt wurden. Für die älteren Kinder zwischen sieben und acht Jahren zeigte sich ein zu vernachlässigender Effekt von $d = 0,063$, was darauf schließen lässt, dass der Hinweis auf bzw. eine Erklärung für wiederholte Befragungen dazu führt, dass keine bis kaum Änderungen in den Aussagen stattfinden, diese Art der Aufklärung aber

nicht bei Kindern zwischen vier und fünf Jahren funktioniert. Des Weiteren zeigte sich, dass eine Aufklärung bei der wiederholten Befragung durch unterschiedliche Interviewer zu einem gegenteiligen Effekt führt und junge als auch ältere Kinder tendenziell mehr Shifts zeigten.

Krähenbühl et al. (2009) konnten die negativen Auswirkungen wiederholter Fragen auf die Aussagegenauigkeit mit variierenden Abständen zwischen den Wiederholungen und Füllfragen nachweisen. Von besonderer Bedeutung sind die Ergebnisse der Autoren für unbeantwortbare Fragen, so stieg mit der Anzahl an Wiederholungen dieses Fragentyps die Wahrscheinlichkeit, dass die Kinder zwischen fünf und neun Jahren ($N = 156$) eine Antwort ausdachten ($d = -0,185$; kleiner Effekt) für prozentuale Angabe der Veränderungen.

Bezogen auf die Prozesse des Abrufs und der Regulation haben Nesbitt und Markham (1999) an $N = 64$ drei- bis sechsjährigen Kindern untersucht, wie sich ein Training bzw. eine Instruktion zum Quellen-Monitoring auf die Leistung in einer Source Monitoring Aufgabe auswirkt. Dazu wurde den Kindern ein Video gezeigt und eine Geschichte erzählt, welche zum Teil identische oder auch nur ähnliche Elemente enthielten. Eine Hälfte der Kinder wurde sowohl im Einzel- als auch im Gruppensetting dahingehend trainiert, dass sie „Ich weiß nicht“ sagen sollen, wenn sie sich bei einer Frage unsicher waren oder eine Antwort nicht wussten. Dabei wurde auf die negativen Folgen falscher Antworten im polizeilichen Kontext hingewiesen. Die trainierten und untrainierten Kinder wurden im Anschluss mit zwei unterschiedlichen Instruktionen zu den Inhalten des Videos und der Geschichte befragt. Je die Hälfte trainierter und untrainierter Kinder wurde zu den Inhalten befragt, ohne den Hinweis zu bekommen, dass auch irreführende Fragen zu nicht enthaltenen Elementen vorkommen. Die andere Hälfte erhielt in der Testphase eine Instruktion zum Quellen-Monitoring, wobei sie darüber aufgeklärt wurden, dass nur manche Teile der Geschichte auch im Video vorkamen und sie bei der Zuordnung der Elemente aufpassen sollten. Der Anteil richtiger Antworten in der Source Monitoring Aufgabe war bei den individuell trainierten Kindern höher ($d = 0,476$) als bei denen ohne Training. Die untrainierten Kinder gaben hingegen signifikant mehr richtige Antworten (*Hits*; siehe Kapitel 1.1.2 „Source Monitoring“) ($d = 0,459$) als die trainierten Kinder, was jedoch an der höheren Anzahl an „Ich weiß nicht“ Antworten der trainierten Kinder lag ($d = 1,778$; sehr große Effektstärke). In Anlehnung an Lange, Guttentag und Nida (1990) schlussfolgerten die Autoren, dass die Entwicklung

kognitiver Fähigkeiten durch explizites Training beschleunigt werden kann und die Fähigkeit zur Unterscheidung von Wissen und Raten, im Sinne des Remember/ Know-Paradigmas, sowie adäquate Monitoringkompetenzen sich zwischen dem dritten und fünften Lebensjahr entwickeln (Johnson & Wellman, 1980; Gopnik & Graf, 1988).

Koriat et al. (2001) nehmen entwicklungspsychologische Defizite beim Monitoring als mögliche Erklärung für die geringere Aussagegenauigkeit von Kindern an. Dabei wird die Metakognition als Mediator verstanden, welcher die strategische Kontrolle der Gedächtnisleistung ermöglicht. Belege für das zugrundeliegende Modell konnte bereits bei einer Stichprobe Erwachsener nachgewiesen werden (Koriat & Goldsmith, 1996; Goldsmith & Koriat, 1999; siehe Kapitel 4.1 „Strategische Regulation der Aussagegenauigkeit“). In einer experimentellen Untersuchung haben die Autoren den freien Bericht und die Rekognitionsleistung erhoben, dabei wurden free- und forced-Report Optionen genutzt. Den Teilnehmern ($N = 160$) im Alter von sieben bis zwölf Jahren wurde entweder freigestellt, eine Antwort zu geben, oder sie wurden dazu forciert, dabei wurden sie für richtige Antworten belohnt und für falsche bestraft. Die Erinnerungsgenauigkeit (accuracy; qualitativ) war bei den jüngeren Kindern für den Recall in der free Report-Bedingung besser als in der forced Report-Bedingung ($d = -1.491$), der Effekt für die Rekognition war hingegen deutlich kleiner ($d = -.886$). Auch die älteren Kinder zeigten genauere Erinnerungsleistungen in der free als in der forced Report-Bedingung ($d = -0,771$), hier erwies sich der Effekt für die Rekognition als größer ($d = -1,53$). Für die Vergleiche muss angemerkt werden, dass die Werte für Quantität und Qualität in den forced-Bedingungen äquivalent sind und demnach die Darstellung verzerren können. In einem zweiten Experiment ($N = 80$) wurde die Wirkung hoher Belohnung untersucht und ob die positiven Effekte auf die Erinnerungsleistung, welche bei Erwachsenen gefunden werden konnten, auch bei Kindern unterschiedlichen Alters auftreten. Für die jüngeren Teilnehmer zeigte sich für den Recall in Abhängigkeit der Stärke der Belohnung ein mittlerer Effekt ($d = 0,427$) auf die Erinnerungsgenauigkeit, bei den älteren war dieser im Vergleich sehr hoch ($d = 1,196$), dementsprechend war die Wirkung der Belohnung auf die Qualität der Erinnerung hier ebenfalls hoch.

Ricci et al. (1996) haben den Bekanntheitsgrad zwischen Interviewer und Befragtem manipuliert und untersucht, wie sich die Aussagegenauigkeit bei der Befragung von $N = 80$ Kindern ($M = 5,75$ Jahre) durch einen Interviewer oder die eigenen Eltern unterscheidet. Zusätzlich wurde ein Line-Up zur Täteridentifikation durchgeführt, wobei die

Anwesenheit des Täters manipuliert wurde. Es zeigte sich, dass die Kinder bei einer Befragung durch ihre Eltern signifikant schlechtere Ergebnisse erzielten, als bei der Verwendung von strukturierten Leitfäden. Bei der Line-Up Identifikation ergaben sich keine Unterschiede, so dass davon ausgegangen werden kann, dass vergleichbare Gedächtnisleistungen vorlagen. Die Berechnung von Effektstärken war aufgrund fehlender Angaben zur Standardabweichung nicht möglich. Es ergab sich weiterhin, dass die Kinder bei der Befragung durch ein vertrautes Elternteil entspannter waren und dies positiv mit der Aussageleistung korrelierte (nach drei Minuten: $r = .79$, $p < .01$). Dabei wurde die Entspannung von Beobachtern, welche nur die Kinder, nicht aber die Interviewer sehen konnten, im Nachgang per Video eingeschätzt.

Garven et al. (2000) haben in Anlehnung an einen Rechtsfall aus den 80er Jahren an der McMartin Vorschule in den USA, bei dem Lehrer fälschlicherweise des Missbrauchs bezichtigt wurden, den Effekt von normativem und informativem sozialen Einfluss auf die Aussagen von Kindern ($N = 93$) untersucht. Dabei manipulierten sie bestätigendes bzw. bestrafendes Verhalten des Interviewers oder Informationen von Co-Zeugen. Die Autoren gingen davon aus, dass im Sinne klassischer Konditionierung, Lob und Anerkennung als primäre Verstärker die Aussage ebenso beeinflussen, wie Bestrafung, Kritik oder Widerspruch. Informationen von Co-Zeugen wurden von den Autoren als Konformitätsdruck operationalisiert, indem ein Interviewer die Befragten über die Aussagen anderer Zeugen aufklärte und somit möglicherweise unter Druck setzte. Die Wirkung von Co-Zeugen-Informationen und der Verstärkung von Ja-Antworten lag bei $\eta^2 = 0,127$, der Haupteffekt der Co-Witness Informationen war ebenfalls signifikant, aber im Vergleich bedeutend kleiner $\eta^2 = 0,033$. Für falsche Informationen, welche verstärkt wurden, lag der Effekt bei $\eta^2 = 0,331$, dabei haben die Kinder in 52,71 % der Fälle fehlleitende Fragen mit Ja beantwortet, im Vergleich zu 6,11 % ohne Verstärkung.

Wie zu Beginn dieses Kapitels erläutert und im Rahmen der Ergebnisse ausgewählter Studien deutlich geworden ist, zeigen sich die Bedenken bei Zeugenaussagen von Kindern nicht hinsichtlich der Aussagemenge, sondern eher bezüglich der Genauigkeit und der Zuverlässigkeit (Koriat et al., 2001). Wie die Zuverlässigkeit beeinflusst werden kann, wird in Kapitel 4 „Zuverlässigkeitsmotivation“ vorgestellt.

3 Aussagepsychologische Begutachtung und Befragungstechniken

Das Ziel einer Befragung oder Vernehmung von Zeugen ist die möglichst präzise und fehlerfreie Erhebung von Informationen zu einem Sachverhalt (Aschermann, Mantwill & Köhnken, 1991; Köhnken, Mantwill & Aschermann, 1991; Krix & Sauerland, 2013; Loohs, 1996; Mantwill et al., 1995). Um bestmögliche Ergebnisse zu erzielen, sollten bereits bei der Aussageentstehung Bedingungen und Voraussetzungen bestehen und geschaffen werden, welche dies begünstigen.

3.1 Glaubhaftigkeitsbegutachtung

Die Statement Validity Analysis (SVA) steht synonym für die Trias der Glaubhaftigkeitsbegutachtung im deutschsprachigen Raum: Zur Bewertung der Frage, ob eine Zeugenaussage auf einem wahren Erlebnishintergrund beruht oder nicht, werden die Aspekte der Aussagetüchtigkeit und Aussagegenese überprüft sowie die kriterienorientierte Inhaltsanalyse der Aussage durchgeführt. Hinsichtlich der Aussagetüchtigkeit werden die notwendigen Voraussetzungen und Eigenschaften des Aussagenden zur Tätigkeit einer Aussage geprüft, dazu gehört die adäquate Wahrnehmung und Speicherung von Ereignissen sowie die Fähigkeit zum Abruf und zur Wiedergabe (Greuel, 2001; Volbert, 2005). Zur Überprüfung der Aussagegenese werden Kontext und Umstände der Erstaussage betrachtet sowie die Aussageentwicklung, hier wird insbesondere das Kriterium der Konstanz genauer untersucht. Mit Hilfe der kriterialen Inhaltsanalyse werden verschiedene kognitive, motivationale und deliktspezifische Eigenschaften der Aussage anhand eines Kriterienkatalogs überprüft. Im internationalen Raum wird von der Criteria based Content Analysis (CBCA) gesprochen, die im Folgenden als CBCA bezeichnet wird.

In Deutschland und auch den Niederlanden werden seit Jahrzehnten psychologische Sachverständige vor Gericht zur Klärung der Frage der Glaubhaftigkeit einer individuellen Zeugenaussage eingesetzt (Oberlader et al., 2016; Volbert & Steller, 2014). Als Folge fehlerhafter Gutachten, welche sich mit sexuellem Kindesmissbrauch beschäftigten und die öffentliche Aufmerksamkeit erregten (Steller, 2008), legte der BGH im Jahr 1999 Mindeststandards für die Begutachtung der Glaubhaftigkeit einer Zeugenaussage fest (Bundesgerichtshof in Strafsachen (BGHSt) 45, 164). Die Mindeststan-

dards orientieren sich an den Erkenntnissen aktueller Forschung, dabei werden Zeugenaussagen als höchstrichterliches Beweismittel betrachtet (Steller, 2008). Der Begutachtung von Zeugen liegt ein hypothesengeleiteter Prozess zugrunde, der von der SVA Gebrauch macht (Vrij, 2008). Basierend auf den verfügbaren Akten und Dokumenten wird ein halbstrukturiertes Interview geplant und mit dem Zeugen durchgeführt, dieses wird aufgenommen, transkribiert und anschließend mit Hilfe der CBCA-Kriterien ausgewertet (Vrij, 2008). Die inhaltliche Qualität der Zeugenaussage wird evaluiert und mit der Aussagekompetenz des Zeugen verglichen (Steller, 2008). Unter der Aussagekompetenz werden beispielsweise deliktsspezifische Erfahrungen oder kognitive Leistungsfähigkeiten verstanden, die eine glaubhafte Falschaussage erleichtern würden (Steller, 2008). Die inhaltliche Qualität einer Aussage wird durch die CBCA-Kriterien bestimmt (Vrij, 2008).

3.2 Kriterienorientierte Inhaltsanalyse

William Stern war Anfang des 20. Jahrhunderts einer der ersten Psychologen, der sich wissenschaftlich mit Unterschieden zwischen wahren und erfundenen Aussagen beschäftigte. Dabei ist ihm die grundlegende Differenzierung zwischen Aussagemenge und Aussagegenauigkeit zuzuschreiben (Neisser, 1982; Undeutsch, 1954; Undeutsch, 1989).

Die Entwicklung inhaltlicher Beurteilungskriterien zur Bewertung von Zeugenaussagen erfolgte vor allem durch die induktive Herleitung von Kriterien aufgrund von Aktenstudien und der Teilnahme von Psychologen an Gerichtsverhandlungen (Wegener & Steller, 1987). Undeutsch postulierte später die sog. Undeutsch-Hypothese, der zufolge Aussagen mit Erlebnisbasis sich in Qualität und Inhalt von einer erfundenen Aussage unterscheiden (Undeutsch, 1954). Basierend auf der Undeutsch-Hypothese entwickelten Steller und Köhnken (1989) die CBCA anhand der Analyse zahlreicher transkribierter Zeugenaussagen. Die CBCA Kriterien sind bei erlebten Aussagen häufiger vorzufinden als bei erfundenen Schilderungen, somit unterstützt ein häufiges Auftreten der Kriterien in einer Aussage die Alternativhypothese zu Gunsten einer Erlebnisbasis (Breuer, 2008). Die CBCA umfasst 19 unterschiedliche Kriterien (Steller & Köhnken, 1989), deren Bewertung aufgrund der Häufigkeit oder Stärke ihres Auftretens in einer Aussage erfolgt (Oberlader et al., 2016). Die Abwesenheit bzw. das nicht-Auftreten

eines Kriteriums schließt jedoch nicht automatisch aus, dass eine Erlebnisbasis vorliegt, vielmehr kann das Fehlen eines oder mehrerer Kriterien auch auf situationale, kognitive oder motivationale Defizite zurückzuführen sein (Oberlader et al., 2016; Volbert, Steller & Galow, 2010). Die einzelnen CBCA-Kriterien können in kognitive, motivationale und deliktspezifische Merkmale unterteilt werden (Steller & Köhnken, 1989). Theoretische und empirische Fundierung für die kognitiven und motivationalen Merkmale der CBCA finden sich im kognitiven Ansatz und dem Kontrollansatz. Die Aspekte des kognitiven Ansatzes sind in den kognitiven Kriterien repräsentiert. Aussagen ohne Erlebnisbasis sind logisch weniger konsistent, sie enthalten weniger Interaktionsschilderungen oder ungewöhnliche Details. Zudem nutzen lügende Personen bei der Konstruktion einer Aussage zumeist Schemawissen, hervorzuheben ist auch die kognitive Belastung, um eine logisch-stringente und widerspruchslose Geschichte zu produzieren (Niehaus, 2008; Volbert et al., 2010). Der Kontrollansatz beinhaltet die Annahme, dass falschaussagende Personen bemüht sind, einen glaubwürdigen Eindruck zu vermitteln und dabei auf Alltagsvorstellungen zurückgreifen, wie ein glaubwürdiger Zeuge aussagen würde (Niehaus, 2008; Volbert et al., 2010). Diese Annahmen spiegeln sich in den motivationalen Merkmalen der CBCA Kriterien wieder, so enthalten falsche Aussagen weniger Selbstkorrekturen, Selbstbelastungen oder Eingeständnisse von Erinnerungslücken. Im Folgenden werden die Kategorien und Unterteilungen erläutert, die einzelnen Kriterien können in Tabelle 5 eingesehen werden, für eine detaillierte Erläuterung und Beispiele siehe Steller und Köhnken (1989).

Den kognitiven Merkmalen liegt die Annahme zugrunde, dass sich die Aussagen mit Erlebnisbasis hinsichtlich ihrer Komplexität von erfundenen Aussagen unterscheiden (Niehaus, 2008). Es werden drei Subkategorien postuliert, allgemeine Merkmale, spezielle Inhalte sowie inhaltliche Besonderheiten (Steller & Köhnken, 1989). Zu den allgemeinen Merkmalen einer Aussage zählen die logische Konsistenz, der Detailreichtum sowie die unstrukturierte Darstellung (Steller & Köhnken, 1989). Die *logische Konsistenz* einer Aussage wird als gegeben angesehen, wenn keine inneren oder äußeren Widersprüche vorhanden sind (Niehaus, 2008), die Aussage an sich logisch ist und die einzelnen Aussageelemente zueinander passen (Vrij, Kneller & Mann, 2000). *Detailreichtum* ist gegeben, wenn die Aussage viele Details hinsichtlich der Beschreibung von (Tat-)Orten, Personen, Ereignissen, Handlungsabläufen oder Objekten be-

sitzt (Steller & Köhnken, 1989; Vrij, Kneller & Mann, 2000). Unter *unstrukturierter Darstellung* wird eine Schilderung des Handlungsablaufs verstanden, die nicht zwangsläufig chronologisch und kohärent erfolgt (Vrij et al., 2000), was der episodischen und hoch individuell gefärbten Erinnerung entspricht. Nichtsdestotrotz müssen die einzelnen Teile der Aussage logisch zusammenpassen und sich in die Gesamterzählung einfinden (Steller & Köhnken, 1989). Dieses Kriterium tritt vor allem auf, wenn der Zeuge das Geschehnis zuvor selten erzählt hat (Vrij, 2008). Spezielle Inhalte beschreiben die kontextuelle Einbettung, Interaktionsschilderungen, Wiedergabe von Gesprächen und die Schilderung von unerwarteten Handlungskomplikationen (Steller & Köhnken, 1989). Dabei wird bei der *kontextuellen Einbettung* ermittelt, inwieweit sich eine Handlung in den raum-zeitlichen Kontext einfügt und inwiefern das Tatgeschehen mit externen Faktoren verknüpft ist (Niehaus, 2008; Volbert et al., 2010). *Interaktionsbeschreibungen* liegen vor, wenn Handlungen zwischen dem Aussagenden und dem Beschuldigten beschrieben werden, die eine entsprechende Reaktion auf beiden Seiten ausgelöst haben (Vrij et al., 2000). Das Kriterium für die *Wiedergabe von Gesprächen* ist erfüllt, wenn neben der Zitation auch einzelne Sprecher identifizierbar werden (Steller & Köhnken, 1989; Vrij et al., 2000; Vrij, 2008). Eine *Schilderung unerwarteter Handlungskomplikationen* liegt vor, wenn intendierte Handlungen durch plötzliche Schwierigkeiten unterbrochen oder gestört werden (Niehaus, 2008; Vrij et al., 2000). Inhaltliche Besonderheiten umfassen Schilderungen ungewöhnlicher und nebensächlicher Details, unverständener Handlungselemente, eigener psychischer sowie fremdpsychischer Vorgänge sowie indirekt handlungsbezogene Schilderungen (Volbert et al., 2010). Ungewöhnliche Details sind ausgefallene Objekteigenschaften oder Merkmale von Personen oder Ereignissen (Niehaus, 2008; Vrij et al., 2000). *Nebensächliche Details* sind Äußerungen, die für den Tathergang nicht relevant sind (Vrij et al., 2000). Eine *Schilderung unverständener Handlungselemente* liegt vor, wenn vom Zeugen Ereignisse geschildert werden, die er (noch) nicht verstanden hat und dementsprechend nur eine phänomenologische Beschreibung des tatsächlichen Geschehens erfolgt (Niehaus, 2008; Vrij et al., 2000). *Indirekt handlungsbezogene Schilderungen* sind Berichte über ähnliche Erlebnisse ohne direkten Zusammenhang zum eigentlichen Ereignis, die jedoch ineinander verschachtelt sind (Niehaus, 2008; Vrij et al., 2000). Eine *Schilderung eigener psychischer Vorgänge* ist die Äußerung eigener Ge-

danken und Gefühle des Zeugen (Steller & Köhnken, 1989; Vrij et al., 2000). *Schilderung fremdpsychischer Vorgänge* ist vorhanden, wenn fremde Gedanken und Emotionen ausdrückt werden (Steller & Köhnken, 1989; Vrij et al., 2000).

Tabelle 5. CBCA Kriterien (nach Steller & Köhnken, 1989).

Kognitive Merkmale	
Allgemeine Merkmale	Logische Konsistenz
	Unstrukturierte Darstellung
	Detailreichtum
Spezielle Inhalte	Kontextuelle Einbettung
	Interaktionsschilderungen
	Wiedergabe von Gesprächen
	Unerwartete Handlungskomplikationen
Inhaltliche Besonderheiten	Ungewöhnliche Details
	Nebensächliche Details
	Schilderung unverstandener Handlungselemente
	Indirekt handlungsbezogene Schilderungen
	Eigenpsychische Vorgänge
	Fremdpsychische Vorgänge
Motivationale Merkmale	
	Spontane Verbesserung der eigenen Aussage
	Eingeständnis von Erinnerungslücken
	Einwände gegen die eigene Aussage
	Selbstbelastungen
	Verteidigung des Beschuldigten
Deliktspezifische Merkmale	
	Deliktspezifische Aussageelemente

Ferner werden motivationale Merkmale durch die CBCA erfasst. Falsch aussagende Personen bemühen sich, im Sinne des Impression-Managements, den allgemeinschematischen Vorstellungen eines glaubhaft aussagenden Zeugen zu entsprechen (Volbert et al., 2010), dabei werden verschiedene Ziele verfolgt (Niehaus, Krause & Schmidke, 2005): Es soll ein kompetenter und makelloser Eindruck entstehen und die Glaubhaftigkeit des Beschuldigten untergraben werden. Die motivationalen Merkmale, welche durch die CBCA Kriterien erfasst werden, stehen im Gegensatz zu den Alltagsvorstellungen über einen „ehrlichen“ Zeugen. Spontane Verbesserung der eigenen Aussage, das Eingeständnis von Erinnerungslücken, Einwände gegen die eigene Aussage, Selbstbelastungen und die Verteidigung des Beschuldigten (Steller & Köhnken, 1989) sind Kriterien, die häufiger in Aussagen mit Erlebnisbasis zu finden sind. *Spontane Verbesserungen der eigenen Aussage* sind Informationen, die während der Aussage hinzugefügt werden oder Korrekturen des Gesagten ohne Aufforderung (Vrij, 2008; Vrij et al., 2000), weil diese im Gedächtnis unterschiedlich abgespeichert sind. Das Kriterium *Eingeständnisse von Erinnerungslücken* ist erfüllt, wenn die befragte Person angibt, bestimmte Teile des Geschehens nicht mehr erinnern zu können (Vrij, 2008; Vrij et al., 2000). *Einwände gegen die Richtigkeit der Aussage* liegen vor, wenn der Zeuge eingesteht, dass Teile seiner eigenen Aussage unglaubwürdig wirken oder falsch sein könnten (Vrij, 2008; Vrij et al., 2000). *Selbstbelastungen* sind Bestandteile einer Aussage, die für den Zeugen nachteilig sind (Vrij et al., 2000; Vrij et al., 2008). Das Kriterium der *Verteidigung des Beschuldigten* gilt als erfüllt, wenn der Zeuge die Handlungen des Täters entschuldigt oder Informationen äußert, welche die Schuld verringern (Vrij, 2008; Vrij et al., 2000).

Neben den kognitiven und motivationalen Merkmalen werden *deliktspezifische Aussageelemente* im Rahmen der CBCA erhoben. Das Kriterium ist erfüllt, wenn die aussagende Person über Wissen verfügt, welches ihr die Abgabe einer glaubhaften Falschaussage erleichtern würde (Volbert et al., 2010). Eine besonders gute Differenzierung zwischen wahren und erfundenen Aussagen bieten die Kriterien der phänomengemäßen Schilderung unverstandener Handlungselemente oder der Bericht von Komplikationen im Handlungsablauf. Die Äußerung eigener Gefühle ist hingegen eher zu vernachlässigen bei der Beurteilung der Erlebnisbasis (Oberlader et al., 2016).

Die CBCA ist ein standardmäßiger Bestandteil bei der aussagepsychologischen Begutachtung im Rahmen der SVA. Die Validität und Reliabilität der Kriterien wurde in

den letzten Jahren durch zahlreiche empirische Studien untersucht (Oberlader et al., 2016; Vrij et al., 2000, Vrij et al., 2008). Nach der Feststellung der Einzelkriterien wird eine psychometrisch und mathematisch begründete Aggregation zur Gesamtbeurteilung vorgenommen. Dadurch wird, trotz der geringen Einzelvaliditäten der Kriterien, eine Aufhebung der statistisch unabhängigen Fehlermerkmale erreicht, womit eine diagnostisch wertvolle Grundlage zur Beurteilung des Wahrheitsgehalts einer Aussage resultiert (Fiedler & Schmid, 1999). Dieses Vorgehen hat auch zur Anerkennung der Methodik im Rahmen der SVA geführt (BGHSt, 2000). Einschränkend ist jedoch anzumerken, dass die Glaubhaftigkeitsbegutachtung nicht zwischen wahren und suggerierten Aussagen unterscheiden kann oder bei sehr kurzen Aussagen nicht anwendbar ist (Volbert et al., 2010). Im Folgenden wird auf Interviewtechniken und Verfahren eingegangen, welche die CBCA Kriterien teilweise als Bestandteil einer Zeugenbefragung abfragen.

Analog zur CBCA soll in der vorliegenden Arbeit eine qualitative Auswertung der Daten erfolgen. Dabei sollen Erinnerungen hinsichtlich ihres Inhalts und der Auftretenshäufigkeit bewertet werden und übergeordneten Kategorien zugeordnet werden.

3.3 Das kognitive Interview

Geiselman, Fisher, MacKinnon und Holland entwickelten 1986 in Zusammenarbeit mit der US-amerikanischen Polizeibehörde die Vernehmungstechnik des kognitiven Interviews. Zeugenbefragungen zählen zu den Routinetätigkeiten eines Polizeibeamten, jedoch fehlt es an methodisch basierten Techniken, um eine möglichst akkurate und fehlerfreie Aussage zu erhalten. Interviewverfahren sind oftmals durch zahlreiche Fragen seitens des Interviewers gekennzeichnet (Geiselman, & Fisher, 2014; Memon & Bull, 1991). Bei Befragungen zu einem konkreten Sachverhalt wird häufig anhand von geschlossenen Fragen Wissen in Ja oder Nein Antworten erfragt. Dieses Frageformat führt oftmals zu Suggestionseffekten, es erfolgt kein episodisches Erinnern und eine Rekonstruktion der eigenen episodischen Erinnerungen wird für den Zeugen erschwert. Insbesondere ängstliche oder angespannte Zeugen sind bei solchen Fragen suggestionsanfällig (Arntzen, 2008).

Ziel war es, ein für Vernehmungsbeamte eigenständig anwendbares Verfahren zu entwickeln, welches mit Hilfe kognitiver Techniken die Informationsmenge einer Zeugenaussage erhöht und die Fehleranfälligkeit vermindert. Das Interviewverfahren ist im

Hinblick auf Ergebnisse der gedächtnispsychologischen Grundlagenforschung konzipiert und um Elemente der Kommunikation und Gesprächsführung ergänzt worden (Fisher, Geiselman, Raymond, Jurkevich & Wahrhaftig, 1987). Gemäß der *encoding specificity* Hypothese (Tulving & Thomson, 1973) können Hinweisreize die Erinnerung an Einzelheiten vereinfachen. Dies kann durch die Vorgabe des Kontextes, in dem die Erinnerung erworben wurde, erfolgen. Die Reaktivierung bezieht sich dabei auch auf Affekte, welche bei Übereinstimmung in der Abruf- und Wahrnehmungssituation die Erinnerung fördern können (Bower, Gilligan & Monteiro, 1981). Auch die Aufforderung zur Schilderung aller Details, auch wenn nicht alles mit dem fraglichen Ereignis zusammenhängt, führt zu einer besseren Erinnerungsleistung. Dies wurde aus der Multi-Komponenten Theorie von Baddeley und Hitch (1974) abgeleitet, der zufolge das Gedächtnis aus einem Netzwerk besteht, welches durch assoziative Verknüpfungen gekennzeichnet ist (Collins & Loftus, 1975) (siehe Kapitel 1 „Gedächtnisprozesse und Modellvorstellungen“). Des Weiteren sind die Änderung der Erzählreihenfolge und die Einnahme einer anderen mentalen Position als der eigenen der Erinnerung zuträglich (Anderson & Pichert, 1978).

Der freie Bericht im Rahmen des Kognitiven Interviews beinhaltet die eigenständige und unsystematische Reproduktion und Wiedergabe eines Ereignisses und damit eingehender Details oder beobachteter Begebenheiten. Hierbei werden weder Abrufhilfen noch spezifische Fragstellungen verwendet (Roebbers & Elischberger, 2002). Zu Beginn wird der Befragte mit einer offenen Anstoßfrage zu einem expliziten Erinnerungsabruf aufgefordert, während dem er nicht unterbrochen wird (Geiselman & Fisher, 2014). Der Interviewer macht sich währenddessen Notizen, um anschließend unter Nutzung von Abrufhilfen gezielte, nicht-suggestive Fragen zu stellen (Geiselman & Fisher, 2014).

Ein entscheidender Unterschied zu anderen Vernehmungsmethoden liegt in den vier Basistechniken des kognitiven Interviews, in welche der Zeuge zu Beginn des Interviews eingewiesen wird (Geiselman & Fisher, 2014). Mit der ersten Technik *context reinstatement* wird er gebeten, sich mental in den Kontext des fragten Ereignisses in Bezug auf den Ort, seine Emotionen und Gedanken zurückzusetzen (Memon, Wark, Bull & Köhnken, 1997). Im Hinblick auf die Kontextspezifität des Gedächtnisses (Buchner & Brandt, 2008; Tulving, 1983) wird hiermit ein besserer Abruf der Erinnerung angenommen. Da durch eine Kongruenz zwischen Lern- und Abrufkontext ein

akkuratere Abruf erfolgen kann. Es wird zwischen mentalem und physischen Reinstatement unterschieden, welches den Abruf erleichtern kann, aber auch schon Kontextinformationen, die 24 Stunden vor einer Befragung gegeben werden, üben einen positiven Effekt aus (Priestley, Roberts & Pipe, 1999). Hershkowitz, Orbach, Lamb, Sternberg und Horowitz (2001) konnten in einer empirischen Untersuchung zeigen, dass die Technik zwar nicht die rein numerische Anzahl der insgesamt erhobenen Informationen beeinflussen kann, jedoch erhöhte sich der Umfang an Informationen die frei und spontan abgerufen werden und mit einer größeren Wahrscheinlichkeit akkurat sind.

Anschließend folgt mit dem *detailed recall* der freie Bericht des Zeugen, bei dem er seine Erinnerung so vollständig wie möglich ohne Beachtung der Reihenfolge wiedergeben soll. Die dritte Technik des *recall events in different orders* fordert den Zeugen auf, seine Erinnerung im selben Gespräch wiederholt möglichst vollständig wiederzugeben, aber in einer anderen, frei gewählten zeitlichen Reihenfolge (Geiselman et al., 1986). Eine weitere Möglichkeit, andere Abrufpfade für Erinnerungen als die bisher angesprochenen zu aktivieren, bietet die vierte Technik der *changed perspectives*. Hierbei soll sich der Zeuge in andere Perspektiven in Bezug auf das beobachtete oder erlebte Ereignis versetzen.

Die trichterförmige Befragung ist Teil des Kognitiven Interviews und erfolgt klassischerweise nach der freien Spontanschilderung (Balloff, 2004). Zunächst sollten die Fragen so offen wie möglich sein und erst im Verlauf spezifischer werden (Steller & Volbert, 1999; Volbert et al., 2010). Suggestive Fragen sollten zu jeder Zeit vermieden werden. Da offene Fragen und die Aufforderung zu freiem Bericht zu längeren Antworten und einer höheren Aussagequalität führen, sind diese geschlossenen vorzuziehen (Vrij, 2005). Kindern, die von einem Vorfall berichten, der nicht stattgefunden hat, zeigen durch offene Fragen mehr CBCA Kriterien in ihren Aussagen, als bei einer Befragung mit geschlossenen Fragen (Craig, Scheibe, Raskin, Kircher & Dodd, 1999). Bei einem Erlebnis mit Erlebnisbasis resultierten die meisten CBCA Kriterien durch eine offene Befragung (Craig et al., 1999).

Die Verwendung von freien oder vorgegebenen Antwortformaten wird unter anderem auch in der Markt-, Meinungs- und Umfrageforschung systematisch untersucht. Freie Antwortformate werden als Produktionsmethode betrachtet, die große Datenmengen

zur Folge haben (Felser, 2015), welche nur schwer zu handhaben sind. Durch die Verwendung eines Forced-choice Antwortformats oder den Verzicht auf Mittelkategorien (Likert-Skala mit ungerader Anzahl an Wahlalternativen) kann eine Tendenz zur Mitte vermieden werden, gleichzeitig resultiert ein relativer Informationsverlust. In einer Untersuchung von Hassebrauck (1995) zeigte sich, dass ein Vergleich der *Produktionsmethode*, welche freie Assoziationen zum Thema „Liebe“ oder der *Urteilsmethode*, welche Zustimmung zu vorgegebenen Begriffen nutzt, zu sehr unterschiedlichen Ergebnissen führen kann. Aufgrund des Mehraufwands bei der Auswertung werden grundsätzlich geschlossene Frageformate empfohlen, jedoch sind die interessierenden Inhalte hier ausschlaggebend. Strack (1994) konstatiert, dass der Einfluss des Frageformats immer auch unter Berücksichtigung von Frageinhalt oder Personenmerkmalen betrachtet werden muss. Beide Methoden können zu unterschiedlichen Ergebnissen führen, ursächlich dafür sind hier die kognitiven Prozesse, die zu einer freien spontanen Produktion oder Zustimmung auf Nachfrage führen. Unklar bleibt allerdings, ob das spontane oder bestätigende Urteil im Rahmen der Meinungsforschung relevanter bzw. nützlicher ist (Felser, 2015; Strack, 1994).

Memon und Bull (1991) berichten positive Ergebnisse aus Studien, in denen Kinder zwischen sieben und zwölf Jahren zum einen mit einem Standardinterviewverfahren und zum anderen mit dem Kognitiven Interview befragt wurden. In beiden Studien zeigte sich das Kognitive Interview im Hinblick auf die Menge korrekt erinnelter Fakten überlegen. In Bezug auf die Fehlerrate zeigen sich keine signifikanten Unterschiede (Köhnken, Kraus & vom Schemm, 2008).

Saywitz, Geiselman und Bornstein (1992) konnten in einer empirischen Evaluationsstudie feststellen, dass die Anwendung des Kognitiven Interviews bei der Befragung von Kindern zwischen acht und zwölf zu signifikant mehr richtig erinnerten Antworten führte (*recall*), ohne dass sich die Anzahl an Fehlern erhöhte. Mit dem Kognitiven Interview befragte Kinderzeugen zeigen genauere Informationen als Kinder, die mit anderen Techniken befragt wurden (Köhnken, Milne, Memon, & Bull, 1999).

Dies führte zu der Schlussfolgerung, dass das Kognitive Interview anderen Interviews bei der Erhebung forensisch relevanter Informationen überlegen ist. Insgesamt wurden in den USA, in Großbritannien und Deutschland zahlreiche empirische Studien zur Überprüfung und Evaluation der Technik des Kognitiven Interviews durchgeführt

(Köhnken et al., 1999; Köhnken, Schimossek, Aschermann & Höfer, 1995), welche die Anwendung in der Praxis als Standard nahelegen.

3.4 Aktuelle Interviewtechniken und -verfahren

Im internationalen Vergleich liegen diverse Interview- oder Vernehmungstechniken vor. Darüber hinaus wurden Leitfäden und Handanweisungen zur Befragung von Zeugen, Opfern und Tätern entwickelt. Dabei unterscheiden sich die Zielsetzungen, der Standardisierungsgrad und die Zielgruppen zur Anwendung der jeweiligen Verfahren. Im Folgenden soll eine Auswahl dieser Techniken und Verfahren in ihren Grundzügen dargestellt werden, dabei werden zunächst zwei Techniken (Inbau, Reid, Buckley & Jayne, 2011; Milne & Bull, 2003) und anschließend Verfahren bzw. Handlungsanweisungen beschrieben (Home Office, 2011; Lamb et al., 2007). Zum Schluss wird ein aktuelles Analyseinstrument zur Bewertung von Experteninterviews vorgestellt (Dodier & Denault, 2018).

Im Rahmen polizeilicher Ermittlungen sind in den USA die *Reid Technique* (Inbau, Reid, Buckley & Jayne, 2011) und in Großbritannien der *PEACE* (Planning & Preparation; Engage & Explain; Account; Closure; Evaluation)-Ansatz (Milne & Bull, 2003) weit verbreitet. Die Reid Technique ist eine konfrontative geständnisorientierte Vernehmungsmethode, wobei das Ablegen eines Geständnisses durch den Beschuldigten als Ziel verfolgt wird. Dabei wird auf geständnishemmende und geständnisbegünstigende Faktoren eingegangen, sodass es für den Beschuldigten günstiger erscheint zu gestehen. Mit Hilfe von Minimierungstechniken soll Verständnis für den Befragten signalisiert werden, dabei werden Rationalisierungs- und Bagatellisierungstendenzen beim Befragten gefördert, sodass negative Konsequenzen für das Selbstbild abgeschwächt werden und Hoffnungen auf eine milde Strafe entstehen. Der *PEACE*-Ansatz hingegen ist ein informationssammelnder Ansatz, bei dem aktives Zuhören und offene Befragungen im Vordergrund stehen. Im Gegensatz zur Reid Technique soll durch den *PEACE*-Ansatz die Gefahr falscher Geständnisse in polizeilichen Befragungen vermindert werden.

Zur Befragung von erwachsenen und insbesondere kindlichen Opferzeugen wurden halbstandardisierte Leitfadeninterviews und Handanweisungen entwickelt, welche auf Basis gedächtnispsychologischer Erkenntnisse die Aussagemenge und Qualität bzw.

Genauigkeit von Aussagen verbessern sollen. Diese werden im folgenden Abschnitt exemplarisch dargestellt.

Das U.S. National Institute of Child Health and Disease (NICHD; Lamb et al., 2007) hat ein Interview-Protokoll veröffentlicht, welches Standards bei der Befragung kindlicher Zeugen festlegt. Dabei wurde empirisches Wissen über gedächtnispsychologische Besonderheiten von Kindern, kommunikative und soziale Fähigkeiten und soziale Tendenzen in Handlungsanleitungen übersetzt, um die Qualität forensischer Interviews zu verbessern. Die Wirksamkeit des NICHD Protokolls konnte in kontrollierten Studien nachgewiesen werden, da sich die Interrater-Reliabilität der Interviews bei Anwendung des NICHD Protokolls signifikant verbesserte. Interviewer, die im Umgang mit dem Protokoll geschult sind, erhalten mehr Informationen durch die Anwendung offener Fragen, sie führen strukturiertere Befragungen durch und können fokussierte Fragen besser einsetzen (Orbach et al., 2000; Pipe et al., 2004). Das NICHD liegt in einer deutschen Übersetzung von Christmann und Wazlawik aus dem Jahr 2016 vor. Zentrale Elemente und Abschnitte der Befragung lauten: I. Einleitung, II. Rapport, III. Training des episodischen Gedächtnisses IV. Befragung zu substantiellen Themen und V. Untersuchung der Vorfälle. Ergänzungen in der Handanweisung betreffen zum Beispiel die Erhebungen von Informationen, welche vom Kind nicht genannt wurden und auch der Abschluss der Befragung mit einem neutralen Thema wird beschrieben.

Das Home Office als Organ der britischen Regierung behandelt Fragen der inneren Sicherheit. In Zusammenarbeit mit dem Justiz- und Gesundheitsministerium ist das Home Office (2011) Herausgeber der *Achieving Best Evidence Guidelines (ABE)* zur Befragung von Zeugen und Opfern, um das bestmögliche Vorgehen bei der Bearbeitung von Kriminalfällen zu sichern. Neben der konkreten Planung einer Zeugenbefragung werden zum Beispiel grundlegende anamnestische Befragungsaspekte, wie Alter, Geschlecht, Ethnie etc. festgehalten, die ein Interviewer erfassen sollte. Zu den erhobenen Aspekten werden immer auch mögliche Effekte auf das Verhalten und die Vulnerabilität der kindlichen und erwachsenen Zeugen dargestellt. So werden z.B. ein Migrationshintergrund und das eventuelle Erleben von Rassismus als mögliche Gründe für einen niedrigen Selbstwert und Angst vor Diskriminierung angeführt. Im weiteren Verlauf werden wie auch im NICHD die vier Phasen des Rapports, des Einübens der Narration und die Befragung sowie das Ende als Interviewphasen beschrieben. Für die schriftliche Dokumentation werden Hinweise zur Zusammenfassung und

Nutzung eventueller Videoaufnahmen der Befragung gegeben. Die Nutzung von Bildern und (anatomischen) Puppen während der Befragung wird in der Handanweisung unter praktischen Gesichtspunkten vor dem Hintergrund empirischer Ergebnisse erläutert.

Dodier und Denault (2018) entwickelten die *Griffith Question Map (GQM)*, welche ein forensisches Werkzeug für Experten darstellt und bei der Beurteilung der Aussagen und Befragungen von Opfern und jungen Kindern helfen soll. Dazu entwickelten die Autoren in Anlehnung an die ABE Guidelines (Home Office, 2011) ein Frage-Gitter, welches eine ausführliche retrospektive Interview-Analyse ermöglicht und dabei die verwendeten Fragen vor dem Hintergrund empirischer Forschungsergebnisse zur forensischen Befragungspraxis beschreibt. Die Autoren der GQM konstatieren, dass die Art der Fragen durch einen Interviewer der signifikanteste Faktor für die Qualität von Aussagen darstellt und demnach einer spezifischen empirischen Betrachtung und Evaluation bedarf. Dabei identifiziert die GQM den Fragentyp als eine Funktion seines Kontexts und differenziert angemessene und unangemessene Fragen als zwei Hauptkategorien. So kann eine offene Frage beispielsweise als unangemessen bewertet werden, wenn sie die Rede des Zeugen unterbricht, oder eine geschlossene Frage bei der Zusammenfassung von Inhalten als angemessen gelten. Die Interrater-Reliabilität erweist sich bei der Anwendung der GQM auf ein Fallbeispiel als äußerst zufriedenstellend (Cohens Kappa $\kappa = .93$, $p < .001$; Dodier & Denault, 2018).

4 Zuverlässigkeitsmotivation

Den qualitativen und motivationalen Aspekten der Gedächtnisleistung wird in der aussagepsychologischen Forschung zunehmend mehr Bedeutung beigemessen. Entgegen der anfänglichen Fokussierung auf die Aussagemenge befassen sich die Arbeiten nun mit der möglichst genauen, das heißt detaillierten und originalgetreuen Wiedergabe von Erinnerungen bei gleichzeitiger Zurückhaltung von unsicheren Informationen (Koriat, Goldsmith & Pansky, 2000; Roebbers, 2010). Während die Aussagemenge an den Input gebunden ist und anhand des Prozentsatzes korrekt beantworteter Fragen gemessen werden kann, ist die Aussagezuverlässigkeit (hier deutsche Übersetzung für engl. „accuracy“) an den Output gebunden und definiert als die Wahrscheinlichkeit dafür, dass eine Aussage korrekt ist (Koriat et al., 2001) und nicht auf Irrtum beruht

(Regber, 2007). Zuverlässigkeitsmotivation ist die deutsche Übersetzung des Begriffs für *accuracy motivation*.

Roebbers und Kollegen konnten in verschiedenen Untersuchungen (zum Beispiel Roebbers & Fernandez, 2002; Roebbers, Moga & Schneider, 2001; Roebbers & Schneider, 2005a) zeigen, dass die Antwortqualität von der Stärke der *Zuverlässigkeitsmotivation* abhängt, wobei Qualität die Richtigkeit und Genauigkeit einer Aussage umfasst. Die Richtigkeit einer Aussage ist von der Behaltens- und Abrufleistung abhängig, deren Wiedergabe nach Koriat und Goldsmith (1994, 1996) einen aktiven Prozess des Handelns unter Abwägung persönlicher Ziele und situativer Anforderungen beinhaltet. Die Genauigkeit einer Aussage ist neben der Behaltensleistung, welche sich auf die Quantität der geäußerten Informationen auswirkt, abhängig von der individuellen strategischen Kontrolle und unterliegt demnach der Zuverlässigkeitsmotivation (Volbert, 2005). Eine Darstellung der empirischen Studien findet sich in Abschnitt 4.2.1 „Ich weiß nicht“-Antwortoption.

Roebbers et al. (2001) operationalisierten Zuverlässigkeitsmotivation über die Möglichkeit zur Zurückhaltung einer Antwort, die sogenannte *IWN*-Antwortoption, welche im Ergebnis mit einer mittleren Zuverlässigkeitsmotivation einherging. Kinder sind demnach bereits ab einem Alter von sechs Jahren in der Lage unsichere Antworten zurückzuhalten, wenn richtige Antworten belohnt und falsche Antworten bestraft werden. Jedoch wurde im Rahmen der Untersuchung die *IWN*-Antwortoption von ihnen weniger genutzt als von Erwachsenen. Es wird somit angenommen, dass jüngere Kinder im Vergleich zu Erwachsenen geringere Fähigkeiten zum Gedächtnismonitoring besitzen.

Eine Belohnung richtiger Antworten resultierte als Ergebnis einer weiteren Studie in einer hohen Zuverlässigkeitsmotivation und qualitativ besseren Antworten (Roebbers & Fernandez, 2002). Roebbers und Fernandez (2002) wiesen nach, dass die Möglichkeit Antworten zurück zu halten, mit einer mittleren Zuverlässigkeitsmotivation einhergeht, während die Möglichkeit, Antworten zurückzuhalten und zusätzlich für jede richtige Antwort belohnt zu werden, zu einer hohen Zuverlässigkeitsmotivation und qualitativ besseren Erinnerungen führt. Hierzu ist jedoch anzumerken, dass die Fehlerrate auch bei einer hohen Zuverlässigkeitsmotivation bei suggestiven Fragen grundlegend höher ist als bei offenen Fragen.

4.1 Strategische Regulation der Aussagegenauigkeit

Wenn Menschen von vergangenen Ereignissen berichten, erzählen sie nicht einfach alles, was ihnen in den Sinn kommt, sondern sie fällen strategische Entscheidungen dahingehend, welche Aspekte der Geschichte sie hervorheben möchten, welche Perspektive ihnen wichtig ist und welches Level an Generalisierung oder Detailliertheit relevant ist (Goldsmith & Koriat, 1999; Koriat & Goldsmith, 1996). Diese Art strategischer Kontrolle kann einen substantiellen Effekt auf die Qualität von Erinnerungsberichten haben. Auf dieser Grundlage sind Koriat et al. (2001) der Frage nachgegangen, inwiefern Kinder in der Lage sind, ihre Aussagen bzw. Erinnerungsberichte im Hinblick auf spezifische Ziele strategisch zu kontrollieren. Koriat und Goldsmith (1996) postulieren in ihrem Modell der strategischen Regulationsleistung der Erinnerungsgenauigkeit und -menge einen Trade-off zwischen Aussagemenge und -genauigkeit, demzufolge die Erinnerungsleistung das Resultat eines komplexen Zusammenspiels von Gedächtnisleistungen (d.h. Speicherung und Abruf) und metakognitiven Überwachungsprozessen (Monitoring und Regulation) sowie von situativen (motivationalen) Faktoren ist. Hier wird zwischen dem, was tatsächlich erinnert wird, und dem, was eine Person entscheidet zu berichten, unterschieden. Der Abruf von Informationen aus dem Langzeitgedächtnis erfolgt unter der Hinzunahme von Monitoringprozessen, welche die Informationen hinsichtlich der Wahrscheinlichkeit ihrer Richtigkeit bewerten (siehe Abbildung 2).

Situative Faktoren, wie z.B. die Betonung von Aussagemenge und/ oder Aussagegenauigkeit oder das Vorhandensein von Anreizen oder Sanktionen für richtige und falsche Antworten (motivationale Faktoren gem. Koriat & Goldsmith, 1996), beeinflussen das sog. *Antwortkriterium*. Das Antwortkriterium stellt eine stetige Variable dar, welche angibt, ab welcher Wahrscheinlichkeit für die Richtigkeit einer Aussage diese kundgetan oder zurückgehalten wird.

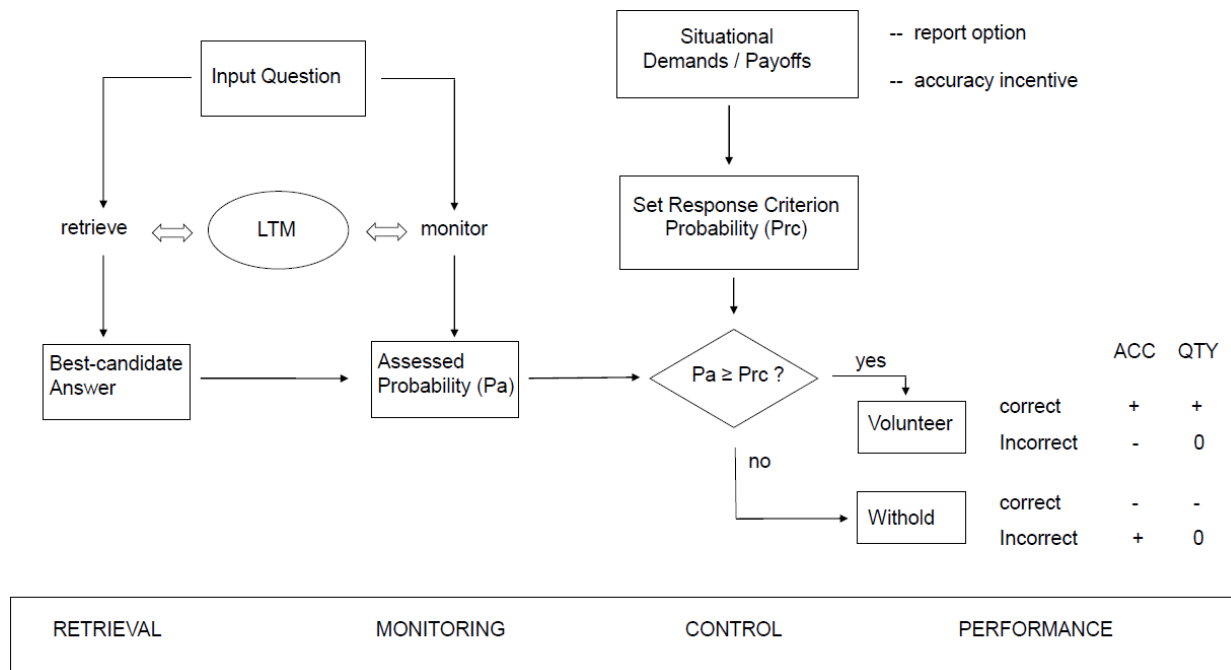


Abbildung 2. Das schematische Modell der strategischen Regulationsleistung der Erinnerungsgenauigkeit und –menge. Leistungseffekte sind mittels Plus (Anstieg), Minus (Abfall) und Null (kein Effekt) dargestellt; LTM= Langzeitgedächtnis, ACC=Aussagegenauigkeit, QTY=Menge, Pa=geschätzte Wahrscheinlichkeit für Richtigkeit der Antwortmöglichkeit, Prc=Antwortkriterium (nach Koriat & Goldsmith, 1996).

Die Beeinflussung des Antwortkriteriums kann über interne und externe Prozesse erfolgen bzw. resultiert aus einem Zusammenspiel dieser. Ist das Antwortkriterium liberal (siehe Abbildung 3), so ist die Wahrscheinlichkeit für falsche Antworten größer und es erfolgt eher bzw. schneller eine Entscheidung für die Preisgabe einer Antwort (siehe Signal Detection Theory von Green & Swets, 1966). Ist das Antwortkriterium strikt, so werden nur Antworten gegeben, die mit hoher Wahrscheinlichkeit als richtig eingeschätzt wurden. Dabei ist die Aussagemenge geringer und die Aussagegenauigkeit hoch, was dem von Koriat und Goldsmith (1996) beschriebenen Trade-off entspricht. Demnach wird eine Antwort zurückgehalten, wenn die Wahrscheinlichkeit, dass die beste, aus dem Langzeitgedächtnis abgerufene Antwortmöglichkeit richtig ist, unterhalb des Antwortkriteriums liegt.

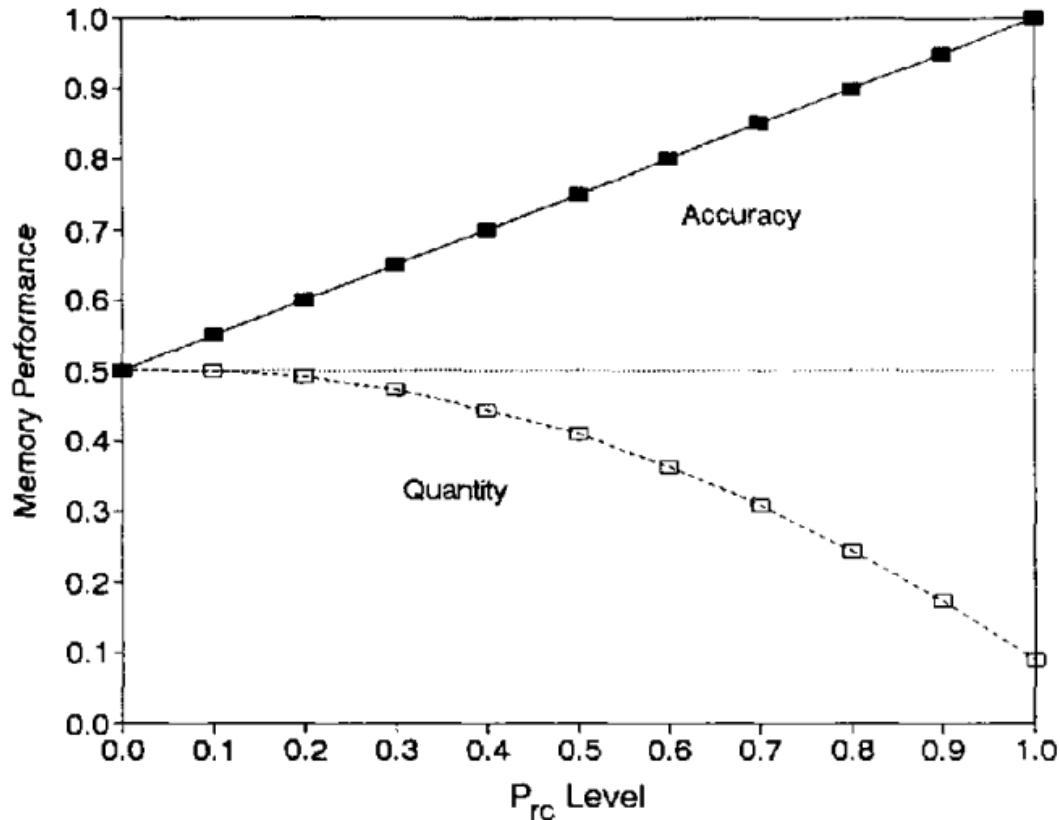


Abbildung 3. Simulierte Gedächtnisleistung hinsichtlich Erinnerungsmenge und -genauigkeit als Funktion des Antwortkriteriums P_{rc} (nach Koriat & Goldsmith, 1996).

Ist diese Antwortmöglichkeit tatsächlich falsch, erhöht dies die Aussagegenauigkeit, ohne Auswirkungen auf die Aussagemenge zu haben (die schließlich definiert ist als die Menge richtiger Aussagen, s.o.). Ist die zurückgehaltene Antwort hingegen richtig, werden Aussagegenauigkeit und -menge reduziert. Wenn die Wahrscheinlichkeit, dass die aus dem Langzeitgedächtnis abgerufene Antwortmöglichkeit richtig ist, das Antwortkriterium übersteigt, erfolgt eine Aussage. Ist diese Aussage falsch, reduziert dies die Aussagegenauigkeit, ohne Auswirkungen auf die Aussagemenge zu haben. Ist diese Aussage richtig, steigen Aussagemenge und -genauigkeit.

Damit kognitive Kontrollprozesse das System effektiv regulieren können, sind Informationen notwendig. Gemäß Nelson und Narens (1994) wird das Monitoring zum einen durch den aktuellen Informations-Input als auch durch nicht aktuell präsente (Gedächtnis) Informationen beeinflusst. Es wird zwischen retrospektivem und prospektivem Monitoring, wobei diese sich hinsichtlich der Beurteilung einer bereits gegebenen oder

bevorstehenden Antwort unterscheiden. Diese Prozesse werden wiederum in drei Kategorien unterteilt: Ease of learning (EOL); Judgements of learning (JOL) und Feeling of Knowing (FOK). EOL stellt Urteile darüber bereit, wie einfach oder schwer etwas zu lernen sein wird, wodurch auch entsprechende Lernstrategien abgeleitet werden. Prozesse des JOL finden nach dem Erkennen statt und stellen auf der Grundlage aktuell verfügbarer bzw. erinnelter Items Vorhersagen über zukünftige Testperformanz bereit. Das FOK tritt während oder nach der Akquisition von Informationen auf und beurteilt, ob ein Item bekannt ist oder in einem nachfolgenden Test erinnert werden wird (Nelson & Narens, 1994). Die drei Kategorien korrelieren nicht untereinander, was eine multidimensionale Struktur der Prozesse nahelegt. Hart hat bereits 1967 den Begriff des Memory Monitoring geprägt und Erkenntnisse zum Feeling of Knowing aufgetan, die sich auf die individuelle Einschätzung beziehen, welche Erinnerungen mit welcher Wahrscheinlichkeit erinnert werden können. Kinder haben Nelson und Narens (1994) zufolge weniger gute Regulations- und Monitoringfähigkeiten als Erwachsene.

Verhalten wird als Funktion aus Person und Umwelt verstanden (Lewin, 1963; Westhoff & Kluck, 2014). Insgesamt kann die Zuverlässigkeitsmotivation durch in der Person oder der Situation liegende Merkmale manipuliert werden. Dabei wird z.B. über Anreizsysteme die Motivation beeinflusst, was in der vorliegenden Arbeit eine der experimentellen Variationen darstellt.

4.2 Ansätze zur Operationalisierung

Eine einheitliche Operationalisierung des Konzepts der Zuverlässigkeitsmotivation ist bisher nicht vorhanden. Vielmehr kann dies im Kontext der Aussagepsychologie als die Einsicht in die Notwendigkeit, valide Angaben zu machen, verstanden werden. Vor Gericht bedeutet das unter anderem, dass ein Befragter sich im Klaren darüber sein muss, dass sich die forensische Befragungssituation von anderen Kommunikationssituationen unterscheidet. Er muss seine Erinnerungsbemühungen dahingehend einstellen, dass Gedächtnisinhalte möglichst vollständig abgerufen werden, aber auch, dass er bei Unsicherheiten Erinnerungslücken angibt (Steller, 2008).

In Anlehnung an Pipe et al. (2004) lässt sich die Zuverlässigkeitsmotivation personalen und situativen Merkmalen zuordnen, wobei die Forschungsarbeiten zur Zeugenbefragung sich in vier Kategorien unterteilen lassen, nämlich Abruf und Regulation, soziale Aspekte und Aspekte der Fragesituation bzw. Befragungstechniken. Diese lassen sich

analog zur Strukturierung der Studien zum kindlichen Source Monitoring (siehe Kap. 1.3.1) den Personenmerkmalen (Abruf und Regulation), der Lernsituation (Soziale Aspekte) und der Abrufsituation (Abruf und Regulation, Soziale Aspekte, Motivationale Aspekte) zuordnen.

In der vorliegenden Arbeit wird die Operationalisierung der Zuverlässigkeitsmotivation im Gegensatz zu bisherigen Untersuchungen neben der Möglichkeit, unsichere Antworten zurückzuhalten (*IWN*-Antwortoption), über eine Bestrafungskomponente realisiert. Damit wird die Relevanz von akkuraten Angaben im forensischen Kontext adäquat repräsentiert, denn inkorrekte Zeugenaussagen können schwerwiegende Konsequenzen haben (Greuel, 2001). Im deutschen Strafrecht sind (bewusste) Falschaussagen strafbar (§153 Strafgesetzbuch (StGB)), dies gilt jedoch nur für strafmündige Personen. Zweifelsohne können aber auch Falschaussagen von Kindern zu Fehlurteilen führen.

4.2.1 „Ich weiß nicht“-Antwortoption

Koriat und Goldsmith (1994, 1996) stellten fest, dass erwachsene Probanden qualitativ höherwertige bzw. akkuratere Antworten liefern, wenn ihnen die Antwortoption „Ich weiß nicht“ angeboten wird. Dabei gehen die Autoren davon aus, dass durch die Möglichkeit einer fakultativen Antwort Monitoring- und Kontrollprozesse bei der Überprüfung von Erinnerungen angeregt werden. Dies geht mit aussagepsychologischen Erkenntnissen im Kontext der Befragung kindlicher Zeugen einher: Kinder sind zwar selbstständig in der Lage, Informationen korrekt abzurufen, neben anderen Faktoren, ist aber die Fähigkeit des Interviewers, Informationen „auszulösen“ und zu erheben von besonderer Relevanz für die Aussagemenge und -qualität (Lamb et al., 2007). Hinsichtlich der inhaltlichen Qualität gehen die Autoren davon aus, dass dann auf ein strikteres Antwortkriterium zurückgegriffen wird, wenn Kontrollprozesse angeregt werden (Lamb et al., 2007). Gleichzeitig werden metakognitive Prozesse aktiviert, die in einem Testformat, welches die Teilnehmer zwingt eine Antwort zu geben und bei Unsicherheit zu raten, nicht angeregt werden. Dies führt dazu, dass die Möglichkeit, Antworten zurück zu halten, die Qualität der gegebenen Antworten ansteigen und die Quantität der Antworten abfallen lässt. Roebbers, Moga und Schneider (2001) konnten zeigen, dass Erwachsene bessere Leistung in Wiedererkennungsaufgaben erbrachten, wenn eine *IWN*-Antwortoption vorhanden war und dass Kinder bereits ab einem

Alter von acht Jahren in der Lage waren, ihre Gedächtnisleistungen strategisch zu kontrollieren. Kinder erzielten hier vergleichbare Ergebnisse wie Erwachsene.

Ungeklärt in diesem Zusammenhang bleibt jedoch der Effekt der *IWN*-Antwortoption, welcher noch undifferenziert erforscht ist (siehe Kapitel 1.2 „Gedächtnisleistung und Source Monitoring von Kindern“). Diesem wird in der vorliegenden Arbeit im Folgenden nachgegangen.

Nesbitt und Markham (1999) konnten die Aussagegenauigkeit von drei- bis fünfjährigen Kindern durch ein Monitoring-Training verbessern, welches auf die Fähigkeit zur Unterscheidung zwischen Wissen und Raten abzielte. Kinder lernten zunächst in Gruppen von acht Teilnehmern und später im Einzelsetting bei Unsicherheit *IWN*-Antworten zu geben, wodurch sich die Anzahl korrekt positiver Antworten (*Hits*) reduzierte, jedoch gleichzeitig die Anzahl falscher Alarme (*false Alarm*) ebenfalls erniedrigte, was dem klassischen Qualitäts-Quantitäts Trade-off entspricht. Die Tragweite falsch-positiver Antworten ist in der forensischen Realität von besonderer Bedeutsamkeit, so ist ggf. eine inkorrekte Identifikation schlimmer als Einbußen in der Antwortmenge.

Brady, Poole, Warren und Jones (1999) konnten bei der Befragung von Kindern zwischen drei und sechs Jahren ($N = 56$) mit geschlossenen Frageformaten (Ja, Nein) keinen Effekt des Frageformats auf ihre Aussagegenauigkeit finden. Sie befragten die Kinder zweimal zu einem videobasierten Ereignis, dabei wurde entweder im Standardformat oder in einem modifizierten, erzwungenen Format befragt. In Einklang mit früheren Studien (Roebbers et al., 2001), zeigten ältere Kinder bessere Leistungen. Das Frageformat hatte weder einen Einfluss auf die Akkuratheit der Erinnerungen noch auf die Neigung, "Ich weiß nicht" zu antworten oder die Konsistenz bei wiederholten Fragen. Für die Mehrheit der Kinder wurden keine eindeutigen Antwortverzerrungen beobachtet, unabhängig vom Frageformat. Die Genauigkeit war bei Fragen zu *Hits* und *Correct Rejections* gleichwertig. Die Konsistenz und die Antworten auf die Fragen der Suggestibilitätstests waren nicht prädiktiv für die Leistung. Den Fehlern bei Ja-Nein-Fragen liegen mehrere Mechanismen zugrunde, damit bleibt das Ziel, die Genauigkeit der Antworten von Kindern nachzuprüfen, schwer fassbar.

Saywitz und Snyder (1993) konnten durch ein Monitoring-Training mit älteren Kindern zeigen, dass der Trade-off im Hinblick auf korrekte Antworten nach einem Gruppen-

training geringer ausfällt als bei einem Individualtraining, sodass hier einer Übergeneralisierung der *IWN*-Antwortoption entgegengewirkt wurde. Koriat et al. (2001) sehen neben entwicklungspsychologischen Defiziten in den Monitoringfähigkeiten eine weitere mögliche Erklärung für die geringere Aussagegenauigkeit von Kindern, so dass angenommen werden kann, dass metakognitive Prozesse als Mediator wirken. Dabei unterscheiden sie beim Testformat Recall (d.h. Narration) und Rekognition (d.h. Multiple-Choice-Frageformat) und die Antwortoption wird in free (der Befragte entscheidet, ob und welche Information er kundtut) und forced Choice unterteilt. Der Recall, also die freie Antwortoption, führt zu erhöhter Aussagegenauigkeit, wohingegen Rekognition und forced choice die Aussagemenge vergrößern. In den meisten Studien seien jeweils free recall und forced recognition konfundiert, weshalb Koriat et al. (2001) von Folgendem ausgehen:

If young children can in fact strategically regulate their memory reporting to produce a more accurate record of past events, then they should be allowed – even encouraged – to decide for themselves which items of information to volunteer and which to withhold. The option of free report, combined with an explicit message that puts a premium on accurate reporting, will be expected to elicit more dependable testimony from a child than questioning methods that take away such control. By contrast, if children cannot be trusted to screen their own memory reports effectively, then perhaps forced-report questioning methods would be preferable (p. 412).

Koriat et al. (2001) befragten $N = 120$ sieben- bis elfjährige Kinder in einem Recall- bzw. Rekognitionsdesign mit free und forced-choice-Antwortoptionen. Dabei wurden Aussagemenge und -genauigkeit erhoben. Im Anschluss an die erste Befragung wurde die Abfolge der Befragungsmethode zwischen den Gruppen getauscht. Des Weiteren wurden in Experiment 1 moderate Anreize bzw. Punkte für korrekte Antworten gegeben. Im zweiten Experiment wurde dieser Anreiz manipuliert: Bei Recall und Rekognition (UV1) war die Antwortoption immer frei. Der Anreiz war entweder moderat oder hoch (bei falscher Antwort wurden alle Punkte abgezogen, UV2). Experiment 3 bestand aus einem Retest der Kinder aus Experiment 2 nach einem Jahr, bei dem es drei Testphasen gab: die beiden Phasen aus Experiment 1 sowie eine weitere, in der wie in Experiment 2 eine Manipulation des Anreizes vorgenommen wurde. Experiment 1 zeigte, dass die freie Antwortoption (81%) die Aussagegenauigkeit im Vergleich zu

forced-choice (68%) erhöhte. Dieser Effekt bestand auch noch in Experiment 3. Dies galt besonders in der Bedingung mit hohem Anreiz (88%) im Vergleich zum moderaten Anreiz (81%, Experiment 2). Auch dieser Effekt blieb in Experiment 3 bestehen. Zu beachten ist aber, dass es einen starken Abfall in Aussagemenge (v.a. bei Recall) und -genauigkeit nach einem Jahr im Vergleich zu den Ergebnissen aus Experiment 2 gab. Insgesamt sind Kinder dieser Altersklasse demnach zu adäquatem Monitoring und Regulation fähig. Ein hoher Anreiz führte zu einer Verstärkung dieses Effekts: Im Vergleich zum moderaten Anreiz wurden mehr Antworten, darunter auch mehr korrekte, zurückgehalten, was dem von Koriat und Goldsmith (1996) postulierten Trade-off entspricht. Die freie Rekognition war am effektivsten, da hierbei eine größere Aussagemenge zustande kam als beim Recall, aber die Aussagegenauigkeit mit der des Recalls vergleichbar war. Hinsichtlich des Alters zeigte sich, dass jüngere Kinder (auch in Experiment 3 noch) im Vergleich zu älteren ein liberaleres Antwortkriterium hatten, weniger effektiv beim Monitoring waren oder ihre Monitoringprozesse weniger in die Entscheidung bzgl. des Kundtuns bzw. Zurückhaltens von Antworten einbezogen. Bei Kindern im Alter von vier bis fünf Jahren liegen also noch entwicklungspsychologische Defizite beim Monitoring und der Regulation vor, die aber mittels Training teilweise überwunden werden können. Bei der älteren Stichprobe (sieben- bis achtjährige) von Koriat et al. (2001) zeigten sich diese Defizite nicht mehr. Dafür spielten hier Antwortoption und Anreiz zu richtigen Aussagen eine Rolle, inwiefern die Fähigkeiten eingesetzt wurden.

Roebbers et al. (2001) gehen davon aus, dass Kinder mehr als Erwachsene von einer Instruktion zur Zuverlässigkeitsmotivation profitieren, unter der Annahme, dass sie ihren Erinnerungsprozess im Prinzip überwachen können. Mit zunehmendem Alter wird davon ausgegangen, dass das Monitoring mehr automatisiert und weniger effektiv ist, was bedeutet, dass eine Manipulation des Antwortkriteriums weniger Einfluss auf die Genauigkeit der Erinnerung hat. Dabei ist der positive Effekt eines strikten Antwortkriteriums vom Ausmaß der individuellen Monitoringfähigkeiten abhängig, welche die Aussagequalität in einer freien oder forcierten Antwortbedingung bestimmen. In ihrer experimentellen Untersuchung wurden $N = 180$ Personen zwischen fünf und 22 Jahren untersucht und den Versuchsbedingungen *free report*, *free report plus incentive* und *forced* zugeordnet. Dabei wurden alle Teilnehmer in den zwei *free report*-Bedingungen instruiert, dass sie auf Fragen, deren Antwort sie nicht wissen, *IWN* antworten sollen.

In der forced-Bedingung sollten sie immer eine Antwort geben, wenn sie sich nicht sicher waren, sollten sie raten. In der free report plus incentive-Bedingung erhielten die Teilnehmer zusätzlich einen Token für eine richtige Antwort, bei einer falschen Antwort verblieb der Token in einem Behältnis. Dabei griffen die Autoren auf die Ergebnisse von O'Sullivan (1993) zurück, welcher zeigen konnte, dass bereits Vierjährige wissen, dass Erinnerungen besser und genauer sind, wenn man sich anstrengt - es zeigte sich, dass Belohnungen zu höherer Anstrengung führen. Den Teilnehmern wurden 27 Fragen mit offenem Ende zu zentralen und peripheren Details eines zuvor gesehenen Films gestellt. Dabei zeigten sich keine signifikanten Unterschiede hinsichtlich der Anzahl korrekter Antworten für die unterschiedlichen Bedingungen. Jedoch zeigte sich, dass 80,9 % der Fragen zu zentralen Details und nur 20,9 % zu peripheren Details korrekt beantwortet wurden und eine signifikante Interaktion zwischen Alter, Bedingung und Fragentyp bestand. In post-hoc Analysen stellte sich heraus, dass in den forced report-Bedingungen signifikant mehr Fragen korrekt beantwortet wurden als in den beiden free report-Bedingungen. Neben den Häufigkeiten wurde ein Qualitätsscore aus dem Verhältnis richtiger zur Gesamtanzahl beantworteten Antworten berechnet: Unabhängig vom Alter war die Qualität der Antworten in den free report plus incentive-Bedingungen am besten und in der forced report-Bedingung am schlechtesten. Dabei ergaben sich Haupteffekte für Alter und Bedingung, jedoch keine Interaktion. Mit zunehmendem Alter stiegen auch die Qualität und die Anzahl der Antworten. In der Bedingung free plus incentive waren die Teilnehmer besser als in der free report-Bedingung, welche wiederum jenen in der forced report-Bedingung überlegen waren. Bei der Analyse der inkorrekten und der *IWN*-Antworten wurde deutlich, dass beide in gleichem Maße zum signifikanten Effekt beitragen, was insofern relevant ist, als dass die Teilnehmer in der forced report-Bedingung nachvollziehbarerweise mehr inkorrekte Antworten gaben als in der free report-Bedingung und dort wiederum mehr als in der free report plus incentive-Bedingung. Entgegen der Annahmen der Autoren, kommt es zu keinem vollständigen Qualitäts-Quantitäts Trade-off i.S. des Modells von Koriat und Goldsmith (1996), was die Autoren auf das Verhältnis zwischen Belohnung und Bestrafung zurückführen, da sie den Kindern nicht alle Token für eine falsche Antwort abnahmen.

Roebers und Fernandez (2002) untersuchten mit einem analogen Versuchsdesign den Effekt der Zuverlässigkeitsmotivation auf die Suggestibilität von kindlichen und erwachsenen Versuchsteilnehmern ($N = 240$; sechs bis 51 Jahre). In den Versuchsbedingungen, welche mit absteigender Zuverlässigkeitsmotivation wiederum free report plus incentive, free report und forced report hießen, wurden vier verschiedene Fragetypen verwendet, beantwortbare und nicht beantwortbare Fragen, welche jeweils offen oder fehlleitend waren. Dabei hatten die Teilnehmer in den free-report Bedingungen eine *IWN*-Antwortoption und erhielten ein Token für eine korrekte Antwort, bei einer falschen Antwort erhielt der Interviewer das Token, was von den Autoren vorab nicht explizit als Bestrafung angesehen wurde. In der free report plus incentive-Bedingung resultierten die besten und akkuratesten Ergebnisse für offene und fehlleitende Fragen. Analog zu Roebers et al. (2001) wurde festgestellt, dass Kinder durch Belohnung in der Lage waren, inkorrekte Antworten zu unterdrücken. Auch der Qualitäts-Quantitäts Trade-off konnte gefunden werden: in der forced report-Bedingung resultierten prozentual mehr korrekte Antworten als in den free report-Bedingungen. Durch eine *IWN*-Option resultierten weniger, aber qualitativ bessere Antworten, durch zusätzliche Belohnung waren die Antworten qualitativ zudem am besten.

Roebers und Schneider (2005) nahmen eine Variation der o.g. Studien (Roebers & Fernandez, 2002; Roebers et al., 2001) vor und untersuchten zusätzlich zur *IWN*-Option die Wirkung von direktem Feedback auf in-/korrekte Antworten und Erinnerungen durch den Interviewer auf die Zuverlässigkeit. Dabei konnten sie feststellen, dass bei Kindern allein durch einen Hinweis zur Überwachung ihrer Antworten keine verbesserte Leistung resultiert. Nur eine Kombination von Feedback und Hinweis hatte eine positive Wirkung auf qualitative Maße der Erinnerung.

4.2.2 Vermeidungslernen

Die operante Konditionierung beschreibt eine Form des Lernens, bei welcher die Auftretenswahrscheinlichkeit von Verhalten durch die Konsequenzen beeinflusst wird. Durch assoziative Verknüpfungen von Reizen und Reaktionen kann ein Verhalten verstärkt oder auch abgeschwächt werden und sich in seiner Häufigkeit in Abhängigkeit der erfahrenen Konsequenzen verändern. Im Rahmen empirischer Untersuchungen von Quellenidentifikations- und Itemrekognitionsleistung wurden bereits Aspekte der positiven Verstärkung bzw. Belohnung für Kinder und auch Erwachsene verwendet (z.B. Roebers et al., 2001). Korrekte Antworten wurden mit Hilfe von Token Systemen

positiv verstärkt, um damit die Zuverlässigkeitsmotivation der Aussagenden zu beeinflussen (Roebbers & Fernandez, 2002; Roebbers et al., 2001; Roebbers & Schneider, 2005). Es fehlen jedoch Studien, im Rahmen derer die Wirksamkeit von Bestrafung übergreifend für free und forced report-Optionen untersucht wird.

In der vorliegenden Forschungsarbeit wird das Konzept der Bestrafung bzw. des Vermeidungslernens in den *Strafe*-Bedingungen verwendet. Es wird davon ausgegangen, dass durch die Bestrafung einer falschen Antwort eine Verschiebung des Antwortkriteriums bzw. der Schwelle (siehe Kapitel 1.1.2 „Source Monitoring – Exkurs: Multinomiale Modelle“) stattfindet. Entgegen der Studien von Roebbers und Kollegen (2001; 2002; 2005), wird hier eine vollständige Kreuzung mit einem freien und einem forcier-ten Antwortformat vorgenommen.

B EMPIRIE

5 Fragestellung

Die vorliegende Forschungsarbeit geht der Frage nach, wie unterschiedliche Variationen der Zuverlässigkeitsmotivation auf die Quellenidentifikationsleistung von Grundschulkindern wirken. Dazu wurde eine Manipulation der Zuverlässigkeitsmotivation kindlicher Aussagen über zwei unterschiedliche Interventionen vorgenommen: zum einen über die Möglichkeit der Zurückhaltung einer Antwort bei Ungewissheit und zum anderen über die Bestrafung falscher Antworten hinsichtlich der Quellenidentifikation. Als abhängige Variablen wurden die Itemrekognition (Sensitivitätsmaß d') und die Quellenidentifikation (u.a. (A)CSIM) sowie die Anzahl der „Ich weiß nicht“-Antworten erhoben. Des Weiteren wurde vor der Quellenidentifikationsaufgabe der Recall, d.h. die ungeleitete Erinnerungsleistung im freien Bericht, erhoben und hinsichtlich qualitativer Dimensionen ausgewertet.

Für die Untersuchung wurde angenommen, dass die Bestrafung falscher Antworten zu komplementären Effekten im Vergleich zur Belohnung richtiger Antworten führt. Die Strafe wurde über den Entzug von Token operationalisiert. Ferner wurde die Annahme getroffen, dass durch die Variation der Faktoren FORCIERUNG und STRAFE eine Verschiebung des Antwortkriteriums (Hochschwelle) bei den Teilnehmern bewirkt wird und Prozesse des Gedächtnismonitoring i.S.v. Koriat und Goldsmith (1994, 1996) angestoßen werden. Den negativen Folgen von falschen oder unsicheren Quellenidentifikationsurteilen von kindlichen Zeugen werden im Kontext der Aussagepsychologischen Forschung immense Bedeutung zuteil. Weiterhin kann davon ausgegangen werden, dass es in Interaktionssituationen zwischen Interviewern und kindlichen Zeugen zu einem Erleben von Strafe durch die Kinder kommt. Zusammen genommen bedürfen diese Aspekte einer experimentellen Untersuchung vor dem Hintergrund gedächtnispsychologischer Erkenntnisse. Ein Alleinstellungsmerkmal der vorliegenden Arbeit ist die Präsentationsmodalität des Lernmaterials und die vollständige Randomisierung der Faktorstufen über die vier Bedingungen. Es handelt sich um live präsentierte und somit von den Teilnehmern erlebte Geschichten. Entgegen dem Großteil empirischer Studien, welche oftmals eine subsequente Testung an die Lernphase anschließen, wurde hier eine Latenz von sieben Tagen zwischen Lernsituation und Abruf angesetzt.

Hinsichtlich der quantitativen Dimension wurde angenommen, dass die Anzahl falscher Antworten in der Bestrafungsbedingung niedriger ausfällt, während die Anzahl der *IWN*-Antworten ansteigt. Für die qualitativen Outcomes konnten keine theoretisch begründeten Vorannahmen getroffen werden, so dass hier eine explorative Untersuchung der Daten erfolgte.

6 Methode

Die Bedingungsvariation bezieht sich auf die beiden unabhängigen Variablen *Befragungssituation* und *Zuverlässigkeitsmotivation*, welche jeweils auf zwei Stufen variiert wurden, womit ein 2x2-faktorielles Design (Faktor FORCIERUNG und Faktor STRAFE) resultiert. Daraus ergaben sich die vier Versuchsbedingungen: *Forcierung-/Strafe-*, *Forcierung-/Strafe+*, *Forcierung+/Strafe-* und *Forcierung+/Strafe+* (siehe Tabelle 6), dabei indiziert ein Minuszeichen (-) die Abwesenheit und ein Pluszeichen (+) die Anwesenheit der Faktorvariation. Vor der gesteuerten Abfrage mit Hilfe der QI-Bögen (siehe Anhang F) wurden die Teilnehmer im *freien Bericht* zur spontanen Reproduktion angeregt (Balloff, 2004). Die Erinnerung an das Versuchsmaterial wurde durch die reine Nachfrage („Du hast letzte Woche zwei Geschichten von zwei Personen gehört, kannst du dich daran erinnern?“) angeregt und während den Ausführungen durch nicht-selektive Nachfragen verstärkt („Was kam noch in den Geschichten vor?“).

6.1 Versuchsdesign

Die unabhängige Variable (UV) Befragungssituation wurde über die Möglichkeit zur Zurückhaltung einer Antwort operationalisiert, welche den Teilnehmern wiederholt offeriert wurde (*free*: *Forcierung-/Strafe-* und *Forcierung-/Strafe+*). Es wurde angenommen, dass diese Option bessere Leistungsergebnisse in der Itemerkennung und Quellenidentifikation zur Folge hat (Roebbers et al., 2001). Das Versuchsdesign ist zur Übersicht in Tabelle 6 dargestellt.

Zur Kontrolle der Bedingungsmanipulation wurde ein Vergleich der forcierungsfreien *free*-Bedingungen mit den *forced*-Bedingungen (*Forcierung+/Strafe-* und *Forcierung+/Strafe+*) realisiert. Die UV Zuverlässigkeitsmotivation wurde in jeweils einer *free*- und einer *forced*-Bedingung über die Bestrafungskomponente bzw. das Vermeidungslernen und über den Entzug von Token operationalisiert.

Tabelle 6. Versuchsdesign.

		STRAFE		
		-	+	
FORCIERUNG	-	<i>Forcierung- /Strafe-</i>	<i>Forcierung- /Strafe+</i>	free
	+	<i>Forcierung+ /Strafe-</i>	<i>Forcierung+ /Strafe+</i>	forced
		Rückmeldung	Strafe	

Anmerkung: Faktoren FORCIERUNG (*free*: mit *IWN*-Antwortoption; *forced*: ohne *IWN*-Antwortoption) und STRAFE (*Rückmeldung* ohne Strafe; *Strafe* mit Strafe) und vier Versuchsbedingungen.

Als abhängige Variablen der Itemerkennung wurden die *Qualität der Itemerkennung*, die *Qualität der Quellenidentifikationsleistung*, die Quantität der *IWN*-Antworten zur Itemerkennung und die Quantität der *IWN*-Antworten bei der Quellenidentifikation untersucht. Um die Qualität der Itemerkennungsleistung zu bestimmen, wurde zunächst die Anzahl der richtig als alt und richtig als neu erkannten Items ausgewertet. Die Anzahl der richtig als alt erkannten und richtig als neu zurückgewiesenen Items wurde in den *free*-Bedingungen an der Gesamtzahl der beantworteten Items relativiert. Die Gesamtanzahl der Items lag bei 28, die Relativierung erfolgte demnach an der Anzahl der Items zur Itemwiedererkennung abzüglich der gegebenen Ich weiß nicht Antworten. Dieser Index wurde entwickelt, da gegenüber Zeugen die Erwartung besteht, dass möglichst viele richtige und wenig falsche Antworten gegeben werden (Greuel, 2001). Dieses Maß wird nachfolgend als *Qualität der Itemerkennung* bezeichnet (siehe Tabelle 7).

Die in Tabelle 7 aufgeführten und in der Analyse verwendeten Indizes wurden von der Autorin in Anlehnung an die klassischen Maße des Source Monitorings, *HR*, *FAR*, *I*, *d'* und *ACSIM selbst* entwickelt.

Eine Übersicht der klassischen Variablen findet sich in Kapitel 1.1.2 „Source Monitoring“, in Tabelle 2. Im Folgenden wird die Herleitung der selbst entwickelten Indizes dargestellt und erläutert. Dabei sind die Indizes (6), (7), (8) und (9) quantitative Maße und die Indizes (10) und (11) qualitative Maße. Eine Übersicht ist in Tabelle 7 dargestellt, welche Indizes der Itemerkennung und Quellenidentifikation zusammengefasst für die *free*- und *forced*-Bedingungen aufführt.

Zur Ermittlung der Itemerkennungsleistung in den *free*-Bedingungen *Forcierung-/Strafe-* und *Forcierung-/Strafe+* wurde die Anzahl der richtig als alt erkannten Items und richtig als neu zurückgewiesenen Items an der Gesamtzahl der beantworteten Items zur Wiedererkennung relativiert, wobei hier die Anzahl der mit *IWN* beantworteten Items die Grundgesamtheit von 28 verändert.

$$Q_{IR_free} = \frac{(X_{AA} + X_{BB} + X_{NN} + X_{AB} + X_{BA})}{(28 - (X_{AIW} + X_{BIW} + X_{NIW}))}$$

(6) Qualität der Itemerkennung für die *free*-Bedingungen

Um die Itemerkennungsleistung in den *forced*-Bedingungen zu ermitteln, wurde die Anzahl der richtig als alt erkannten und richtig als neu zurückgewiesenen Items an der Gesamtzahl der Items zur Wiedererkennung relativiert.

$$Q_{IR_forced} = \frac{X_{AA} + X_{BB} + X_{NN} + X_{AB} + X_{BA}}{28}$$

(7) Qualität der Itemerkennung für die *forced*-Bedingungen

Um die Qualität der Quellenidentifikationsleistung zu bestimmen, erfolgte ein ähnliches Vorgehen. In den *free*-Bedingungen wurde die Anzahl der richtig den beiden Quellen A und B zugeordneten und richtig als neu zurückgewiesenen Items an der Gesamtzahl der beantworteten Items relativiert:

$$Q_{QL_free} = \frac{X_{AA} + X_{BB} + X_{NN}}{(28 - (X_{IAW} + X_{IWB} + X_{IWIW}))}$$

(8) Qualität der Quellenidentifikation für die *free*-Bedingungen

Dieses Vorgehen wurde gewählt, da die Teilnehmer die Frage nach der Quellenidentifikation eindeutig beantworteten konnten und eine Vermischung mit der Itemwiedererkennung ausgeschlossen werden konnte. Die Häufigkeiten der *forced*-Bedingungen wurden lediglich an der Gesamtzahl der Items zur Quellenidentifikation relativiert.

$$Q_{QI_forced} = \frac{X_{AA} + X_{BB} + X_{NN}}{28}$$

(9) Qualität der Quellenidentifikation für die *forced*-Bedingungen

Zur Bestimmung der Quantität der *IWN*-Antworten für die Itemrekognition wurden die *IWN*-Antworten der Items für Quelle A, Quelle B und der Distraktoritems aufsummiert.

$$X_{AIW} + X_{BIW} + X_{NIW}$$

(10) Missings der Itemrekognition

Die Ermittlung der Quantität der „*Ich weiß nicht*“ Antworten der Quellenidentifikation erfolgte analog über die Addition der „*Ich weiß nicht*“ Antworten bei Fragen zur Quellenidentifikation.

$$X_{IAW} + X_{IWB} + X_{IWIW}$$

(11) *IWN*-Antworten der Quellenidentifikation

Neben den oben aufgeführten Indizes, welche an die vorliegende Untersuchung angepasst sind, werden auch die klassischen Maße d' aus der Signalentdeckungstheorie und *ACSIM* in den Ergebnissen berichtet.

Tabelle 7. Berechnungsformeln der verwendeten Indizes.

<i>free</i>	<i>forced</i>
(6) Qualität der Itemerkennung	(7) Qualität der Itemerkennung
$Q_{IR_free} = \frac{(X_{AA} + X_{BB} + X_{NN} + X_{AB} + X_{BA})}{(28 - (X_{AIW} + X_{BIW} + X_{NIW}))}$	$Q_{IR_forced} = \frac{X_{AA} + X_{BB} + X_{NN} + X_{AB} + X_{BA}}{28}$
(8) Qualität der Quellenidentifikation	(9) Qualität der Quellenidentifikation
$Q_{QI_free} = \frac{X_{AA} + X_{BB} + X_{NN}}{(28 - (X_{IAW} + X_{IWB} + X_{IWIW}))}$	$Q_{QI_forced} = \frac{X_{AA} + X_{BB} + X_{NN}}{28}$
(10) Missings Itemerkennung	(11) <i>IWN</i> -Antworten Quellenidentifikation
$X_{AIW} + X_{BIW} + X_{NIW}$	$X_{IAW} + X_{IWB} + X_{IWIW}$

Anmerkung: *free*-Bedingungen: Forcierung-/Strafe- und Forcierung-/Strafe+; *forced*-Bedingungen: Forcierung+/Strafe- und Forcierung+/Strafe+.

6.2 Hypothesen

In Anlehnung an die Ergebnisse zur *IWN*-Antwortoption, Strafe und die im vorangehenden Kapitel dargestellten unabhängigen und abhängigen Variablen wurden die nachfolgenden Hypothesen aufgestellt. Zur Übersicht der Indizes siehe Kapitel 6.1 „Versuchsdesign“.

Es wird angenommen, dass sich die Möglichkeit zur Zurückhaltung einer Antwort (*IWN* bzw. FORCIERUNG) positiv auf die Itemerkennungs- und Quellenidentifikationsleistung auswirkt. Des Weiteren wird angenommen, dass die Bestrafung falscher Antworten einen positiven Effekt auf die Itemerkennungs- und Quellenidentifikationsleistung hat. Zusammenfassend wird also ein Anstieg der Leistungen über die vier Bedingungen angenommen.

Hypothesen der Itemerkennung

Zur Überprüfung der Hypothesen der Itemerkennung werden die Auswirkungen der unabhängigen Variablen auf die Indizes *Qualität der Itemerkennung* und d' betrachtet. Die angenommenen Effekte werden im Folgenden dargestellt.

H1: Haupteffekt FORCIERUNG: In den *free*-Bedingungen resultieren qualitativ bessere Ergebnisse der Itemerkennungsleistung als in den *forced*-Bedingungen.

H2: Haupteffekt STRAFE: In den *Strafe*-Bedingungen resultieren qualitativ bessere Ergebnisse der Itemerkennungsleistung als in den *Rückmeldung*-Bedingungen.

Hypothesen der Quellenidentifikation

Zur Überprüfung der Hypothesen der Quellenidentifikation werden die Auswirkungen der unabhängigen Variablen auf die Indizes *Qualität der Quellenidentifikation* und *ACSIM* betrachtet. Die angenommenen Effekte werden im Folgenden dargestellt.

H3: Haupteffekt FORCIERUNG: In den *free*-Bedingungen resultieren qualitativ bessere Ergebnisse der Quellenidentifikationsleistung als in den *forced*-Bedingungen.

H4: Haupteffekt STRAFE: In den *Strafe*-Bedingungen resultieren qualitativ bessere Ergebnisse der Quellenidentifikationsleistung als in den *Rückmeldung*-Bedingungen.

Annahmen zu einer Interaktion der genannten Faktoren konnten nicht aus der bestehenden Literatur abgeleitet werden. Im nachfolgenden Abschnitt werden zunächst die Stichprobe sowie der Ablauf der Untersuchung vorgestellt. Anschließend werden die deskriptive Statistik, die Ergebnisse der statistischen Analysen und die Auswertung der erhobenen Interviews im Rahmen der qualitativen Inhaltsanalyse dargestellt.

6.3 Stichprobe, Material und Untersuchungsablauf

Die Akquise erfolgte im Vorfeld per Telefon und E-Mail. Die Teilnehmer stammten aus drei Kölner Grundschulen in den Stadtteilen Rodenkirchen ($N = 38$), Merheim ($N = 45$) und Mülheim ($N = 21$). Alle Kinder waren Schüler der vierten Jahrgangsstufe. Von den 105 Teilnehmern waren 54 weiblich, der Altersrange lag bei neun bis elf Jahren ($M = 9,81$; $SD = 0,54$), der Modus bei zehn Jahren (siehe Abbildung 4).

Vor der Untersuchung wurde das Einverständnis der Sorgeberechtigten mit einem Informationsschreiben und beigefügtem Rückgabeabschnitt eingeholt (siehe Anhang B). Die Teilnehmer wurden durch Reihenfolgen-Randomisierung gleichverteilt und zufällig den vier Untersuchungsbedingungen *Forcierung-/Strafe-* ($N = 27$), *Forcierung-/Strafe+* ($N = 26$), *Forcierung+/Strafe-* ($N = 26$), *Forcierung+/Strafe+* ($N = 26$) zugeordnet. Dabei wurde sowohl innerhalb der Schule als auch zwischen den Bedingungen eine Variation vorgenommen.

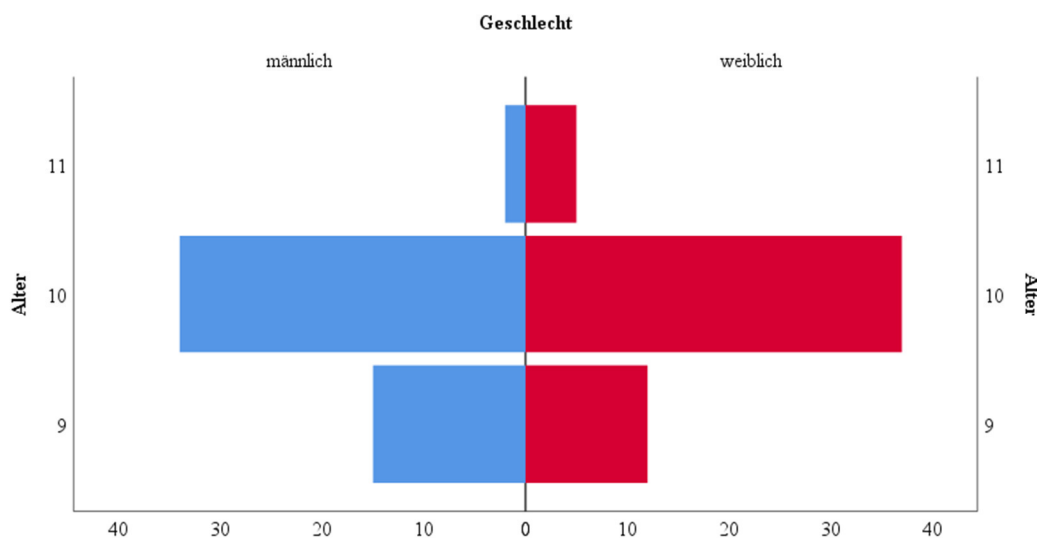


Abbildung 4. Alters- und Geschlechtsverteilung.

Die Antworten aus den QI-Fragebögen (siehe Anhang F) wurden vor der statistischen Analyse entsprechend ihrer Bedingung in einen von zwei Auswertungsbögen übertragen (siehe Abbildung 5 und 6), was der Vorbereitung der Dateneingabe in SPSS diente. Die hier verwendeten Variablen sind an die Antwortmatrix des klassischen QI-Paradigmas angelehnt. Zu beachten ist hinsichtlich des Auswertungsbogens für die *free*-Bedingungen (Abbildung 5), dass die *IWN*-Antworten im tabellarischen Schema nicht den Quellen zuordenbar sind.

		Quelle			Ich weiß nicht
		Quelle 1A	Quelle 2B	Neu	
Antwort	Quelle 1A	<i>Treffer</i> X_{AA}	<i>Fehler</i> X_{BA}	<i>False Alarm</i> X_{NA}	<i>Treffer</i> $IR_{korrekt}; QI_{IWN}$ X_{IWA}
	Quelle 2B	<i>Fehler</i> X_{AB}	<i>Treffer</i> X_{BB}	<i>False Alarm</i> X_{NB}	<i>Treffer</i> $IR_{korrekt}; QI_{IWN}$ X_{IWB}
	Neu	<i>Miss</i> X_{AN}	<i>Miss</i> X_{BN}	<i>Treffer/ Correct</i> <i>Rejection</i> X_{NN}	$IR_{falsch}; QI_{IWN}$ X_{IWIW}
	Ich weiß nicht	<i>Miss</i> X_{AIW}	<i>Miss</i> X_{BIW}	<i>Miss</i> X_{NIW}	
SUMME					

Abbildung 5. Auswertungsbogen 1 für die *free*-Bedingungen.

Der Auswertungsbogen für die *forced*-Bedingungen ergänzt die Antwortmatrix um die Antwortkategorien „Ich weiß nicht“ und „Geraten“. Dabei steht die Kategorie „Geraten“ hier sowohl für *IWN*-Antworten hinsichtlich der Itemerkennung als auch für die Quellenidentifikation (siehe Abbildung 6).

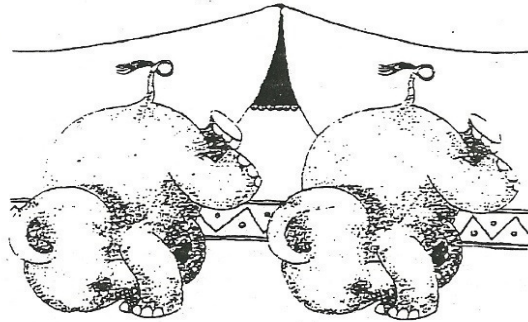
		Quelle				
		Quelle 1A	Quelle 2B	Neu	Geraten	
					Item- erkennung	Quellen- identifikation
Antwort	Quelle 1A	<i>Treffer</i> Y _{AA}	<i>Fehler</i> Y _{BA}	<i>False Alarm</i> Y _{NA}	G _{IRa}	G _{Q1a}
	Quelle 2B	<i>Fehler</i> Y _{AB}	<i>Treffer</i> Y _{BB}	<i>False Alarm</i> Y _{NB}		
	Neu	<i>Miss</i> Y _{AN}	<i>Miss</i> Y _{BN}	<i>Treffer/ Correct Rejection</i> Y _{NN}	G _{IRn}	/
SUMME						

Abbildung 6. Auswertungsbogen 2 für die *forced*-Bedingungen.

Die Teilnehmer saßen in Kleingruppen von je zehn bis zwölf Personen in einen für den Versuch vorbereiteten Raum (abgedunkelt, Stühle im Halbkreis angeordnet) der Schule. Dort warteten zwei Vorleserinnen, die einzeln nacheinander vor die Teilnehmer traten und das Stimulusmaterial, jeweils eine circa zehnminütige Zirkusgeschichte frei sprechend vortrugen. Beide Geschichten enthielten einen identischen einleitenden Text, zehn Zirkusnummern mit ähnlichen, aber unterscheidbaren Inhalten und einen identischen abschließenden Textteil. Die einzelnen Zirkusattraktionen wurden in Sequenzen von ein bis fünf Sätzen umschrieben. Die Originalversionen der Zirkusgeschichten A und B sind im Anhang C aufgeführt. Um das verbal-auditive Stimulusmaterial visuell zu verstärken und die Gedächtnisleistung der Kinder zu unterstützen, wurden die Zirkusnummern auf entsprechenden schwarz-weiß Illustrationen gezeigt (siehe Abbildung 7). Anhang H zeigt diese originär im DIN-A3 Format präsentierten Bilder und die jeweils entsprechende Textstelle. Die Präsentation der Geschichten

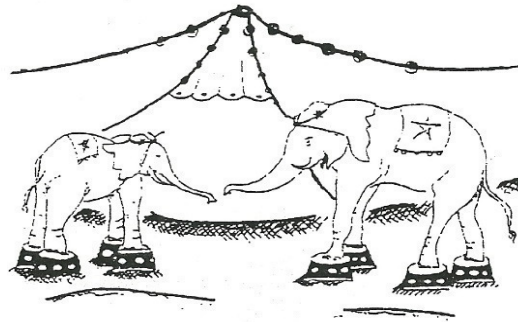
wurde auf Video aufgezeichnet, um eventuelle Zwischenfälle und damit mögliche Verzerrungen der Situation und des Materials kontrollieren zu können.

Geschichte A



Zwei kleine Elefanten machten Handstand

Geschichte B



Elefanten liefen über Podeste

Abbildung 7. Beispielillustrationen aus Geschichte A und B.

Sieben Tage nach der Lernphase wurden die Teilnehmer einzeln zu den Zirkusgeschichten von drei geschulten und den Kindern unbekanntem Versuchsleiterinnen in separaten Räumen interviewt. Zur Herstellung eines Rapports wurden die Befragungen mit der Vorführung der Funktionsweise des Diktaphons begonnen, bei welchem die Versuchsleiterinnen das Datum nannten und die Teilnehmer auf Nachfrage ihren Namen nennen sollten. Diese Aufnahme wurde vor der eigentlichen Befragung gemeinsam angehört. Eine offene und ungezwungene Stimmung in der Interviewsituation fördert die Wiedergabe eines freien Berichtes (Willis, 2005). Der erste Teil der Untersuchung beinhaltete eben diesen *freien Bericht* als Teil des kognitiven Interviews (siehe Abschnitt 1.3.2 Das kognitive Interview). Es wurde einleitend daran erinnert, dass eine Woche zuvor zwei Geschichten vorgelesen wurden. Die Teilnehmer wurden gebeten, all ihre Erinnerungen an beide Zirkusgeschichten wiederzugeben. Es wurde darauf hingewiesen, dass Reihenfolge und Vollständigkeit keine Rolle spielen. Der Bericht wurde durch die Interviewerin nicht-selektiv verstärkt (Kopfnicken und „Mmh“). Bei einer längeren Pause der Teilnehmer wurde zunächst 90 Sekunden abgewartet, bis wiederholt darauf hingewiesen wurde, dass alles berichtet werden soll, was erinnert wird. Die Instruktionen des halb-standardisierten Interviewleitfadens für die Rapportbildung und für den freien Bericht sind Anhang D zu entnehmen. Nach Beendigung

des freien Berichts, durch Hinweis der Teilnehmer oder auf Nachfrage der Versuchsleiterin wurde das Interview mit einem von sechzehn voll-standardisierten Quellenidentifikationsfragebögen (siehe Anhang F) fortgesetzt. Diese strukturierte Befragung zielte auf die korrekte Rekognition von 14 tatsächlichen (sieben Items je Zirkusgeschichte A und B) und die Zurückweisung von 14 falschen Distraktoritems. Pro Bedingung gab es je vier verschiedene Bögen, für welche die Reihenfolge der Items randomisiert wurde, womit insgesamt 16 unterschiedliche QI-Bögen Verwendung fanden. Zunächst wurden die Teilnehmer entsprechend dem vorab festgelegten Untersuchungsplan, instruiert. Hier wurde auf die *IWN*-Antwortoption ohne (*Forcierung-/Strafe-*) oder mit Strafe (*Forcierung-/Strafe+*) hingewiesen. Die Möglichkeit zur „Ich weiß nicht“ Antwort wurde jeweils nach drei Items wiederholt. In den *forced*-Bedingungen wurde entsprechend darauf hingewiesen, dass auch bei Unsicherheit oder Nichtwissen eine Antwort gegeben werden muss.

Nur in den *Strafe*-Bedingungen erhielten die Teilnehmer 56 bunte Büroklammern (28 pro Rekognitions-/ Distraktoritem und 28 pro QI- Items) in einem durchsichtigen Behälter mit Deckel, welche als Token dienten. Diese wurden den Teilnehmern während der Instruktion übergeben. Zu jedem Rekognitionsitem wurde die entsprechende Quelle bzw. die jeweilige Vorleserin erfragt. Dafür wurden den Teilnehmern zwei Fotos von Vorleserin 1 und 2 vorgelegt. Bei einer korrekten Zurückweisung eines Items als „neu“ oder einer „*Ich weiß nicht*“ Antwort in den *free*-Bedingungen, wurde die Frage nach der Quellenidentifikation ausgelassen. Die Interviewerinnen markierten die Antworten während der Befragung in einem der QI-Bögen, der durch eine farbliche Markierung, welche die Teilnehmer nicht kannten, die Versuchsbedingung indizierte. In den *Rückmeldung*-Bedingungen erfolgte kein Entzug von Token für eine falsche Antwort, vom Versuchsleiter wurde jedoch zurückgemeldet, dass die Antwort nicht richtig war (siehe Anhang D - Unterpunkt 11.).

Am Ende der Befragung sollten die Teilnehmer angeben, wie oft sie bereits im Zirkus gewesen sind, um später mögliche Konfundierungen und Einflüsse hinsichtlich kognitiver Schemata zu prüfen. Dies wurde auf den QI-Bögen numerisch erfasst. Des Weiteren wurde erfragt, welche der Vorleserinnen sympathischer war bzw. wen sie besser fanden.

Zur Vermeidung von Verzerrungen und Versuchsleitereffekten erhielten die Vorleserinnen und Interviewerinnen vorab ein Training im Hinblick auf Wortwahl, Betonung, Gestik und Mimik im Versuchs- und Interviewverlauf. Die zwei Vorleserinnen trainierten unter Anleitung die Präsentation der Geschichten mit Hilfe der Bilder, bis sie die Geschichten frei vortragen konnten. Die drei Interviewerinnen trainierten den Umgang mit dem Diktaphon, studierten den Leitfaden und übten jeweils an mehreren Laien-Schauspielern das Führen eines Interviews in den jeweiligen Versuchsbedingungen. Im Detail wurde auch der Versuchs- und Ablaufplan der Untersuchung besprochen, damit die randomisierte Zuteilung der einzelnen Schüler zu den jeweiligen Versuchsbedingungen korrekt erfolgte und auch die Reihenfolge und wechselnde Präsentation der Geschichten durch die Vorleserinnen plangemäß ablief. Des Weiteren wurden vorab standardisierte Antworten auf mögliche Rückfragen formuliert (siehe Anhang E). Darüber hinaus wurde auf einheitliche und unauffällige Bekleidung geachtet: Alle Beteiligten waren schwarz gekleidet. Zur Vorbeugung von Reihenfolge- und Kombinationseffekten (Vorleserin 1/2 + Geschichte A/B) sowie hinsichtlich der Präsentationsreihenfolge der Geschichte wurde wie folgt über die Schule und die Kleingruppen randomisiert: An der ersten Schule trug Vorleserin 1 durchgängig Geschichte A und Vorleserin 2 Geschichte B vor. An der zweiten Schule tauschten die Erzählerinnen die Geschichten, an der dritten Schule wurde wie an der ersten verfahren. Innerhalb der Kleingruppen in den Schulen startete zunächst Vorleserin 1 mit Geschichte A und Vorleserin 2 mit Geschichte B folgte. Die nächste Gruppe hörte erst Geschichte B und anschließend Geschichte A. Die Geschichtsreihenfolge der dritten Gruppe entsprach wiederum der ersten Abfolge. Die Abwechslung der Geschichten erfolgte so fortlaufend. Die Kleingruppengröße lag bei maximal zwölf Teilnehmern, so dass eine ruhige Atmosphäre und ausreichende Seh- und Hörverhältnisse im Hinblick auf die Vorleserinnen und Illustrationen sichergestellt werden konnte. Zudem sollte eine Störung der Teilnehmer durch Gespräche untereinander vermieden werden. Des Weiteren wurden Maßnahmen zur Minimierung möglicher Störvariablen und Kontexteffekte getroffen. Hierzu wurden z.B. ähnliche Rahmenbedingungen in allen drei Schulen geschaffen (Stuhlkreise, Raumverdunkelung) und die Nutzung eines gesonderten Raumes für die Versuchsdurchführung sowie das spätere Interview realisiert.

7 Ergebnisse

Im folgenden Kapitel werden zunächst die quantitativen und anschließend die qualitativen Ergebnisse der Untersuchung vorgestellt. Dazu werden erst die deskriptiven Berechnungen dargelegt und darauffolgend die inferenzstatistische Überprüfung der Hypothesen. Abschließend werden weiterführende Analysen präsentiert. Die statistische Auswertung der Daten erfolgte anhand des Programms IBM SPSS Statistics (Version 25). Die qualitative Datenanalyse wurde mit dem Verfahren der qualitativen Inhaltsanalyse nach Mayring (1991) durchgeführt.

7.1 Deskriptive Statistik

Vor der inferenzstatistischen Überprüfung der aufgestellten Hypothesen erfolgt eine deskriptive Betrachtung der Daten. Im Folgenden sind die Treffer, d.h. *Hits* und *Correct Rejections*, getrennt für die Itemerkennung und Quellenidentifikation über alle vier Bedingungen dargestellt. Anschließend erfolgt die Betrachtung der *IWN*-Antworten.

7.1.1 Itemerkennung

Im Mittel resultieren $M = 18,59$ ($SD = 3,816$) Hits und Correct Rejections in der Bedingung *Forcierung-/Strafe-*, welche von insgesamt 28 Items richtig als alt erkannt bzw. als neu zurückgewiesen wurden. In der Bedingung *Forcierung-/Strafe+* wurden $M = 18,88$ ($SD = 3,374$) Items richtig identifiziert und korrekt zurückgewiesen. In der Bedingung *Forcierung+/Strafe-* wurden im Mittel $M = 22,04$ ($SD = 3,194$) und in der Bedingung *Forcierung+/Strafe+* $M = 22,54$ ($SD = 2,121$) Items korrekt als alt bzw. neu erkannt. Über alle vier Gruppen wurden im Mittel $M = 20,5$ ($SD = 3,624$) Items richtig als alt erkannt bzw. als neu zurückgewiesen (siehe Tabelle 8).

Tabelle 8. Mittlere Anzahl beantworteter Items in den vier Versuchsbedingungen.

	<i>Forcierung- /Strafe-</i>	<i>Forcierung- /Strafe+</i>	<i>Forcierung+ /Strafe-</i>	<i>Forcierung+ /Strafe+</i>	Gesamt
IR	18,59 (3,816)	18,88 (3,374)	22,04 (3,194)	22,54 (2,121)	20,5 (3,624)
QI	14,81 (3,648)	15,15 (3,209)	17,69 (3,642)	18,04 (2,69)	16,41 (3,586)

Anmerkung: Mittelwert und Standardabweichung in Klammern für Itemerkennung (IR) und Quellenidentifikation (QI).

Über die vier Bedingungen ergab sich ein Anstieg hinsichtlich des Index *Qualität der Itemerkennung* (siehe S. 18 Abschnitt: Abhängige Variablen) mit $M_{Q_IR} = 0,747$ ($SD = 0,115$) für *Forcierung-/Strafe-*, $M_{Q_IR} = 0,762$ ($SD = 0,098$) für *Forcierung-/Strafe+*, $M_{Q_IR} = 0,787$ ($SD = 0,114$) für *Forcierung+/Strafe-* und $M_{Q_IR} = 0,804$ ($SD = 0,075$) für *Forcierung+/Strafe+* (siehe Tabelle 9).

Tabelle 9. Mittlere Qualität der Itemerkennungsleistung in den vier Versuchsbedingungen.

		STRAFE		
		-	+	
FORCIERUNG	-	$M_{Q_IR} = 0,747$ ($SD = 0,115$)	$M_{Q_IR} = 0,762$ ($0,098$)	<i>free</i>
	+	$M_{Q_IR} = 0,787$ ($0,114$)	$M_{Q_IR} = 0,804$ ($0,075$)	<i>forced</i>
		<i>Rückmeldung</i>	<i>Strafe</i>	

Anmerkung: M_{Q_IR} = Mittlere Qualität der Itemerkennung; SD = Standardabweichung in Klammern.

7.1.2 Quellenidentifikation

Hinsichtlich der Maße für die Quellenidentifikation muss beachtet werden, dass eine zweistufige Befragung stattfand (siehe Kapitel 6.3 „Stichprobe, Material und Untersuchungsablauf“). Zuerst wurde die Itemerkennung und bei korrekter Alt-Neu- Unterscheidung die Quellenidentifikation erfragt, womit sich eine variierende Grundgesamtheit für jeden einzelnen Teilnehmer ergab. Für die Quellenidentifikation wurden in der Bedingung *Forcierung-/Strafe-* $M = 14,81$ ($SD = 3,648$) Items korrekt beantwortet. In der Bedingung *Forcierung-/Strafe+* wurden im Mittel $M = 15,15$ ($SD = 3,209$) Quellen richtig identifiziert. In der Bedingung *Forcierung+/Strafe-* wurden im Mittel $M = 17,69$ ($SD = 3,642$) korrekt zugeordnet und $M = 18,04$ ($SD = 2,69$) in der Bedingung *Forcierung+/Strafe+*. Insgesamt wurden $M = 16,41$ ($SD = 3,586$) Quellen richtig identifiziert (siehe Tabelle 8).

In der Bedingung *Forced-/Strafe-* lag der Mittelwert bei $M_{Q_{QI}} = 0,578$ ($0,14$) und in der Bedingung *Forced-/Strafe+* bei $M_{Q_{QI}} = 0,547$ ($0,125$). In der Bedingung *Forced+/Strafe-* lag der Mittelwert bei $M_{Q_{QI}} = 0,632$ ($0,128$) und in der Bedingung *Forced+/Strafe+* bei $M_{Q_{QI}} = 0,644$ ($0,095$). (siehe Tabelle 10).

Tabelle 10. Mittlere Qualität der Quellenidentifikationsleistung in den vier Versuchsbedingungen.

		STRAFE		
		-	+	
FORCIERUNG	-	$M_{Q_{QI}} = 0,578$ ($0,14$)	$M_{Q_{QI}} = 0,574$ ($0,125$)	<i>free</i>
	+	$M_{Q_{QI}} = 0,632$ ($0,128$)	$M_{Q_{QI}} = 0,644$ ($0,095$)	<i>forced</i>
		Rückmeldung	Strafe	

Anmerkung: $M_{Q_{QI}}$ = Mittlere Qualität der Quellenidentifikation; SD = Standardabweichung in Klammern.

7.1.3 IWN-Antwortoption

Für die zwei *free*-Bedingungen (*Forcierung-/Strafe-* und *Forcierung-/Strafe+*) wurde die Nutzung der *IWN*-Antwortoption untersucht, insgesamt ergab sich ein $N = 53$ in der Teilstichprobe. Im Mittel wurden $M_{IWN-IR} = 3,25$ ($SD = 2,616$) der insgesamt 28 Rekognitionsitems mit *IWN* beantwortet. Von den Quellen wurden im Mittel $M_{IWN-QI} = 1,87$ ($SD = 2,184$) Items mit *IWN* beantwortet (siehe Tabelle 11).

Tabelle 11. Verteilung der *IWN*-Antworten.

	<i>N</i>	<i>Max</i>	M_{IWN}	SD_{IWN}
<i>IWN</i> _{IR_Forcierung-/Strafe-}	27	9	3,22	2,806
<i>IWN</i> _{IR_Forcierung-/Strafe+}	26	8	3,27	2,459
<i>IWN</i> _{IR_Gesamt}	53	9	3,25	2,616
<i>IWN</i> _{QI_Forcierung-/Strafe-}	27	10	2,22	2,665
<i>IWN</i> _{QI_Forcierung-/Strafe+}	26	5	1,5	1,503
<i>IWN</i> _{QI_Gesamt}	53	10	1,87	2,184
<i>IWN</i> _{IR/QI_Rückmeldung}	27	/	5,444	4,126
<i>IWN</i> _{IR/QI_Strafe}	25	/	4,844	2,478
<i>IWN</i> _{Gesamt}	53	18	5,113	3,395

Anmerkungen: *IWN*_{IR_Forcierung-/Strafe-} = Anzahl der *IWN*-Antworten für die Itemerkennung mit Rückmeldung. *IWN*_{IR_Forcierung-/Strafe+} = Anzahl der *IWN*-Antworten für die Itemerkennung mit Strafe. *IWN*_{IR_Gesamt} = Anzahl der *IWN*-Antworten für die Itemerkennung zusammengefasst für die *free*-Bedingungen. Die Erläuterung der Indizes ist äquivalent für *QI*. *N* = Anzahl der Nennungen; *Max* = maximale Anzahl der *IWN*-Antworten; M_{IWN} = Mittelwert der *IWN*-Antworten, *SD* = Standardabweichung der *IWN*-Antworten.

7.2 Inferenzstatistische Analyse

Im Rahmen dieser Arbeit wurden in Anlehnung an die klassische Antwortmatrix im Quellenidentifikationsparadigma auf die Untersuchung abgestimmte, abhängige Variablen und entsprechende Berechnungsformeln entwickelt (siehe Kapitel 6). Die Indizes *Qualität der Itemerkennung*, *Qualität der Quellenidentifikation*, *ACSIM*, *d'* und *Quantität der IWN-Antworten gesamt/ Itemerkennung/ Quellenidentifikation* wurden im

Rahmen der Hypothesentestung anhand von Varianzanalysen untersucht. Die Voraussetzung der Normalverteilung wurde mit dem Kolmogorov-Smirnov-Anpassungstest überprüft. Dabei erwiesen sich d' ($p = .200$) und *Qualität der Itemerkennung* ($p = .181$) als normalverteilte Variablen über alle vier Bedingungen, für alle weiteren Indizes war die Voraussetzung nicht gegeben. Aufgrund des Stichprobenumfangs von $N = 105$ Probanden und einer annähernden Gleichverteilung der Werte über die einzelnen Zellen konnte angenommen werden, dass die varianzanalytischen Verfahren robust auf die Verletzung der Voraussetzung reagieren (Eid et al., 2013). Die Gruppenvergleiche wurden mit parametrischen und nicht-parametrischen Verfahren vorgenommen. Alle numerischen Werte und Effektstärken wurden auf die dritte Nachkommastelle gerundet.

7.2.1 Itemerkennung

Im Folgenden werden die Ergebnisse der Itemerkennung getrennt für die zwei Faktoren FORCIERUNG und STRAFE dargestellt. Hierbei werden jeweils die entsprechenden Variablen (Q_{IR} , d' , IWN) berichtet. Dabei werden die IWN -Antworten nur für den Faktor STRAFE untersucht, da diese Antwortoption nur im Rahmen der Versuchsbedingungen mit *Forcierung+* nicht bestand.

7.2.1.1 Untersuchungen des Faktors FORCIERUNG

Im nächsten Abschnitt werden die Ergebnisse der Untersuchung zur Auswirkung des Faktors FORCIERUNG untersucht. Dazu wurden die Indizes der *free*-Bedingungen (Forcierung-/Strafe- und Forcierung-/Strafe+) gegen die der *forced*-Bedingungen (Forcierung+/Strafe- und Forcierung+/Strafe+) getestet.

Qualität der Itemerkennungsleistung. In die Berechnung des Index *Qualität der Itemerkennung* (Q_{IR}) gingen die Treffer (X_{AA} , X_{BB}), Correct Rejections (X_{NN}) und Treffer bzw. korrekten Itemerkognitionen, welche aber gleichzeitig Quellenzuordnungsfehler beinhalteten (X_{AB} , X_{BA}) (siehe Kapitel 6), relativiert an der Gesamtanzahl insgesamt beantworteter Items, getrennt für die *free*- und *forced*-Bedingungen, mit ein. Zur Erläuterung: Da ein zweistufiges Befragungsverfahren verwendet wurde, ergab sich der Antworttyp, bei welchem ein Item zunächst als richtig erkannt wurde, anschließend aber der falschen Quelle zugeordnet wurde. Der Vergleich der *free*- versus *forced*-

Bedingungen bzw. die Untersuchung der Auswirkung des Faktors FORCIERUNG hat gezeigt, dass die Q_{IR} in den *forced*-Bedingungen signifikant besser war, [$F_{(1,103)} = 4,354$; $p = .039$; $\eta^2 = 0,406$] (siehe Tabelle 12).

Tabelle 12. Übersicht zu den Ergebnissen der ANOVA für die Itemerkennung.

		IR			
		<i>df</i>	<i>F</i>	<i>Sig.</i>	η^2
Faktor FORCIERUNG	Q_{IR}	1,103	4,354	.39	0,406
	d'	1,103	0,001	.954	0,002
Faktor STRAFE	Q_{IR}	1,103	0,683	.411	0,006
	d'	1,103	0,001	.511	0,043

Die mittlere Qualität der Itemerkennungsleistung lag für die *free*-Bedingungen zusammengefasst bei $M_{free} = 0,755$ ($SD = 0,106$). In den *forced*-Bedingungen ergab sich im Mittel $M_{forced} = 0,796$ ($SD = 0,103$) (siehe Abbildung 8).

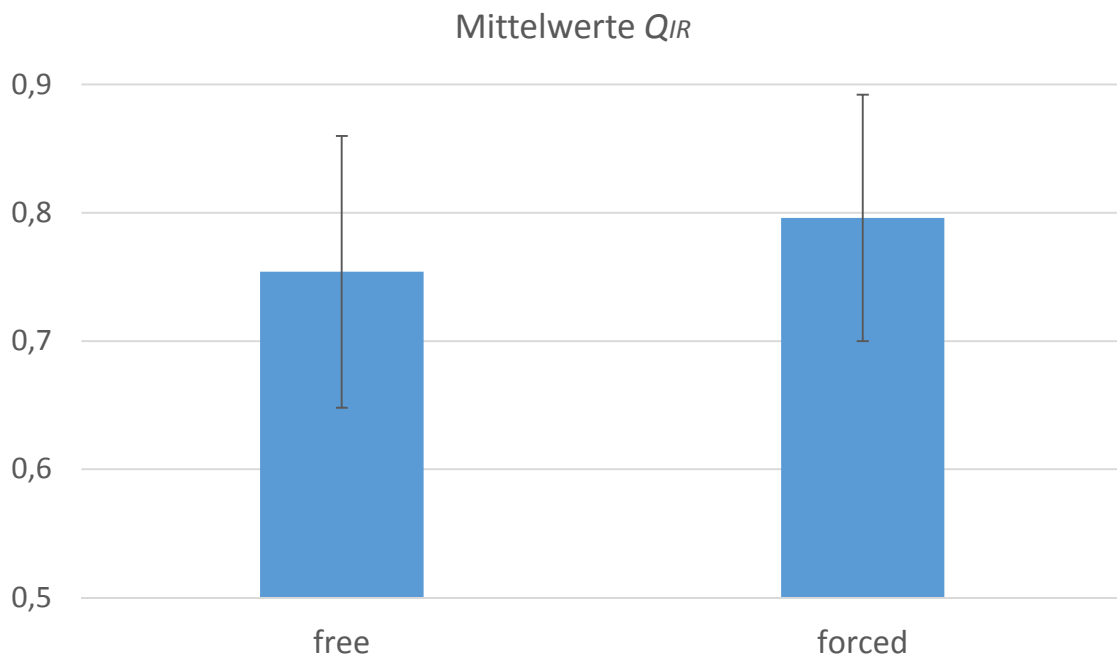


Abbildung 8. Vergleich der mittleren Qualität der Itemerkennungsleistung (Q_{IR}) in den *free*- und *forced*-Bedingungen. Die Fehlerbalken indizieren die jeweilige Standardabweichung.

d' . Der standardisierte Sensitivitätsindex d' wurde aus der Differenz der z-transformierten Hit-Rate und False Alarm Rate berechnet (siehe Kapitel 1.1.2 „Source Monitoring“). Die Varianzanalyse zum Vergleich von d' in Abhängigkeit vom Faktor FORCIERUNG zeigte keinen signifikanten Haupteffekt [$F_{(1,103)} = 0,001$; $p = .994$; $\eta^2 = 0,002$]. Die mittlere Itemrekognitionsleistung d' lag in den *free*-Bedingungen bei $M_{Q_IR_free} = 0,001$ ($SD = 1,396$) versus $M_{Q_IR_forced} = 0,002$ ($SD = 1,446$).

7.2.1.2 Untersuchungen des Faktors STRAFE

Im folgenden Abschnitt werden die Ergebnisse der Untersuchung zur Auswirkung des Faktors STRAFE untersucht. Dazu wurden die Indizes der *Rückmeldung*-Bedingungen (*Forcierung-/Strafe-* und *Forcierung+/Strafe-*) gegen die der *Strafe*-Bedingungen (*Forcierung-/Strafe+* und *Forcierung+/Strafe+*) getestet; kurz Bedingungen 1. & 3. versus 2. & 4.).

Qualität der Itemrekognitionsleistung. Die Untersuchung des Faktors STRAFE ergab keinen signifikanten Haupteffekt [$F_{(1,103)} = 0,683$; $p = .411$; $\eta^2 = 0,006$], die mittlere Q_{IR} war in den *Rückmeldung*-Bedingungen jedoch tendenziell höher als in den *Strafe* Bedingungen ($M_{Rückmeldung} = 0,784$ vs. $M_{Strafe} = 0,767$).

d' . Die Varianzanalyse zum Vergleich der Itemrekognitionsleistung in Abhängigkeit vom Faktor Strafe zeigte keinen signifikanten Haupteffekt für den Quellenidentifikationsindex d' [$F_{(1, 103)} = 0,001$; $p = .511$; $\eta^2 = 0,043$]. Die mittlere Quellenidentifikationsleistung d' lag in den *Rückmeldung*-Bedingungen bei $M_{Rückmeldung} = 0,002$ ($SD = 1,389$) versus $M_{Strafe} = 0,001$ ($SD = 1,453$).

IWN-Antwortoption. Die rein numerische Anzahl der *IWN*-Antworten für die Itemerkennung unterschied sich nicht signifikant in Abhängigkeit vom Faktor STRAFE. Die mittlere Anzahl der *IWN*-Antworten war in der *Strafe*-Bedingung *Forcierung-/Strafe+* ($M_{IWN_IR_Forcierung-/Strafe+} = 3,22$; $SD = 2,806$) nicht signifikant verschieden von der Bedingung *Forcierung-/Strafe-* ($M_{IWN_IR_Forcierung-/Strafe-} = 3,27$; $SD = 2,459$); $t_{(50)} = -0,78$; $p = .938$. Auch der nicht-parametrische Man-Whitney U-Test konnte keine signifikanten Unterschiede zwischen der Anzahl an *IWN*-Antworten in Abhängigkeit vom Faktor STRAFE identifizieren ($U = 323,5$; $p = .795$).

7.2.2 Quellenidentifikation

Im Folgenden werden die Ergebnisse der Quellenidentifikation getrennt für die zwei Faktoren FORCIERUNG und STRAFE dargestellt. Darunter werden jeweils die entsprechenden Variablen (Q_{QI} , $ACSIM$, IWN) aufgeführt. Dabei werden die IWN -Antworten nur für den Faktor STRAFE untersucht, da diese Antwortoption nur im Rahmen der Versuchsbedingungen mit *Forcierung+* nicht bestand.

7.2.2.1 Untersuchungen des Faktors FORCIERUNG

Qualität der Quellenidentifikationsleistung. In die Berechnung des Index *Qualität der Quellenidentifikation* (Q_{QI}) gingen die Treffer (X_{AA} , X_{BB}) und Correct Rejections (X_{NN}) (siehe Kapitel 6.1 „Versuchsdesign“), relativiert an der Gesamtanzahl insgesamt beantworteter Items, getrennt für die *free*- und *forced*-Bedingungen, ein. Die varianzanalytische Untersuchung in Abhängigkeit des Faktors FORCIERUNG zeigte einen signifikanten Haupteffekt [$F_{(1,103)} = 6,673$; $p = .011$; $\eta^2 = 0,061$] (siehe Tabelle 13).

Tabelle 13. Übersicht zu den Ergebnissen der ANOVA für die Quellenidentifikation.

		QI			
		<i>df</i>	<i>F</i>	<i>Sig.</i>	η^2
Faktor FORCIERUNG	Q_{QI}	1,103	6,673	.011	0,062
	$ACSIM$	1,103	0,75	.148	0,2
Faktor STRAFE	Q_{QI}	1,103	0,383	.847	0,061
	$ACSIM$	1,103	4,35	.511	0,004

Die mittlere QI-Leistung liegt bei $M_{Q_{QI_free}} = 0,576$ ($SD = 0,132$) im Vergleich zu $M_{Q_{QI_forced}} = 0,638$ ($SD = 0,638$). Die Ergebnisse sind in Abbildung 9 dargestellt. Zum Vergleich sind die Ergebnisse der Itemerkennung ebenfalls aufgeführt.

ACSIM. Der Index $ACSIM$ wurde berechnet durch die Relativierung der Treffer an der Summe der Treffer sowie Relativierung der fehlerhaften Quellenzuordnungen an der Gesamtzahl der Quellen (siehe Kapitel 1.3.2 „Source Monitoring“). Die Varianzanalyse zum Vergleich der Quellenidentifikationsleistung in Abhängigkeit vom Faktor FORCIERUNG zeigte keinen signifikanten Haupteffekt für den Index der Quellenidentifi-

kation *ACSIM* [$F_{(1, 103)} = 0,75$; $p = .148$; $\eta^2 = 1,89$]. Die mittlere Quellenidentifikationsleistung *ACSIM* lag in den *free*-Bedingungen bei $M_{free} = 0,534$ ($SD = 0,205$) versus $M_{forced} = 0,587$ ($SD = 0,167$).

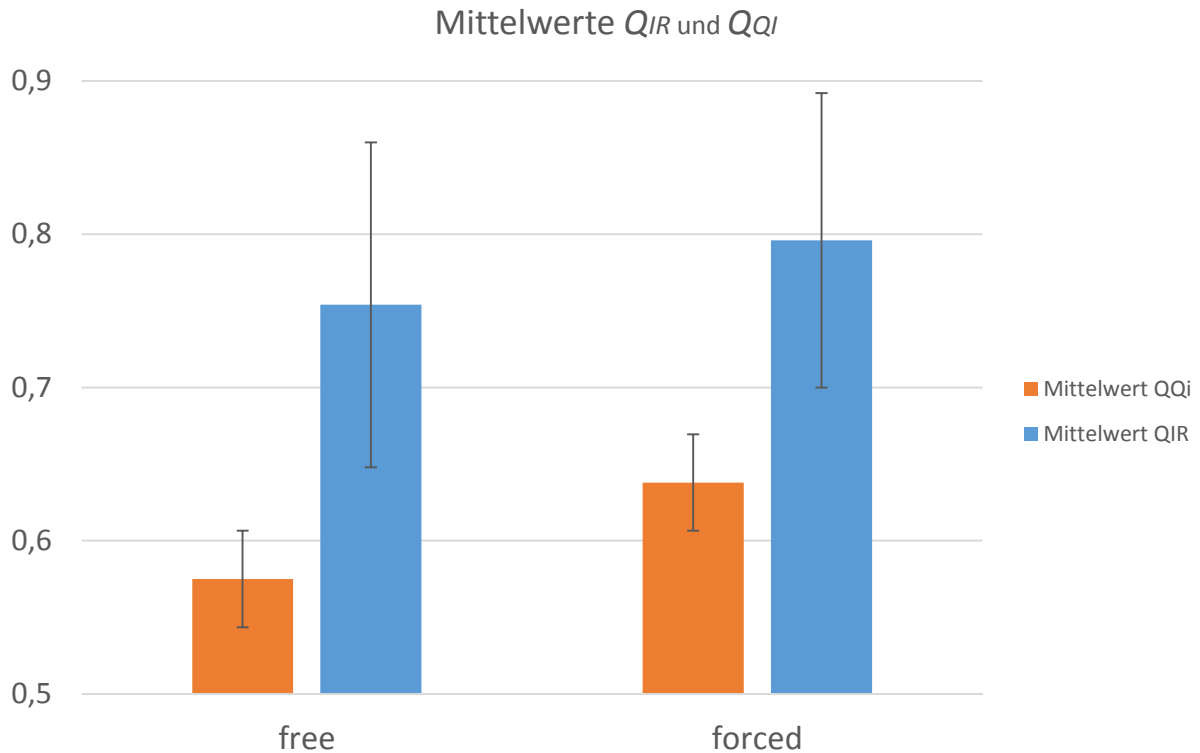


Abbildung 9. Vergleich der mittleren Qualität der Quellenidentifikationsleistung (Q_{QI}) und der Qualität der Itemerkennungslleistung (Q_{IR}) in den *free*- und *forced*-Bedingungen. Die Fehlerbalken indizieren jeweils die Standardabweichung.

7.2.2.2 Untersuchungen des Faktors STRAFE

Qualität der Quellenidentifikationsleistung. Die Qualität der Quellenidentifikation Q_{QI} unterschied sich in Abhängigkeit vom Faktor STRAFE nicht signifikant [$F_{(1,103)} = 0,383$; $p = .847$; $\eta^2 = 0,061$] und ($M_{Rückmeldung} = 0,604$ ($SD = 0,136$) versus $M_{Strafe} = 0,609$ ($SD = 0,117$)).

ACSIM. Die Varianzanalyse zum Vergleich der Quellenidentifikationsleistung in Abhängigkeit vom Faktor STRAFE zeigte keinen signifikanten Haupteffekt für den QI-Index *ACSIM* [$F_{(1, 103)} = 4.35$; $p = .511$; $\eta^2 = 0,004$]. Die mittlere QI-Leistung *ACSIM* lag in den *Rückmeldung*-Bedingungen bei $M_{Rückmeldung} = 0,548$ ($SD = 0,186$) versus $M_{Strafe} = 0,573$ ($SD = 0,192$).

IWN-Antwortoption. Die Anzahl der *IWN*-Antworten für die Quellenidentifikation unterschied sich nicht signifikant in Abhängigkeit vom Faktor STRAFE. Die mittlere Anzahl der *IWN*-Antworten lag in der straffreien Bedingung *Forcierung-/Strafe-* ($M_{IWN_QI_Forcierung-/Strafe-} = 2,22$; $SD = 2,665$) rein deskriptiv höher als in der Bedingung *Forcierung-/Strafe+* ($M_{IWN_IR_Forcierung-/Strafe+} = 1,5$; $SD = 1,503$); $t_{(50)} = 1,092$; $p = .28$). Auch der nicht-parametrische Man-Whitney U-Test konnte keine signifikanten Unterschiede zwischen der Anzahl an *IWN*-Antworten in Abhängigkeit vom Faktor STRAFE identifizieren ($U = 333,5$; $p = .645$). Die mittlere Anzahl der *IWN*-Antworten zusammengefasst für die Itemerkennung und Quellenidentifikation in den Bedingungen *Forcierung-/Strafe-* und *Forcierung-/Strafe+* unterschied sich im t-Test für unabhängige Stichproben nicht signifikant ($t_{(50)} = 0,634$; $p = .529$). Die mittlere Anzahl der *IWN*-Antworten für die Itemerkennung und Quellenidentifikation lag in der Bedingung *Forcierung-/Strafe-* ($M_{IWN_IR/QI_Forcierung-/Strafe-} = 5,444$; $SD = 1,126$) deskriptiv höher als in der Bedingung *Forcierung-/Strafe+* ($M_{IWN_IR/QI_Forcierung-/Strafe+} = 4,484$; $SD = 2,478$). Im nicht-parametrischen Man-Whitney U-Test resultierten keine signifikanten Unterschiede zwischen der Anzahl an *IWN*-Antworten in Abhängigkeit vom Faktor STRAFE ($U = 328,5$; $p = .868$).

7.2.3 Zusammenfassung der inferenzstatistischen Ergebnisse

Insgesamt ergab sich in der vorliegenden Untersuchung ein Haupteffekt für den Faktor FORCIERUNG, allerdings in unerwarteter Richtung: Die Teilnehmer zeigten signifikant bessere Leistungen, wenn sie keine *IWN*-Antwortoption hatten. Die *Qualität der Itemerkennung* und die *Qualität der Quellenidentifikation* waren in den *forced*-Bedingungen signifikant besser als in den *free*-Bedingungen. Für die Maße d' und *ACSIM* ergaben sich keine statistisch bedeutsamen Unterschiede in Abhängigkeit vom Faktor FORCIERUNG. Der Faktor STRAFE hatte keinen signifikanten Effekt auf die erhobenen Maße der Itemerkennungs- oder Quellenidentifikationsleistung, Q_{IR} , Q_{QI} , d' , *ACSIM* und *IWN*-Antworten. Die Nutzungshäufigkeit der *IWN*-Antwortoption war demnach unabhängig von der STRAFE.

7.3 Qualitative Inhaltsanalyse

Zur Untersuchung der freien Berichte wurde ein qualitativ-inhaltsanalytisches Vorgehen im Sinne der von Mayring postulierten qualitativen Inhaltsanalyse gewählt. Nach Mayring (1991, 2015) besteht das Ziel der Inhaltsanalyse darin, protokolliertes Material

einer Kommunikation systematisch, regelgeleitet und anhand einer theoretischen Fragestellung zu analysieren. Ein zentraler Bestandteil qualitativer Inhaltsanalysen ist die Bildung eines Kategoriensystems zur Klassifizierung des Datenmaterials. Dies ermöglicht eine systematische und nachvollziehbare Aussagenanalyse unter Nutzung individuell erstellter Kodierleitfäden (Mayring, 1991; siehe Anhang I bis N). Dieses Vorgehen wurde in Anlehnung an die Analyse und Kategorisierung frei berichteter Erinnerungen, ähnlich zu den CBCA Kriterien, gewählt.

7.3.1 Der freie Bericht im Interview

Die qualitative Inhaltsanalyse der vorliegenden Berichte soll Aussagen über Art, Detailliertheit und Ausmaß bzw. Menge der Erinnerungsleistung ermöglichen. Das Textmaterial lag in qualitativen Analyseeinheiten vor und konnte nach Erfassung der Treffer quantitativ ausgewertet werden. Dabei sollten mögliche Effekte im Zusammenhang mit Vorwissen, Quellengedächtnis und Vorleserin aufgedeckt werden. Unter Berücksichtigung der in Kapitel 3.2 „Kriterienorientierte Inhaltsanalyse“ und 3.4.2 „Der freie Bericht“ dargestellten theoretischen Grundlagen wurden fünf übergeordnete Kategorien definiert: *korrekte Handlungsinhalte*, *Konfabulation*, *Emotion*, *Schemawissen* und *Vorlesermerkmale*. Eine Übersicht des Kategoriensystems ist in Abbildung 10 dargestellt. Jeder Kategorie wurden bestimmte Items als Analyseeinheit zugeordnet und es wurden entsprechende Kodierregeln hergeleitet. Die Itemzuordnungen zu den Illustrationen sind in Anhang H aufgeführt. Als zusätzliche inhaltliche Kategorie wurde die Kategorie *Rahmentext* eingerichtet. Mit dieser sollte erfasst werden, inwieweit die jeweils identischen Eingangs- und Abschlusspassagen der Erzählung erinnert werden. Mit der Kategorie *Emotion* wurden Erinnerungsleistungen an emotional getönte Textpassagen erfasst, dadurch sollte geprüft werden, ob emotionale Handlungsinhalte besser im Gedächtnis gespeichert werden als andere Informationen (Bell, Mieth & Buchner, 2017).

Nach Zuordnung der einzelnen Audiodateien zu den jeweiligen Teilnehmern, wurden die Interviews einzeln abgehört und die Aussagen innerhalb des freien Berichts auf die zuvor festgelegten Kategorien der Erinnerungselemente hin untersucht und kodiert. Bei Vorliegen der definierten Voraussetzung einer Aussage als erinnerte Information wurde das entsprechende Item als „erinnert“ kodiert. Items, die nicht genannt wurden, wurden als „nicht erinnert“ kodiert.

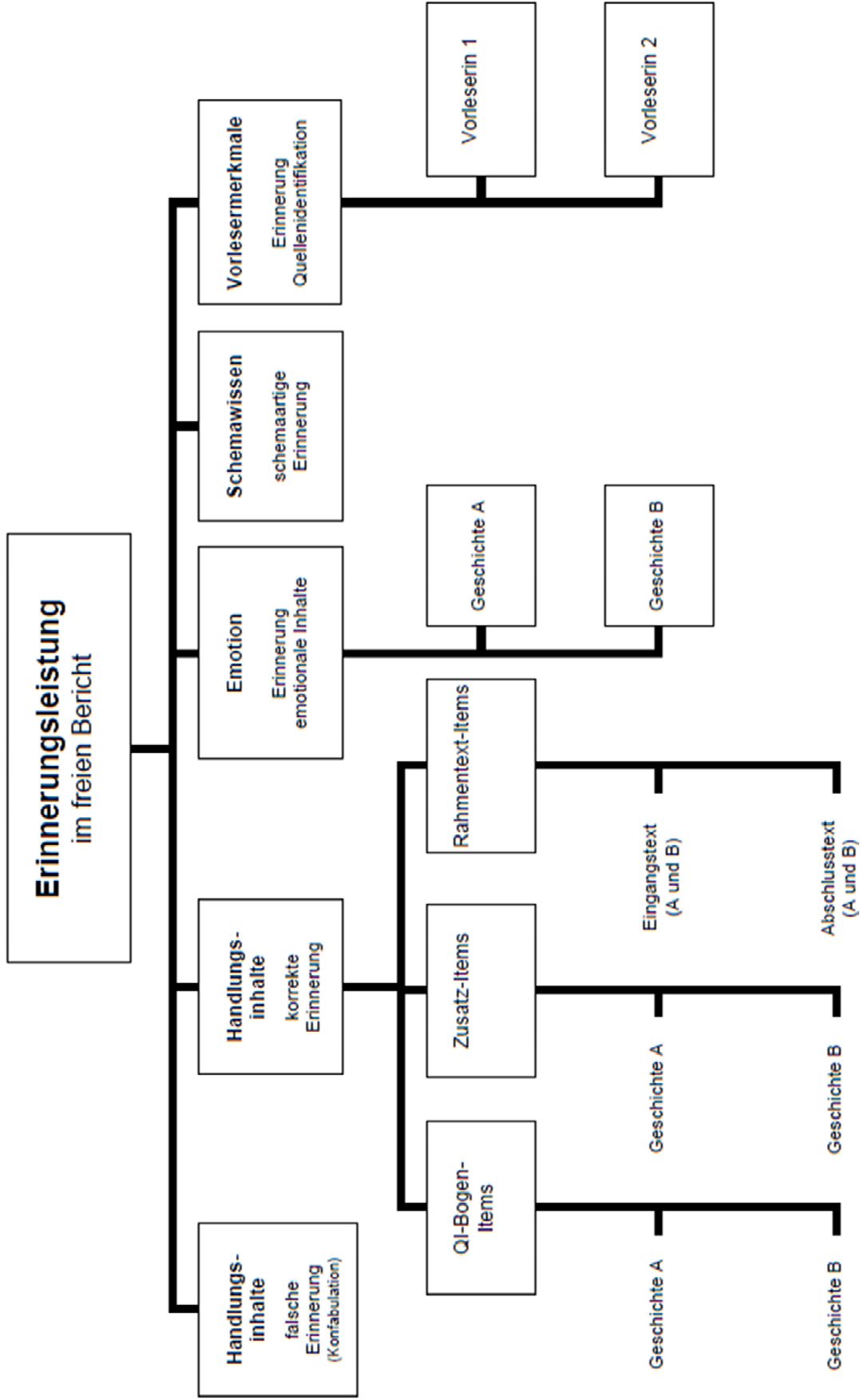


Abbildung 10. Das Kategoriensystem der qualitativen Inhaltsanalyse zur Untersuchung der kindlichen Erinnerungsleistungen im freien Bericht.

Zur Untersuchung der freien Berichte hinsichtlich der korrekten Erinnerung von Handlungsinhalten und hinsichtlich des Detailreichtums der Erinnerungen wurden die Kategorien *QI-Bogen* und *Zusatz* festgelegt. Die Basis bildeten die 14 Rekognitionsitems aus den QI-Bögen, diese wurden als QI-Bogen-Items bezeichnet, welche jedoch nicht alle Handlungen abdeckten. Um weitere oder fehlende Inhalte und eine stärkere Differenzierung aller Geschichtsdetails erfassen zu können, wurden entsprechende Items unter der Kategorie *Zusatz* zusammengestellt. Zur Festlegung, welche Aussagen als wahr im Sinne von korrekt erinnerter Information eingestuft werden, wurden für beide Kategorien die einzelnen Textbausteine der Zirkusattraktionen als Grundlage und Referenzrahmen herangezogen.

Konfabulationen sind objektiv falsche Aussagen oder Erzählungen, die möglicherweise auf falschen Wahrnehmungen oder aber auch auf Fehlleistungen des Gedächtnisses basieren. Die Kategorie Konfabulationen hat keine Referenz zum QI-Bogen. Dieser liegen keine Items aus dem Rekognitionstest zugrunde, als Treffer wurden die individuellen Aussagen kodiert, welche keine Überschneidungen mit den Handlungsinhalten des Stimulus-Materials aufwiesen. Für die vorliegende Untersuchung werden Konfabulationen als Kategorie für die Wiedergabe falscher Inhalte festgelegt. Diese wurden als Erinnerungen definiert, welche zwar komplette Handlungssequenzen mit zirkusbezogenem Inhalt darstellen, aber keine Übereinstimmungen mit den Handlungen der Stimulusgeschichten aufweisen. Da im freien Bericht keine Fragen zum oder Hinweise auf konkrete Inhalte der Zirkusgeschichten gegeben wurden, konnte der Einfluss von irreführenden Informationen, die falsche Erinnerungen zur Folge haben können, ausgeschlossen werden. Gemäß Larsson (2005) wird die fehlende Nennung eines Items nicht als Fehler betrachtet, da die Gedächtnisinhalte möglicherweise vorhanden sind, aber im freien Bericht nicht abgerufen werden können. Als Fehlerrate wird im Rahmen der qualitativen Analyse der Begriff Konfabulation als eine geäußerte, aber in Bezug auf die Referenzgeschichten nicht korrekte Information definiert (Heideler, 2010). Beispiele für präsentierte Konfabulationen sind in Anhang O aufgeführt. Diese Benennung ist nicht deckungsgleich mit der theoretischen Grundannahme von Köhnken et al. (1995), denen zufolge Konfabulationen eine Hinzufügung von gänzlich neuen Inhalten beinhaltet, Fehler jedoch als eine Abänderung vorhandenen Informationen verstanden werden. Im vorliegenden Fall wurden Konfabulationen und Fehler als Kategorie für die qualitative Analyse zusammengefasst.

Als Repräsentation von *Schemawissen* wurden Aussagen definiert, die einzelne Substantive im Singular oder Plural im Sinne einer „Überschrift“ für Zirkusattraktionen und keine Wiedergabe von konkreten Handlungsabläufen beinhalteten. Dabei wurde nicht unterschieden, ob genannte Oberbegriffe für Attraktionen stehen, die in den Geschichten vorkamen oder nicht. Die Anzahl und inhaltlichen Bezeichnungen der Items zur Erfassung schemaartiger Erinnerungsleistungen wurden während der Kodierung der Interviews erfasst, in Folge dessen wurden die bereits kodierten Interviews erneut auf dieses Item untersucht.

Zur Erfassung der Quellenidentifikationsleistungen im freien Bericht wurde die Kategorie Vorleser eingerichtet, um die Erinnerung personenspezifischer Merkmale der beiden Vorleserinnen zu erfassen. Da beide Vorleserinnen ähnliche Kleidung trugen, beziehen sich die generierten Items auf Merkmale wie Name, Haarfarbe und -struktur sowie Frisur. Unabhängig von diesen Merkmalen wurde angenommen, dass die Anzahl der erinnerten Items für Geschichte A oder B, je nachdem welche Vorleserin die Variante erzählte, einen Hinweis auf die Quellenidentifikation liefern kann.

7.3.2 Deskriptive Ergebnisse der Inhaltsanalyse

Die Darstellung der korrekten Erinnerungen erfolgt für die Gesamtwerte der Items innerhalb der Kategorien *QI-Bogen* und *Zusatz* für Geschichtsvariante A und B sowie als Gesamtwert aller Items beider Geschichten je Kategorie (siehe Tabelle 14). Es zeigten sich mehr korrekte Wiedergaben für die Items aus der Kategorie *QI-Bogen* als für die inhaltlich detaillierteren *Zusatz-Items*. Inhalte aus der Geschichte A wurden in beiden Kategorien tendenziell häufiger berichtet als Inhalte aus Geschichte B. Die Ergebnisse der Kategorie Rahmentextitems werden ebenfalls in Tabelle 14 aufgeführt. Hier zeigte sich eine vermehrte Repräsentation des Eingangstextes gegenüber der Schlusspassage der Erzählung.

Tabelle 14. Mittelwerte und Standardabweichungen korrekter Erinnerungsleistungen über alle Items der Kategorien QI-Bogen, Zusatz und Rahmentext über die gesamte Stichprobe und nach Altersklassen.

	Gesamt (<i>N</i> = 105) <i>M</i> (<i>SD</i>)	9 Jahre (<i>N</i> = 27) <i>M</i> (<i>SD</i>)	10 Jahre (<i>N</i> = 71) <i>M</i> (<i>SD</i>)	11 Jahre (<i>N</i> = 7) <i>M</i> (<i>SD</i>)
Items-QI-Bogen				
Geschichte A	1,05 (1,05)	1,11 (1,21)	1,08 (1,01)	0,43 (0,53)
Geschichte B	1,03 (1,03)	1,11 (1,15)	1,03 (1,00)	0,71 (0,95)
Geschichte A und B	2,08 (1,68)	2,22 (1,86)	2,11 (1,63)	1,14 (1,34)
Zusatz-Items				
Geschichte A	1,00 (1,10)	1,11 (1,12)	1,00 (1,14)	0,57 (0,53)
Geschichte B	0,71 (0,99)	0,56 (0,84)	0,77 (1,00)	0,71 (1,49)
Geschichte A und B	1,71 (1,54)	1,67 (1,35)	1,77 (1,60)	1,29 (1,70)
Rahmentext-Items				
Eingangstext	0,33 (0,47)	0,30 (0,46)	0,34 (0,47)	0,43 (0,53)
Abschlusstext	0,21 (0,40)	0,22 (0,42)	0,21 (0,41)	0,14 (0,37)

In der Gesamtstichprobe berichteten 28 Teilnehmer Konfabulationen. Innerhalb dieser Gruppe variierte die Anzahl falscher Erinnerungen zwischen einer und drei Konfabulationen. Eine Ausnahme stellte ein Teilnehmer dar, welcher sechs Konfabulationen zeigte. Der Mittelwert der präsentierten Konfabulationen lag bei $M = 1,43$ ($SD = 1,03$). Die anderen 74 % der Teilnehmer zeigten gemäß Kategoriendefinition keine falsche Erinnerungsleistung der Handlungsinhalte. Über die Gesamtstichprobe lag der Mittelwert für die Wiedergabe falscher Erinnerungen bei $M = 0,38$ ($SD = 0,82$). Emotionale Handlungsinhalte bzw. schemaartige emotionale Äußerungen traten nur vereinzelt auf.

Mit einem Mittelwert von $M = 0,14$ ($SD = 0,42$) wurden Vorleserinnenmerkmale erinnert und geäußert. Korrekte Erinnerungen von QI-Bogen Items der Geschichte A traten mit einem Mittelwert von $M = 1,22$ ($SD = 0,99$) häufiger auf, wenn Vorleserin 1 Geschichte A erzählte (Vorleserin 2: $M = 0,92$, $SD = 1,07$). Im Mittel ergab sich $M = 1,02$ ($SD = 1,11$) für QI-Bogen Items der Geschichte B für Vorleserin 1 und $M = 1,04$ ($SD = 0,92$) für Vorleserin 2. Über alle Items der Geschichten A und B der Kategorie QI-Bogen ergab sich für die experimentell randomisierte Kombination „Vorleserin 1 liest Geschichte A und Vorleserin 2 Geschichte B“, ein Mittelwert von $M = 2,27$ ($SD = 1,55$) und vice versa von $M = 1,92$ ($SD = 1,77$). Für die Zusatzitems fand sich für korrekt wiedergegebene Items der Geschichte A ein Mittelwert von $M = 1,07$ ($SD = 1,26$) für Vorleserin 1 von $M = 0,95$ ($SD = 0,98$) für Vorleserin 2. Für korrekt erinnerte Items aus Geschichte B sich ergab bei Präsentation durch Vorleserin 1 ein Mittelwert von $M = 0,85$ ($SD = 1,17$) und $M = 0,53$ ($SD = 0,78$) bei Vorleserin 2.

Insgesamt zeigten sich über alle Items der Kategorie Zusatz korrekte Erinnerungsleistungen mit $M = 1,60$ ($SD = 1,46$), wenn Vorleserin 1 Geschichtsvariante A vorgetragen und $M = 1,80$ ($SD = 1,60$), wenn Vorleserin 2 diese vorgetragen hatte. Bezüglich der Anzahl inkorrekt erinnerte Items in Form von Konfabulationen ergab sich ein Wert von $M = 0,51$ ($SD = 1,12$) wenn Vorleserin 1 Geschichte A und ein Mittelwert von $M = 0,28$ ($SD = 0,49$), wenn Vorleserin 2 Geschichte A vortrug. Die Präsentation von Schemawissen zeigte sich bei Erzählung der Geschichte A durch Vorleserin 1 mit einem Wert von $M = 1,24$ ($SD = 0,20$) und durch Vorleserin 2 mit $M = 1,27$ ($SD = 0,17$). Die Kontrollfrage nach einem bereits erlebten Zirkusbesuch zur Erfassung von vorhandenem Vorwissen beantworteten 78 Teilnehmer mit *Ja* und 27 mit *Nein* bei einem N von 105 gültigen Datensätzen. Schemawissen bei vorhandenem Vorwissen trat im Mittel mit $M = 1,27$ ($SD = 1,37$) und ohne Vorwissen mit $M = 1,17$ ($SD = 1,33$) auf. Bei der Auswertung von Konfabulationen im Zusammenhang mit Vorwissen berichteten im Durchschnitt $M = 0,27$ ($SD = 0,44$) Teilnehmer mit Vorwissen und $M = 0,25$ ($SD = 0,45$) Teilnehmer ohne Vorwissen Konfabulationen. Die Häufigkeit der Konfabulationen lag bei jenen mit Vorwissen bei $M = 0,40$ ($SD = 0,86$), solche ohne Vorwissen wiesen $M = 0,25$ ($SD = 0,45$) auf. Die Untersuchung der Mittelwerte für die korrekte und inkorrekte Wiedergabe der Handlungsinhalte und der Wiedergabe von Schemawissen im Zusam-

menhang mit der Interviewerin zeigte eine deskriptiv bessere Erinnerungsleistung korrekter Items wie auch weniger Schemapräsentation und Konfabulation bei Interviewerin 3.

7.4 Explorative Analyse

Zusätzlich zur deskriptiven und inferenzstatistischen Analyse wurden weiterführende Auswertungen vorgenommen, welche die Ergebnisse unter Aspekten beleuchten, die nicht in den vorab aufgestellten Hypothesen Berücksichtigung fanden. Hierbei wurde unter anderem eine varianzanalytische Untersuchung möglicher Kovariaten, wie *Schule* und der Kontrollvariablen *Anzahl der Zirkusbesuche* und *Vorleser Präferenz* durchgeführt. Die statistischen Tests wurden, entgegen der inferenzstatistischen Hypothesentest (diese erfolgten bei $p \leq .05$) auf dem Testniveau $p \leq .10$ durchgeführt. Im Rahmen der Kovarianzanalyse zeigte sich, dass in Abhängigkeit von der Schule (siehe Kapitel 6.3) signifikant unterschiedliche Ergebnisse für die *Qualität der Itemerkennung* Q_{IR} ($F_{(2,97)} = 2,81, p = .065, \eta^2 = .05$) resultierten. Im post-hoc Test erwies sich der Unterschied zwischen der Schule in Rodenkirchen ($M = 0,81; SD = 0,08$) und der Schule in Mülheim ($M = 0,75; SD = 0,09$) als signifikant. Die Überprüfung der Kontrollvariable *Zirkusbesuche* zeigte ebenfalls einen signifikanten Einfluss auf die *Qualität der Itemerkennung* Q_{IR} ($F_{(1,98)} = 4,05, p = .047, \eta^2 = .04$). Mit einem $M = 0,79$ ($SD = 0,09$) zeigten die Kinder, welche zuvor schon mal im Zirkus waren, bessere Leistungen in der Itemerkennung, als Kinder die noch nie im Zirkus waren ($M = 0,74; SD = 0,07$).

8 Diskussion

Die zentralen Fragestellungen der Studie behandeln die Auswirkungen variierender Instruktionen auf die Aussagegenauigkeit und Quellenidentifikationsleistung von Kindern. Dabei wurde die Zuverlässigkeitsmotivation über zwei Faktoren manipuliert: Es wurde die Relevanz von Bestrafung auf falsche Antworten in Kombination mit einer freien und einer forcierten Antwortbedingung untersucht. Hierfür wurden vier Befragungsbedingungen realisiert, welche sich durch hohe oder niedrige Schwellen der Zuverlässigkeitsmotivation unterschieden. Diese wurde über die Möglichkeit einer „Ich weiß nicht“ (*IWN*)-Antwortoption manipuliert (Faktor FORCIERUNG). Dabei wurde davon ausgegangen, dass somit Kooperationseffekten (Hughes & Grieve, 1980) entgegen-

gewirkt und das Gedächtnismonitoring erleichtert wird (Koriat & Goldsmith, 1996; Roebbers et al., 2001). Eine Bestrafung falscher Antworten im Sinne der operanten Konditionierung (Faktor STRAFE), sollte entsprechend der real-negativen Konsequenzen von Falschaussagen im rechtlichen Kontext eine Verschiebung des Antwortkriteriums hin zu einer höheren Schwelle bewirken. Als abhängige Maße wurden die Anzahl erinnerter Items im freien Bericht und traditionelle Maße der Quellenidentifikationsleistung erhoben.

Bei der Betrachtung der Ergebnisse bleibt vorweg anzumerken, dass es sich bei der vorliegenden Studie um ein Feldexperiment handelt, dessen Vorteil die hohe externe Validität ist, welche jedoch mit einer niedrigeren internen Validität (Eid et al., 2013) einhergeht. Somit ist der Einfluss möglicher Dritt- und Störvariablen nicht vollständig kontrollierbar.

Die Diskussion orientiert sich, entgegen dem inhaltlichen Aufbau des Ergebnisteils, am tatsächlichen Ablauf der Untersuchung: Nach einer Zusammenfassung der zentralen Ergebnisse, werden zunächst die qualitativen und anschließend die quantitativen Ergebnisse vor theoretischem Hintergrund kritisch erörtert.

8.1 Freier Bericht

Auf qualitativer Ebene zeigte die deskriptive Untersuchung der Erinnerungsleistung im freien Bericht, dass spontan mehr korrekte Items der Kategorie *Handlungsinhalte – Rahmentext* berichtet wurden. Es wurden häufiger Erinnerungen an den Eingangstext berichtet als Items der Schlusspassage. Darüber hinaus wurden im Mittel mehr Items der Geschichte A berichtet. Konfabulationen im Sinne falscher Erinnerungen oder hinzugefügter Items wurden von einem Viertel der Stichprobe berichtet.

Es fanden sich große Streuungen in den freien Berichten hinsichtlich der drei hergeleiteten Kategorien *Rahmentext*, *QI-Bogen* und *Zusatz*. Insgesamt wiesen die Kinder durchschnittlich mehr Erinnerungen an Handlungsinhalte der Geschichte A auf, was in einem möglichen Zusammenhang mit der Quellenidentifikationsleistung steht. Die Auswertung der Kategorie *Zusatzitems*, zur Feststellung der Detailliertheit über die mit dem QI-Bogen erfragten Items hinaus, weist darauf hin, dass keine weiteren detaillierteren Handlungsinhalte frei abgerufen wurden als jene, die mit der Kategorie *QI-Bogen* erfasst wurden. Überraschend war die geringe spontane Nennung von Items im freien

Bericht. Eine mögliche Begründung für das Zustandekommen der Ergebnisse könnte die geringe persönliche Bedeutsamkeit der Untersuchung für die Teilnehmer sein oder mangelnder Aufmerksamkeit bzw. Störungen in der Lernphase geschuldet sein. Die mangelnde persönliche Relevanz könnte auch das seltene Auftreten von Emotionsitems erklären. Die Ablenkung durch externe Ereignisse wurde im vorliegenden Fall durch eine videographische Aufzeichnung der Lernphase kontrolliert.

Grundsätzlich fällen Personen bei dem Bericht von vergangenen Ereignissen Entscheidungen, welche Aspekte einer Geschichte sie hervorheben möchten, welche Perspektive sie einnehmen und/ oder welcher Grad an Detailliertheit oder Generalisierung intendiert wird (Goldsmith & Koriat, 1999; Koriat & Goldsmith, 1996). Diese Art der strategischen Kontrolle kann die heterogenen Ergebnisse im freien Bericht erklären.

Unter der Annahme des seriellen Positionseffekts (Deese & Kaufmann, 1957) bzw. des *primacy*- und *recency*-Effekts (Atkinson & Shiffrin, 1968), sollten *Rahmentextitems* gegenüber anderen Items besser erinnert werden, da sie als erst- und letztgenannte Informationen präsentiert wurden und demnach besser gespeichert werden sollten. Dies spiegelt sich in den Ergebnissen nur hinsichtlich der Items des Eingangstextes wider.

Bezüglich der Quellenidentifikation zeigte sich, dass die Teilnehmer beide Vorleserinnen spontan nennen und Erinnerungen über diese wiedergeben. Da im freien Bericht keine gezielte Wiedererkennung erfragt wurde, kann auch keine Aussage darüber gemacht werden, welcher Quelle die erinnerten Items zugeordnet wurden. Es zeigten sich rein deskriptiv jedoch Erinnerungsvorteile für Items der Geschichte A bei Präsentation durch Vorleserin 1, trotz Vorlesertraining und einer Randomisierung der Geschichten pro Vorleserin und Rotation bei der Durchführung in den verschiedenen Schulen. Daraus kann geschlossen werden, dass die unterschiedlichen Quellen wahrgenommen, aber im freien Bericht von den Versuchsteilnehmern nicht ausreichend kenntlich gemacht wurden. Aus den Studien von Thierry und Pipe (2009) und Kovacs und Newcombe (2006) ist bekannt, dass eine gute Unterscheidbarkeit verschiedener Quellen für sehr junge Kinder im Vergleich zu Grundschulkindern von besonderer Bedeutung ist. Möglicherweise sind die Quellen im vorliegenden Fall nicht ausreichend differenzierbar wahrgenommen worden. Eine Erhebung der Quelleninformation im

freien Bericht wurde nicht vorgenommen, es kam jedoch zu Nennungen der Vorleserinnen. Die Kinder haben versucht Inhalte zu Vorleserin 1 oder 2 zuzuordnen oder die Inhalte entsprechend der Chronologie berichtet, in dem die Reihenfolge konkret benannt wurde.

Insgesamt zeigten wenige Kinder inkorrekte Erinnerungsleistungen in Form von Konfabulationen. Die Kontrolle möglicher Drittvariablen, beispielsweise zwischen Vorwissen durch einen bereits erlebten Zirkusbesuch und der Produktion von Konfabulationen zeigte, dass sowohl Kinder mit Vorwissen als auch Kinder ohne Vorwissen falsche Erinnerungen im Sinne von Konfabulationen präsentierten. Systematische Unterschiede fanden sich dabei nicht. Ergebnisse der Studien von Greenhoot (2000) und Sutherland, Pipe, Schick, Murray und Gobbo (2003) deuten zwar auf einen positiven Effekt von Vorwissen hin, jedoch kann Vorwissen auch negative Einflüsse auf Maße der Erinnerung ausüben (Ornstein, Baker-Ward, Gordon & Merritt, 1997). Saywitz und Snyder (1993) haben gezeigt, dass skriptbasierte Fehler durch Vorab-Interviews reduziert werden können, hier könnte durch entsprechende Hinweise vorgebeugt werden. Im Hinblick auf inkorrekte Erinnerungen zeigte sich weiterhin, dass Kinder tendenziell weniger bei Vorleserin 2 konfabulierten. In Abhängigkeit von der Interviewerin zeigten die Kinder eine unterschiedliche Anzahl an Erinnerungen, trotz intensivem Training resultierten bei Interviewerin 3 rein deskriptiv die geringsten Leistungen. Die Ergebnisse verdeutlichen, dass die Fähigkeit des Interviewers, im freien Bericht Erinnerungen anzustoßen und Informationen zu erheben, bei Befragungen nicht zu vernachlässigen ist (Lamb et al., 2007; Larsson, 2005). Dies ist für die Erinnerungsleistung bei der Rekognition und Quellenidentifikation nicht wirksam, da die Gedächtnisleistung mit Hilfe einer Liste abgefragt wurde.

Es bleibt weiterhin zu diskutieren, ob schemakonforme Items als Konfabulationen zu werten sind, da durch einen einmaligen Zirkusbesuch der Kinder im Vorfeld und unabhängig von der Erhebung nicht zwingend Vorwissen generiert worden sein muss. Der hier aufgestellten Definition der Schema-Items liegt die Annahme zugrunde, dass schemaartige Wiedergaben aufgrund der vorgelesenen Geschichten erinnert und nicht frei erfunden werden (Heidler, 2010). Hier sei darauf verwiesen, dass Köhnken und Kollegen (1995) eine Unterscheidung von Konfabulationen und Fehlern postulierten, welche sich im Hinzufügen neuer oder der Abänderung bestehender Inhalte unter-

scheidet. Vice versa stellt sich die Frage, ob die als Konfabulationen erfassten Aussagen frei erfunden sind oder ob diesen nicht auch Vorwissen zugrunde liegt. Denkbar ist auch, dass Kinder zu konfabulierten Antworten im Rahmen der sozialen Erwünschtheit neigen (Roebbers, Howie & Beuscher, 2007). Statt Erinnerungslücken zuzugeben, werden naheliegende Antworten gegeben, die ihren Ursprung in Schemawissen oder auch in Erinnerungsfehlern haben können.

Insgesamt sollte zu den qualitativen Ergebnissen erwähnt werden, dass die Nennung von Items im freien Bericht nicht zwangsläufig gleichzusetzen ist mit tatsächlich vorhandener Erinnerung. Es ist aber grundsätzlich auch möglich, dass mehr Erinnerungen vorliegen, diese aber nicht frei berichtet wurden (Larsson, 2005).

8.2 Itemerkennung und Quellenidentifikation

Für die Itemerkennung und Quellenidentifikation wurden zwei Haupteffekte angenommen. Die negative Ausprägung des Faktors FORCIERUNG (Vorhandensein der Option einer *IWN*-Antwort) sollte zu besseren Gedächtnisleistungen führen. Auf theoretischer und empirischer Grundlage (Koriat & Goldsmith, 1994, 1996; Roebbers et al., 2001) wurde davon ausgegangen, dass durch eine freie Antwortoption Prozesse des Gedächtnismonitorings angestoßen werden und daraus ein Qualitäts-Quantitäts Trade-off resultiert. Eine weitere Vorannahme betraf die Auswirkungen des Faktors STRAFE und deren Abstufungen. So sollte in Anlehnung an Roebbers et al. (2001) und Garven et al. (2001) Strafe die Leistung verbessern, komplementär zu den Effekten der Belohnung. Annahmen zu einer Interaktion der Faktoren konnten aus der bestehenden Literatur nicht abgeleitet werden.

Die statistische Analyse ließ erkennen, dass bei obligatorischen Antworten signifikant akkuratere Aussagen resultierten als bei der Beantwortung mit *IWN*-Option. Die Bestrafung hingegen hatte keinen Einfluss auf die Aussagequantität und -qualität. Die Bereitstellung einer *IWN*-Antwortoption zeigte keine positiven Effekte auf die Aussagequalität hinsichtlich Itemerkennung und Quellenidentifikation. Damit zeigte sich in der Gesamtschau ein signifikanter Effekt entgegen der angenommenen Wirkrichtung. Die Richtung der Effekte war für die Itemerkennungs- und die Quellenidentifikationsleistung gleich. Es wurden jedoch quantitativ mehr Items wiedererkannt als Quellen identifiziert, was als Hinweis dafür gewertet werden kann, dass es sich um unterscheidbare und differenzierte Prozesse handelt. Die Ergebnisse auf deskriptiver

Ebene entsprechen den Annahmen von Aschermann (1998), die zeigen konnte, dass die Itemschwierigkeit (Döring & Bortz, 2016) der zwei Prozesse Itemerkennung und Quellenidentifikation unterschiedlich ist - Itemerkennung ist einfacher.

Für die abhängigen Maße der Itemerkennung zeigte sich ein signifikanter Haupteffekt für den Faktor FORCIERUNG bezüglich der Qualität der Itemerkennungsleistung (Q_{IR}). Eine obligatorische Antwort führte in der vorliegenden Untersuchung dazu, dass Kinder bessere Leistungen bei der Wiedererkennung von Items aus zwei verschiedenen Zirkusgeschichten zeigen. Weiterhin stehen die Ergebnisse im Gegensatz zu den Forschungsarbeiten von Roebbers und Elischberger (2002), Roebbers und Fernandez (2002) und Roebbers et al. (2001), welche für Erwachsene einen signifikant positiven Effekt der *IWN*-Antwortoption finden konnten.

Die Ergebnisse von Koriat et al. (2001) weisen darauf hin, dass Kinder ein liberales Antwortkriterium haben und im Vergleich zu Erwachsenen nur wenig Gebrauch von einer *IWN*-Antwortoption machen, wenn die Möglichkeit besteht. Auch in der vorliegenden Studie wurde nur etwa in 11,1 % der Fälle eine *IWN*-Antwort gegeben. Zudem haben die Kinder die *IWN*-Antwortoption nicht zur Antwortverbesserung nutzen können, sie erwies sich sogar als nachteilig. Ursächlich hierfür könnte sein, dass kein Einüben der Antwortmöglichkeit vor der eigentlichen Untersuchung stattgefunden hat.

Eine weitere mögliche Erklärung für die seltene Nutzung der *IWN*-Antwortoption ist die mangelnde Möglichkeit zur Kontrolle des Instruktionsverständnisses (hierzu auch Kapitel 8.3 „Limitationen“). Es wurde zwar systematisch auf diese Antwortoption hingewiesen, es kann aber nicht ausgeschlossen werden, dass die *IWN*-Option für die Kinder nicht als gleichwertige Antwort eingeschätzt und genutzt wurde. Außerdem überschätzten die Kinder möglicherweise ihre Erinnerungsfähigkeit und nutzten deshalb die *IWN*-Option zu selten. Gemäß Foley und Johnson (1985) bestehen intraindividuelle Unterschiede im internalen und externalen Source Monitoring, deren Zusammenhang mit der Entwicklung exekutiver Funktionen genauer untersucht werden sollte.

Der hypotheseninkonforme Effekt kann zum einen durch die unausgereiften Fähigkeiten zum Gedächtnismonitoring bei Kindern (Koriat & Goldsmith, 1994) erklärt werden; bei erwachsenen Probanden zeigte sich durch die *IWN*-Antwortoption eine Verbesserung der Erinnerungsleistung (Roebbers et al., 2001). Alternativ könnte gerade der

Zwang, eine Antwort zu geben (analog zu den Schulerfahrungen der Kinder), eine Verschiebung des Antwortkriteriums bewirkt haben, so dass vermehrte Monitoringprozesse angestoßen wurden. Durch die verpflichtende Antwort, durch die Ratetendenzen herangezogen werden müssen, erhöhte sich somit allein die theoretische Wahrscheinlichkeit einer richtigen Antwort.

Erkenntnisse aus empirischen Untersuchungen zum schulischen Testverhalten von Grundschulkindern zwischen acht und elf Jahren legen nahe, dass grundsätzlich adäquate Monitoringprozesse angewandt werden, wenn einfach unterscheidbare Items oder Aufgaben hinsichtlich ihrer Korrektheit differenziert werden sollen (Krebs & Roebbers, 2010). Hier fand sich ein Alterseffekt für das Monitoring von Unsicherheit: Ältere Kinder können besser als jüngere falsche oder unsichere Antworten aus einem vorab selbst ausgefüllten Test herausstreichen und geben für leichte Items höhere Einschätzungen der Sicherheit. Die Ergebnisse lassen darauf schließen, dass sich zunächst regulative Fähigkeiten für den Gedächtnisabruf von nicht-ambivalenten Informationen entwickeln. Roebbers, Schmid und Roderer (2009) konnten zudem zeigen, dass für inkorrekte Antworten geringere Sicherheitsurteile abgegeben werden. Im Lehr-Lern-Kontext ist die Fähigkeit, falsche Antworten zu identifizieren weitaus relevanter als bei der Erinnerung an selbst erlebte Ereignisse. Acht- und neunjährige sind durchaus in der Lage ihre Gedächtnisleistung zu kontrollieren, jedoch sind sie noch immer durch entwicklungsbedingte Besonderheiten in ihrer Gedächtnisleistung limitiert (Schneider & Bjorklund, 1992). Für die vorliegenden Daten kann geschlussfolgert werden, dass die im Vergleich zur Stichprobe von Krebs und Roebbers (2010) relativ jungen Kinder in diesem Prozess der Überprüfung von Antworten noch nicht fortgeschritten waren, so dass sie hier nicht im gleichen Maße wirksam werden konnten. Für die eigene Studie bedeuten diese Ergebnisse, dass die Unterscheidbarkeit der gelernten Inhalte nicht ausreichend gegeben war.

In der weiterführenden Analyse zeigte sich ein signifikanter Effekt der Kontrollvariable Zirkusbesuche auf die Itemrekognitionsleistung. Dies hat möglicherweise durch die Aktivierung von Vorwissen erleichternd gewirkt (Greenhoot, 2000; Sutherland et al., 2003). Auch durch irreführende Fragen können Prozesse des Gedächtnismonitorings bei Kindern forciert werden (Thierry et al., 2001).

Merkmale der Befragungssituation wird im gedächtnispsychologischen und vor allem forensischen Kontext besondere Bedeutung beigemessen. Diese beziehen sich auf die konkrete Befragung, deren Kontext, Fragenformulierungen oder auch auf den Interviewer. Ansätze, die sich Fragen mit offenem Ende bedienen, können mit größerer Wahrscheinlichkeit akkurate Erinnerungen im Vergleich zu reinen Rekognitionsfragen erheben (Pipe et al., 2004). Zusätzlich sind aber auch fokussierte Fragen für den Interviewer von hoher Bedeutsamkeit. Damit werden spezifische Inhalte abgefragt und es entsteht indirekt ein größerer Druck auf die Befragten eine Antwort zu geben oder sich mit dem Interviewer abzustimmen, unabhängig davon, ob sie sich der Antwort sicher sind oder nicht (Thierry et al., 2001). Möglicherweise wirkten *IWN*-Antwortoptionen stärker bei offenen Fragen und weniger stark in Rekognitionsformaten.

Im Gegensatz führen Rekognitionsfragen mit größerer Wahrscheinlichkeit zu fehlerhaften Antworten in Augenzeugenkontexten aufgrund von Antwortverzerrungen, wie beispielsweise Ja-/Nein-Tendenzen und falscher Wiedererkennung von Details, die in früheren Interviews erwähnt wurden oder auf die geschlossen wurde (Brainerd & Reyna, 2012). Betrachtet man die Unterschiede zwischen freiem Bericht und Rekognition bzw. Quellenidentifikation lassen sich jedoch keine systematischen Ergebnisse in diese Richtung finden.

Positive Verstärkung durch den Interviewer kann Aussagen bis hin zu fantastischen, falschen Anschuldigungen verfälschen (Garven et al., 2001; in Deutschland sei hier auf die Wormser Prozesse verwiesen, vgl. Kapitel 2.3 „Kinder als Zeugen“). Auch die Wiederholung von Fragen kann zu unerwünschten Revisionen und Änderungen der Erstantwort führen (Krähenbühl, Blades & Eiser, 2009). Wird die Wiederholung einer Frage durch den Interviewer aber sinnvoll begründet, reduzierte dies die Wahrscheinlichkeit einer unerwünschten Antwortverschiebung in Richtung Ungenauigkeit, nicht aber das Auftreten von Response Shifts an sich (Howie et al., 2004). Die wiederholte Befragung durch denselben Interviewer hat bei Kindern entgegen den Erwartungen positive Auswirkungen auf die Aussageleistung (Bjorklund et al., 2000). In der Gesamtschau kann geschlussfolgert werden, dass wiederholte Fragen suggestives Potential besitzen und bei der Befragung kindlicher Zeugen möglicherweise zu einem liberaleren Antwortkriterium führen. Im vorliegenden Fall wurde nur einmalig und nicht suggestiv befragt, so dass ein Einfluss der Befragungstechnik hier vernachlässigt werden kann.

Die Bestrafung falscher Antworten bewirkte in der vorliegenden Untersuchung keine statistisch bedeutsamen Unterschiede in der Itemrekognitions- und Quellenidentifikationsleistung. Die undifferenzierte Wirkung könnte mit einer multinomialen Modellierung der Antwortprozesse weitergehend untersucht werden. In der Studie von Schütz (2011) wird darauf eingegangen, dass zur Erfassung von Diskriminations- und Raterprozessen der Alt-Neu-Rekognition und des Quellengedächtnisses zum einen Modelle der Signalentdeckungstheorie und zum anderen multinomiale Schwellenmodelle genutzt werden können. Die Modellierung anhand der Signalentdeckungstheorie (Green & Swets, 1966) wird seit langem als valideres Messverfahren betrachtet, da eine gute Passung mit bestehenden Vorstellungen von Gedächtnisstrukturen und Funktionen besteht. Außerdem bietet sich die Anwendung in binären Antwortformate an. Bei einer Erweiterung der Antwortoptionen ist die Nutzung von Receiver Operating Characteristics (ROC)-Kurven und multinomialen Schwellenmodellen angezeigt (Schütz, 2011). Eine Modellierung mit multinomialen Modellen, welche den Bias des Ratens oder die kriteriale Verschiebung abbilden können, steht als nächster Schritt des vorliegenden Forschungsvorhabens an.

Für den Faktor STRAFE ergeben sich in der Gesamtschau keine systematischen Variationen. Der antizipierte Entzug von Token konnte im Kontrast zur Belohnung (Robbers et al., 2001) keine signifikante Verbesserung der Gedächtnisleistung bewirken. Trotz der Annahme, dass negative Folgen im Kontext von Falschaussagen näher an der Realität im rechtlichen und auch aussagepsychologischen Kontext sind, konnte die Wirkung von Strafe in der vorliegenden Untersuchung nicht nachgewiesen werden.

Die Versuchsteilnehmer wurden vor dem Interview über den Wert der Token informiert und darüber, dass diese im Anschluss gegen Süßigkeiten eingetauscht werden könnten. Möglicherweise war die individuelle Wertbeimessung dieser bevorstehenden Belohnung bei den Kindern unterschiedlich hoch ausgeprägt, weshalb eine angedrohte Strafe, der Entzug der Token, nicht den intendierten Effekt der Erhöhung der Zuverlässigkeitsmotivation bewirkte. Diese Vermutung kann analog zu den Ergebnissen von Baumert und Demmrich (2001) betrachtet werden, welche den Einfluss von Anreizen auf Motivation und Performanz untersuchten und feststellen konnten, dass Testergebnisse eher durch Anstrengung und Konzentration als durch antizipierte Konsequenzen zustande kommen. Baumert und Demmrich (2001) realisierten in Anlehnung an das

„Programme for International Student Assessment“ (PISA) vier Untersuchungsbedingungen, die sich in der Bearbeitungsinstruktion unterschieden. 1. Bedingung, identisch mit der original PISA-Instruktion: Die Ergebnisse dieser Studie helfen Politikern und Schulen herauszufinden wie und was Schüler lernen - wir bitten dich, dein Bestes zu geben; 2. Bedingung: Die Ergebnisse werden deinen Lehrern mitgeteilt; 3. Bedingung: Dein Klassenlehrer wird deine Leistung benoten; 4. Bedingung: Für gute Leistungen erhältst du eine finanzielle Belohnung. Im Ergebnis zeigte sich, dass sich weder die beabsichtigten noch die investierten Anstrengungsbemühungen der Schüler oder deren Leistungen in den unterschiedlichen Bedingungen voneinander unterschieden. Baumert und Demmrich (2001) nahmen an, dass alle Schüler unabhängig von der Untersuchungsbedingung bemüht waren, gute Ergebnisse zu erzielen. Dies stützt die Ergebnisse zur ersten Hypothese der vorliegenden Untersuchung: Die unrelativierten Mittelwerte aller Gruppen liegen über der fünfzigprozentigen Ratewahrscheinlichkeit (siehe Kapitel 7.1 „Deskriptive Statistik“).

In den im Theorieteil vorgestellten Studien variierte die Modalität, über welche das Wissen in der Lernphase vermittelt wurde, zwischen Tonband, Video, Vorlesen (Bright-Paul, Jarrold & Wright, 2005) oder live bzw. direkter Interaktion (Hala et al., 2013) und definierte damit die Komplexität der Lernsituation und die Realitätsnähe. In der vorliegenden Studie wurde eine live und direkte Interaktion durch geschulte Vorleserinnen realisiert. Dieser Modalitätseffekt zeigte bei Tobey und Goodman (1992), dass durch Interaktion eine höhere Akkuratheit im freien Recall erfolgte. Für die vorliegende Studie kann davon ausgegangen werden, dass eine im Vergleich hohe Interaktion zwischen dem präsentierten Lernmaterial und den Teilnehmern stattfand.

8.3 Limitationen

Es müssen verschiedene methodische und inhaltliche Begrenzungen bedacht werden, welche eine Generalisierung der Ergebnisse auf Anwendungskontexte einschränken. Auf diese Aspekte wird im nächsten Abschnitt eingegangen.

Generell wird auf der Grundlage empirischer Untersuchung davon ausgegangen, dass Kinder Ereignisse besser erinnern, welche sie selbst erlebt haben, als wenn sie nur Beobachter waren. Unmittelbares Erleben geht jedoch immer mit einer Reduktion der experimentellen Kontrolle einher, da nicht sichergestellt werden kann, dass alle Grup-

pen exakt dasselbe Material präsentiert wurde oder identische Lernsituationen durchlaufen haben. Durch die Schulung der Vorleserinnen, die Videokontrolle und die kleinen Versuchsgruppen wurden relevante Anstrengungen unternommen, um diese Einschränkungen zu kontrollieren. Ob dies hinreichend gelungen ist, kann nicht sicher beurteilt werden. Die Rezeptionsphase und das Material, welche Störeinflüsse auf die Leistung ausüben konnten, sind im Sinne von inzidentellem Lernen möglicherweise ursächlich für die nicht-signifikanten Effekte entgegen der in der Theorie gefundenen Ergebnisse.

Des Weiteren wurde unter Abwägung der Realitätsnähe, eine Befragungssituation, ähnlich zu einer Glaubhaftigkeitsbegutachtung geschaffen, welche ggf. gleichzeitig aber die externe Validität der Ergebnisse reduziert. Es ist denkbar, dass es durch die freie Befragung zu Interaktionseffekten kam, welche die Rekognitionsleistung beeinflusst haben können.

Für die Erinnerungsleistungen in der vorliegenden Untersuchung kann festgestellt werden, dass Kinder im freien Bericht mehr korrekte Informationen präsentieren als falsche. Die Ergebnisse könnten zukünftig auf ihren Bedeutungsinhalt, beispielsweise angelehnt an die Kriterien der qualitativen Inhaltsanalyse (CBCA), untersucht werden. Die Vielfältigkeit der gezeigten Erinnerungen könnte auch im Hinblick auf mögliche Wechselwirkungen mit individuellen aufmerksamkeitsbedingten Zuständen während der Präsentation der Geschichten untersucht werden.

So äußerten einige Kinder, dass sie zum Zeitpunkt des Interviews schulische oder private Probleme gehabt haben und ihnen daher das Erinnern schwer falle. Auch andere Mitschüler wurden als Störquelle beim Hören der Geschichten erwähnt. Diese Anmerkungen der Kinder spiegeln aber auch die Realität von Augenzeugen wieder, deren abzurufende Erinnerung oft unter komplexen und schwierigen Umständen entstanden ist. Denkbar ist aber auch, dass es sich hierbei um eine externale Attributionsstrategie gehandelt hat, die zur Erklärung schlechter Leistungen im schulischen Kontext häufig auftritt. Dies zeigt aber auch, wie gut Kinder bereits über sozialrekonstruktive Gedächtnisprozesse reflektieren und diese auch benennen.

Hinsichtlich methodischer Einschränkung muss erwähnt werden, dass bei der Auswertung der freien Berichte keine Interraterreliabilität erfasst werden konnte, da nur eine

Kodiererin die Daten ausgewertet hat. Diesem Umstand sollte in zukünftigen Untersuchungen Rechnung getragen werden.

Die Manipulation der Zuverlässigkeitsmotivation über den Faktor STRAFE setzt voraus, dass den Token und der Belohnung ein individueller Wert beigemessen wird. Es kann für die untersuchte Stichprobe nicht ausgeschlossen werden, dass diese nicht von Relevanz war und deshalb keinen Effekt auf die Metakognition hatte.

Weiterhin könnte die geringe Erinnerungsleistung durch die einwöchige Latenz zwischen Lernphase und Abruf zustande gekommen sein, so dass Bodeneffekte auf Grund des langen Erinnerungsintervalls resultierten. Andererseits sind die Ergebnisse aus Untersuchungen, die nur das unmittelbare Erinnern fokussieren, nur bedingt auf Anwendungskontexte der Zeugenpsychologie übertragbar – hier sind Latenzen von einigen Tagen bis mehreren Wochen eher die Regel.

Das Versuchsmaterial wurde unter Berücksichtigung ethischer Aspekte so gewählt, dass es emotional zu keinen negativen Folgen kommen sollte. Es handelte sich um ein positives Ereignis (Zirkusvorstellung), wohingegen in anderen Studien häufig negative Vorgänge gezeigt werden (z.B. Diebstahlszenen). Dieser Materialunterschied kann ein Grund sein für die von den Vorstudien abweichenden Ergebnisse und stellt auch eine Einschränkung hinsichtlich der unmittelbaren Übertragbarkeit auf die Praxis dar.

Eine Ursache für die geringe Nutzungshäufigkeit der *I/WN*-Antwortoption könnte eine fehlende Übungsphase sein. Poole und Lindsay (2002) haben festgestellt, dass Kinder die Quelle von gesehenen und gehörten Ereignissen durch eine Trainingsphase besser unterscheiden konnten. Der Gebrauch dieser Antwortoption sollte in weiterführenden Untersuchungen vor der eigentlichen Testphase beispielhaft erläutert und von den Teilnehmern vor der Untersuchung durchgeführt werden. Auch muss überlegt werden, inwieweit die Durchführung während der Schulzeit bei den Kindern unerwünschte Effekte gehabt haben könnte. Die soziale Interviewsituation wurde möglicherweise unter der Annahme, „ich muss eine Antwort geben, wenn ich gefragt werde“, wahrgenommen und entsprechend schulischen Interaktionsregeln unreflektiert auf die Untersuchung übertragen.

Das Verständnis von Instruktionen stellt eine zentrale Einflussvariable auf das Zustandekommen der hiesigen Ergebnisse dar. Dabei handelt es sich um einen Komplex an

Fähigkeiten und Fertigkeiten, wobei kommunikative Kompetenzen im Fokus stehen. Darunter fällt aber auch Aufgabenverständnis. Sader (1957) stellt in Frage, ob durch eine Standardisierung der äußeren Versuchsbedingungen, wie sie Instruktionen zu meist darstellen, eine tatsächliche Konstanzhaltung der gegebenen kognitiven und sozio-emotionalen Bedingungen für die Versuchsperson möglich ist. Dabei beschränkt er den Instruktionsbegriff auf die Vorphase experimenteller Untersuchungen, welche in der Regel nicht thematischer Teil eines Experiments sind. Im vorliegenden Fall konnte, trotz wiederholtem Hinweis auf die *IWN*-Option, keine experimentelle Kontrolle sicherstellen, dass die Kinder das Angebot zur Nutzung einer *IWN*-Antwort verstanden, internalisiert und als gleichwertig zu einer Ja- oder Nein-Antwort wahrgenommen haben. Hier könnte wiederum eine Übungsphase Abhilfe schaffen, auf der anderen Seite könnte bei Erwachsenen eine bereits fortgeschrittene Automatisierung dieser metakognitiven Überwachungsprozesse stattgefunden haben, welche die von Vorstudien abweichenden Ergebnisse begründen.

Bei der Betrachtung der Auswertungsbögen wird ersichtlich, dass in den *forced*-Bedingungen die Kategorie „geraten“ erhoben wurde – für jegliche Antworten der Kinder kann jedoch nicht ausgeschlossen werden, dass diese durch Rateprozesse entstanden oder beeinflusst wurden. Dazu wurden die Teilnehmer nicht befragt und äußerten dies auch nicht spontan. Eine Abfrage zur Sicherheit der Antworten hätte diese Störeinflüsse in gewissem Maße kontrollieren können (Krebs & Roebbers, 2010). Eine Modellierung der Daten mit multinomialen Modellen, welche möglichen Rateprozessen Rechnung trägt, erfolgt im Anschluss an die vorliegende Forschungsarbeit.

Alternativ zum Source-Monitoring Framework kann die Fuzzy-Trace Theory (Brainerd & Reyna, 1995) im Rahmen der Entwicklung von gedächtnispsychologischen Theorien genannt werden. Diese Theorie geht davon aus, dass Informationen in zwei distinkten Strängen gespeichert werden, in Oberflächeninformationen und bedeutungshaltige Informationen. Kinder speichern der zufolge eher wörtliche, detaillierte Informationen und sind demnach in der Speicherung von Inhalten begrenzt, was hier als Grund für die Bodeneffekte angeführt werden kann.

In diesem Zusammenhang bleibt einschränkend anzumerken, dass es sich bei dem vorliegenden Versuch um eine vergleichsweise schwere Aufgabe für die Teilnehmer

gehandelt hat: Die Präsentation des Lernmaterials kann analog zu einem Listenlernparadigma betrachtet werden, wobei ähnliche Inhalte aus zwei verschiedenen Geschichten mit einer Woche Abstand erfragt wurden. Das Erlernen von insgesamt 28 Attraktionen aus zwei Quellen hat die Kinder im Schulalltag möglicherweise überfordert.

Zusammenfassend wird deutlich, dass sich die Argumentationslinien in diesem Feld noch nicht stabilisiert haben und weitere Untersuchungen der Zusammenhänge zwischen metakognitiven Fähigkeiten und aussagepsychologischen Befragungen von Kindern ausstehen.

C RESÜMEE

9 Fazit und Implikationen

Bereits kleine Kinder können sich an selbst erlebte Ereignisse über lange Zeiträume hinweg erinnern und davon berichten, besonders dann, wenn die Erfahrung eine hohe Salienz hatte. Erinnerungen an eine von mehreren ähnlichen Erfahrungen können hingegen schwierig sein und es ist wahrscheinlich, dass Details zwischen den Episoden verwechselt werden. Die Identifikation der Quellen von Wissen und Erinnerungen stellt in diesem Zusammenhang einen wichtigen Bereich im Rahmen der Entwicklung metakognitiver Gedächtnisprozesse bei Kindern dar. Im Alltag werden Kinder regelmäßig zu zurückliegenden Ereignissen interviewt, sei es im privaten oder schulischen Umfeld. Dabei werden durch den Befrager, häufig auch unbewusst, strategische Hinweise zur Steigerung metakognitiver Fähigkeiten benutzt, zum Beispiel: „Versuche dich genau zu erinnern, wo du den Gegenstand verloren hast“ oder „Wenn du dich richtig erinnerst, bekommst du eine gute Note“. Diese und weitere Instruktionen können die Gedächtnisleistung von Kindern beeinflussen.

Die Genauigkeit von Erinnerungen kann verbessert werden, wenn Kinder sich dabei jeweils auf einzelne Ereignisse konzentrieren. Verwechslungen zwischen selbst erlebten Episoden oder Quellenfehler lassen aber nicht notwendigerweise Zweifel an der Richtigkeit oder Glaubhaftigkeit der Aussage von kindlichen Zeugen im forensischen Kontext aufkommen. Wie Kinder befragt werden, kann jedoch dramatische Auswirkungen auf die Genauigkeit ihrer Berichte haben und es bleibt auch in Zukunft zu prüfen, auf welche Art und Weise dies erfolgen sollten.

Das Alleinstellungsmerkmal der Arbeit bezieht sich auf Kombination der drei Faktoren Präsentationsmodalität (Garven et al., 2000; Tobey & Goodman, 1992), Latenz und Auswertungsmethodik. Hinsichtlich der Lernphase handelte es sich im vorliegenden Fall um live präsentiertes, das heißt selbst erlebtes Material, entgegen der häufig verwendeten Präsentation von Videos.

Die vorliegende Forschungsarbeit beschäftigte sich mit den Auswirkungen einer Manipulation von Zuverlässigkeitsmotivation über eine *Ich weiß nicht*-Antwortoption und über eine Bestrafung falscher Antworten. Dabei wurden die Prozessleistungen der Itemerkennung und Quellenidentifikation von vorab erlebten Ereignissen erhoben. Es wurde festgestellt, dass Kinder durch obligatorische Antworten die besten Leistungen

erzielen, was im Gegensatz zu Vorstudien und auch der empfohlenen Vorgehensweise in der forensischen Praxis steht. Grundsätzlich sollte es Kindern im aussagepsychologischen Kontext frei gestellt werden eine Antwort zu geben. Damit kann Suggestionseffekten und Konfabulationen vorgebeugt werden; weiterhin sollten Berichte nur non-selektiv verstärkt werden. Obwohl es sich bei dem hier verwendeten Material um Lerninhalte ohne forensische Relevanz gehandelt hat, stand die Erhebung gedächtnispsychologischer Leistung im Vordergrund, womit die Ergebnisse übertragbar auf und von Bedeutung für Anwendungskontexte der Befragung sind.

In der psychologischen Literatur findet sich häufig die Diskussion, ob Erinnerungen an schmerzhaft oder traumatische Erlebnisse einzigartigen Mechanismen der Speicherung unterliegen oder durch die gleichen erklärt werden können, die sich auf Erinnerungen an andere Ereignisse beziehen (Cordón, Pipe, Sayfan, Melinder & Goodman, 2004, Volbert 2004). Traumatische Erlebnisse sind oft unverwechselbar und einzigartig, so dass Erinnerungen daran im Laufe der Zeit besser erhalten bleiben als Erinnerungen an weniger ausgeprägte Erlebnisse oder bedeutende Ereignisse (Ornstein et al., 1997). Im Gegensatz zu der Annahme der verdrängten Erinnerungen hat die grundlagenorientierte experimentelle Forschung nachgewiesen, dass nicht alle Fälle von sexuellem Missbrauch zwangsläufig traumatisierend sind und der Effekt von leichterem Abruf durch eine höhere Salienz oder Bedeutsamkeit des Ereignisses somit nicht gegeben ist (Pipe et al., 2004; Volbert, 2004). Damit begründet sich auch die Wahl des ethisch unbedenklichen Materials der vorliegenden Arbeit zur Erfassung von Gedächtnisfunktionen.

Bei Testaufgaben zur Wiedererkennung verwechseln Kinder selbst erlebte Episoden, mit solchen, von denen sie nur gehört (Poole & Lindsay, 2001) oder welche sie beobachtet (Roberts & Blades, 2000; Thierry, Spence, & Memon, 2001) haben. Es ist unklar, wie viele Informationen aus anderen Quellen in die Erzählungen während des Experiments mit einbezogen wurden. Betrachtet man die Anzahl an Konfabulationen und den signifikanten Einfluss der Kovariate Vorwissen, wird deutlich, dass Schemawissen durchaus berichtet wurde und die Ergebnisse mit Vorsicht interpretiert werden sollten.

Abschließend wird die Bedeutsamkeit der Ergebnisse für die Praxis betrachtet. Dabei werden Implikationen für die gutachterliche Tätigkeit in der aussagepsychologischen Begutachtung angerissen.

Die Ergebnisse der Studie weisen auf die bei Kindern im besonderen Maße von Erwachsenen abweichenden Gedächtnisfunktionen hin. Insbesondere die verbale Erläuterung von Antwortalternativen kommt hier zum Tragen.

Die Vernehmungsmethode des kognitiven Interviews findet in der Fachliteratur großen Anklang, im vorliegenden Fall wurde eine Grundform dieser Befragungsart durchgeführt, nämlich der freie Bericht. Dazu kann festgehalten werden, dass das kognitive Interview eine gut untersuchte Vernehmungsmethodik darstellt, um den Aussageanteil korrekter Informationen bei der Befragung von Zeugen zu erhöhen (Larsson, 2005). Es hat sich in der Praxis vor allem in der Anwendung mit Erwachsenen gut bewährt.

Für Kinder zeigen sich nach Memon und Bull (1991) die Techniken als teilweise schwer verständlich, daher wird eine Anpassung an die besonderen Anforderungen entwicklungsbedingter Gedächtnisfähigkeiten von Kindern gefordert. Hier sei auf die Leitfäden des NICHD (Lamb et al., 2007) verwiesen. Neuere Forschung setzt hier an und entwickelt an kindliche Fähigkeiten angepasste Methoden mit angemessenen Strukturierungshilfen zur Förderung korrekter Gedächtniswiedergaben bei Kindern (Roebbers & Elichberger, 2002).

Um Aussagen zur Effektivität des freien Berichtes auf die hier gezeigten Erinnerungsleistungen im Vergleich zu anderen Methoden zu bestimmen, müssten Untersuchungen mit Experimental- und Kontrollbedingungen durchgeführt werden. Nach Roebbers und Elichberger (2002) ist der freie Bericht alleine nicht geeignet, um Erinnerungsleistungen hervorzubringen, die in Informationsgehalt und -qualität den Anforderungen an zuverlässige und vollständige Aussagen erfüllen. Die forensische Praxis braucht demnach unkomplizierte Interviewmethoden für Kinder, die möglichst ohne weitere Hilfsmittel und langwierige Trainingsmaßnahmen einsetzbar sind und dabei möglichst viele korrekte und forensisch wichtige Informationen liefern können. In Anlehnung daran wird deutlich, dass das Forschungsfeld der Aussagepsychologie, im Spannungsfeld zwischen experimenteller Grundlagen- und Anwendungsforschung steht und methodischen Herausforderungen bei der Ableitung allgemeingültiger Aussagen unterliegt.

Literatur

- American Professional Society on the Abuse of Children (APSAC) [WWW Dokument]
URL <https://www.apsac.org/guidelines> [letzter Zugriff: 01.02.2018]
- Anderson, J. R. (1976). *Language, memory and thought*. Oxford: Lawrence Erlbaum.
- Anderson, R. C. & Pichert, J. W. (1978). Recall of previously unrecalable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, 17, 1-12.
- Arntzen, F. (1993). *Psychologie der Zeugenaussage. System der Glaubwürdigkeitsmerkmale* (3. Auflage). München: Beck.
- Arntzen, F. (2008). *Vernehmungspsychologie*. München: C.H. Beck.
- Arntzen, F. (2011). *Psychologie der Zeugenaussage*. München: C.H. Beck.
- Aschermann, E. (1998). *Interviews mit Kindern. Die imaginale Erinnerungsstrategie als Befragungsmethode für Grundschul Kinder*. Unveröffentlichte Habilitationsschrift, Christian-Albrechts-Universität zu Kiel.
- Aschermann, E., Mantwill, M. & Köhnken, G. (1991). An Independent Replication of the Effectiveness of the Cognitive Interview. *Applied Cognitive Psychology*, 5, 489–495.
- Atkinson, R. & Shiffrin, R. M. (1968). Human memory. A proposed system and its control processes. In J. T. Spence & K. W. Spence (Ed.), *The psychology of learning and motivation* (pp. 89-195). New York: Academic Press.
- Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In G.A. Bower (Ed.), *Recent Advances in Learning and Motivation* (Vol. 8) (pp. 47–89). New York: Academic Press.
- Balloff, R. (2004). Grundlagen von Klärung und Diagnostik bei sexuellem Missbrauch. *Sexueller Missbrauch*. Band 1 (Bd. 1, S. 107-120). Göttingen: Hogrefe.
- Batchelder, W. H. & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review*, 97, 4, 548-564.

- Baumert, J. & Demmrich, A. (2001). Test Motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education, 16*(3), 441-462.
- Bayen, U. J., Murnane, K. & Erdfelder, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(1), 197-215.
- Bederian-Gardner, D. & Goldfarb, D. (2014). Expectations of emotions during testimony: the role of communicator and perceiver characteristics. *Behavioral sciences & the law, 32*(6), 829-845. <https://doi.org/10.1002/bsl.2146>
- Bell, R., Mieth, L. & Buchner, A. (2017). Emotional memory: No source memory without old-new recognition. *Emotion, 17*(1), 120-130.
- Bemis, R. H., Leichtman, M. D. & Pillemer, D. B. (2013). I Remember When You Taught Me That! Preschool Children's Memories of Realistic Learning Episodes. *Infant and Child Development, 22*(6), 603-621. <https://doi.org/10.1002/icd.1807>
- Beuscher, E. & Roebbers, C. M. (2005). Does a warning help children to more accurately remember an event, to resist misleading questions, and to identify unanswerable questions? *Experimental Psychology, 52*(3), 232-241. <https://doi.org/10.1027/1618-3169.52.3.232>
- Bjorklund, D. F., Cassel, W. S., Bjorklund, B. R., Brown, R. D., Park, C. L., Ernst, K. et al. (2000). Social demand characteristics in children's and adults' eyewitness memory and suggestibility: the effect of different interviewers on free recall and recognition. *Applied Cognitive Psychology, 14*(5), 421-433. [https://doi.org/10.1002/1099-0720\(200009\)14:5<421::AID-ACP659>3.0.CO;2-4](https://doi.org/10.1002/1099-0720(200009)14:5<421::AID-ACP659>3.0.CO;2-4)
- Bower, G. H., Gilligan, S. G. & Monteiro, K. P. (1981). Selectivity of learning caused by affective states. *Journal of Experimental Psychology, 110*, 451-473.
- Brady, M. S., Poole, D. A., Warren, A. R. & Jones, H. R. (1999). Young Children's responses to Yes-No Questions: Patterns and Problems. *Applied Developmental Science 3*(1), 47-57.

- Brainerd, C. J. & Reyna, V. F. (1995). Autosuggestibility in memory development. *Cognitive Psychology*, 28, 65–101. <http://dx.doi.org/10.1006/cogp.1995.1003>.
- Brainerd, C. J. & Reyna, V. F. (2012). Reliability of children's testimony in the era of developmental reversals. *Developmental Review*, 32, 224-267.
- Breuer, M. M. (2008). *Entdeckung von Täuschung: von Alltagsvorstellungen zu empirisch fundiertem Wissen*. Dissertation, Justus-Liebig-Universität Gießen.
- Brewer, W. F. & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, 13, 207-230.
- Bright-Paul, A., Jarrold, C. & Wright, D. B. (2005). Age-appropriate cues facilitate source monitoring and reduce suggestibility in 3- to 7-year-olds. *Cognitive Development*, 20(1), 1–18.
- Bröder, A. & Meiser, T. (2007). Measuring Source Memory. *Zeitschrift für Psychologie*, 215(1), 52-60.
- Bruck, M. & Melnyk, L. (2004). Individual differences in children's suggestibility. A review and synthesis. *Applied Cognitive Psychology*, 18(8), 947-996.
- Bundesgerichtshof in Strafsachen (BGHSt) 45, 164 [WWW Dokument] URL <https://www.hrr-strafrecht.de/hrr/1/98/1-618-98.php3> [letzter Zugriff: 01.02.2018]
- Büttner, G. (2003). Gedächtnisentwicklung im Kindes- und Jugendalter. *Sprache, Stimme, Gehör*, 27, 24-30.
- Cassel, W. S., Roebbers, C. M. & Bjorklund, D. F. (1996). Developmental Patterns of Eyewitness Responses to Repeated and Increasingly Suggestive Questions. *Journal of Experimental Child Psychology*, 61, 116-133.
- Ceci, S. J. & Bruck, M. (1993). Suggestibility of the child witness: a historical review and synthesis. *Psychological Bulletin*, 113(3), 403-439.
- Christmann, B. & Wazlawik, M. (2016). Professionelle Herausforderungen bei Befragungen von Kindern in (Verdachts-)Fällen von sexueller Gewalt – Das NICHD-

Protocol als Beispiel leitfadengestützter Gesprächsführung. *Kindesmisshandlung und –vernachlässigung*, 19, 2, 112-123.

<https://doi.org/10.13109/kind.2016.19.2.112>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed)*.

Hillsdale, N.J.: L. Erlbaum Associates.

Cohen, S. B., Leslie, A. M. & Frith, U. (1985). Does the autistic child have a „theory of mind“? *Cognition*, 21, 37-46.

Cole, C. B. & Loftus, E. F. (1987). The memory of children. In M. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Children’s eyewitness memory* (pp. 178–208). New York: Springer.

Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407-428.

Connolly, D. A. & Lindsay, D. S. (2001). The influence of suggestions on children’s reports of a unique experience versus an instance of a repeated experience. *Applied Cognitive Psychology*, 15, 205-223.

Cordón, I. M., Pipe, M. E., Sayfan, L., Melinder, A. & Goodman, G. S. (2004). Memory for traumatic experiences in early childhood. *Developmental Review*, 24, 101-132.

Craig, R., Scheibe, R., Raskin, D. C., Kircher, J. C. & Dodd, D. H. (1999). Interviewer Questioning and Content Analysis of Children’s Statements of Sexual Abuse. *Applied Developmental Science*, 3(2), 77-85.

Crawley, S. L., Newcombe, N. & Bingman, H. (2010). How focus at encoding affects children's source monitoring. *Journal of Experimental Child Psychology*, 105, 273-285.

Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39, 2, 223-238.

Deese, J. & Kaufman, R. A. (1957). Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology*, 54(3), 180–187. <https://doi.org/10.1037/h0040536>

- Dodier, O. & Denault, V. (2018). The Griffiths Question Map: A Forensic Tool For Expert Witnesses' Assessments of Witnesses and Victims' Statements. *Journal of Forensic Sciences*, 63(1), 266-274.
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Auflage). Berlin, Heidelberg: Springer.
- Eid, M., Gollwitzer, M. & Schmitt, M. (2013). Unterschiede zwischen mehreren unabhängigen Stichproben: Varianzanalyse und verwandte Verfahren. In M. Eid, M. Gollwitzer & M. Schmitt (Hrsg.). *Statistik und Forschungsmethoden* (S.371-446). Basel: Beltz Verlag.
- Erdfelder, E. & Bredenkamp, J. (1998). Recognition of script-typical versus scriptatypical information: Effects of cognitive elaboration. *Memory & Cognition*, 26, 922-938.
- Felser, G. (2015). *Werbe- und Konsumentenpsychologie* (4.Auflage). Berlin: Springer.
- Fiedler, K. & Schmid, J. (1999). Gutachten über Methodik und Bewertungskriterien für Psychologische Glaubwürdigkeitsgutachten. *Praxis der Rechtspsychologie*, 9(2), 5-45.
- Fisher, R. P. (2010). Interviewing cooperative witnesses. *Legal and Criminological Psychology*, 15(1), 25-38. <http://dx.doi.org/10.1348/135532509X441891>.
- Fisher, R., Geiselman, R. E., Raymond, D. S., Jurkevich, L. M. & Warhaftig, M. L. (1987). Enhancing enhanced eyewitness memory: Refining the Cognitive Interview. *Journal of Police Science & Administration*, 15, 291-297.
- Fivush, R. & Hammond, N. (1990). Autobiographical memory across the preschool years: Towards reconceptualizing childhood amnesia. In R. Fivush and J. A. Hudson (Eds), *Knowing and remembering in young children* (pp. 223-248). New York: Cambridge University Press.
- Foley, M. A. & Johnson, M. K. (1985). Confusions between Memories for Performed and Imagined Actions. A Developmental Comparison. *Child Development*, 56, 1145-1155.

- Friedrich, H. & Mandl, H. (1992). Lern- und Denkstrategien: Ein Problemaufriß. In H. F. Friedrich & H. Mandl (Hrsg.), *Lern- und Denkstrategien: Analyse und Intervention* (S. 3–54). Göttingen: Hogrefe.
- Garven, S., Wood, J. M. & Malpass, R. S. (2000). Allegations of Wrongdoing, The Effect of Reinforcement on Children's Mundane and Fantastic Claims. *Journal of Applied Psychology*, *85*(1), 38-49.
- Geiselman, R. E. & Fisher, R. P. (2014). *Interviewing Witnesses and Victims*. [WWW Dokument] URL <https://www.psych.ucla.edu/faculty/page/rgeiselm> Worddokument *Current CI Research* [Letzter Zugriff: 01.03.18]
- Geiselman, R. E., Fisher, R. P., MacKinnon, D. P. & Holland, H. L. (1986). Enhancement of eyewitness memory with the cognitive interview. *American Journal of Psychology*, *99*(3), 385-401.
- Goldsmith, M. & Koriat, A. (1999). The strategic regulation of memory reporting: Mechanisms and performance consequences. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII — Cognitive regulation of performance: Interaction of theory and application* (pp. 373–400). Cambridge, MA: MIT Press.
- Goodman, G. S. (1984), The Child Witness: An Introduction. *Journal of Social Issues*, *40*, 1-7.
- Gopnik, A. & Graf, P. (1988). Knowing How You Know: Young Children's Ability to Identify and Remember the Sources of Their Beliefs. *Child development*, *59*(5), 1366–1371. Verfügbar unter <http://www.jstor.org/stable/1130499>
- Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley
- Greenhoot, A. F. (2000). Remembering and Understanding: The Effects of Changes in Underlying Knowledge on Children's Recollections. *Child development*, *71*(5), 1309–1328. doi: 10.1111/1467-8624.00230.
- Greuel, L. (2001). *Wirklichkeit, Erinnerung, Aussage*. Weinheim: Beltz.
- Haden, C. A., Haine, R. A. & Fivush, R. (1997). Developing Narrative Structure in Parent-Child Reminiscing Across the Preschool Years. *Developmental Psychology*, *33*, 295-307.

- Haigh, S. N. & Robinson, E. J. (2009). What children know about the source of their knowledge without reporting it as the source. *European Journal of Developmental Psychology*, 6(3), 318–336.
- Hala, S., Brown, A. M., McKay, L.-A. & San Juan, V. (2013). Two-and-a-Half-Year-Old's Memory for Sources of Actions: Contextual Support Facilitates Recall. *Journal of Cognition and Development*, 14(2), 343–358.
- Hala, S., Rasmussen, C. & Henderson, A. M. (2005). Three Types of Source Monitoring by Children With and Without Autism: The Role of Executive Function. *Journal of Autism and Developmental Disorders*, 35(1), 75–89.
- Hart, J. T. (1967). Memory and the memory-monitoring process. *Journal of Verbal Learning and Verbal Behavior*, 6(5), 685–691. [https://doi.org/10.1016/S0022-5371\(67\)80072-0](https://doi.org/10.1016/S0022-5371(67)80072-0)
- Hassebrauck, M. (1995). Kognitionen von Beziehungsqualität. Eine Prototypenanalyse. *Zeitschrift für Sozialpsychologie*, 26(3), 160–172.
- Heidler, M.-D. (2010). Konfabulationen - Alte und neue Theorien eines rätselhaften Phänomens. *Fortschritte der Neurologie-Psychiatrie*, 78(5), 256–268. <https://doi.org/10.1055/s-0029-1245217>
- Hershkowitz, I., Orbach, Y., Lamb, M. E., Sternberg, K. J. & Horowitz, D. (2001). The effects of mental context reinstatement on children's accounts of sexual abuse. *Applied Cognitive Psychology*, 15, 235–248.
- Hughes, M., & Grieve, R. (1980). On asking children bizarre questions. In M. Donaldson, R. Grieve & C. Pratt (Eds.), *Early childhood development and education* (pp. 104–114). Oxford, England: Blackwell.
- Home Office (2011). Achieving Best Evidence in Criminal Proceedings. Guidance on interviewing victims and witnesses, and guidance on using special measures. [WWW Dokument] URL https://www.cps.gov.uk/sites/default/files/documents/legal_guidance/best_evidence_in_criminal_proceedings.pdf [Letzter Zugriff: 15.05.2018]
- Howie, P., Sheehan, M., Mojarrad, T. & Wrzesinska, M. (2004). 'Undesirable' and 'desirable' shifts in children's responses to repeated questions: age differences

- in the effect of providing a rationale for repetition. *Applied Cognitive Psychology*, 18(9), 1161–1180. <https://doi.org/10.1002/acp.1049>
- Howitt, D. (2011). *Introduction to Forensic and Criminal Psychology* (4th Ed.). Prentice Hall: Pearson Education limited.
- Inbau, F. E., Reid, J. E., Buckley, J. P. & Jayne, B. C. (2011). *Criminal interrogations and confessions* (5. Auflage). Aspen: Gaithersberg.
- Innocence Project [WWW Dokument] URL <https://www.innocenceproject.org/> [Letzter Zugriff: 01.03.18]
- Johnson, M. K., Hashtroudi, S. & Lindsay, S. D. (1993). Source monitoring. *Psychological Bulletin*, 1, 3–28.
- Johnson, M. K. & Raye, C. L. (1981). Reality Monitoring. *Psychological Review*, 88, 1, 67–85.
- Johnson, C. N. & Wellman, H. M. (1980). Children's Developing Understanding of Mental Verbs: Remember, Know, and Guess. *Child Development*, 51(4), 1095–1102. Verfügbar unter <http://www.jstor.org/stable/1129549>
- Köhnken, G., Kraus, U. & vom Schemm, K. (2008). Das Kognitive Interview. In R. Volbert & M. Steller (Hrsg.), *Handbuch der Rechtspsychologie* (S. 232 – 243). Göttingen: Hogrefe.
- Köhnken, G., Mantwill, M. & Aschermann, E. (1991). Das Kognitive Interview - experimentelle Evaluation einer neuen Vernehmungsmethode. In R. Egg (Hrsg.), *Brennpunkte der Rechtspsychologie* (S. 247-258). Bonn: Forum Verlag Godesberg.
- Köhnken, G., Milne, R., Memon, A. & Bull, R. (1999). The cognitive interview: A meta-analysis. *Psychology, Crime and Law*, 5, 3–27.
- Köhnken, G., Schimossek, E., Aschermann, E. & Höfer, E. (1995). Statement Validity Analysis and the cognitive interview. *Journal of Applied Psychology*, 80, 671-684.

- Koriat, A. & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General*, *123*, 297-316.
- Koriat, A. & Goldsmith, M. (1996). Monitoring and Control Processes in the Strategic Regulation of Memory Accuracy. *Psychological Review*, *103* (3), 490-517.
- Koriat, A., Goldsmith, M. & Pansky, A. (2000). Toward a psychology of memory. *Annual Review of Psychology*, *51*, 481-537.
- Koriat, A., Goldsmith, M., Schneider, W. & Nakash-Dura, M. (2001). The Credibility of Children's Testimony: Can Children Control the Accuracy of Their Memory Reports? *Journal of Experimental Child Psychology*, *79*(4), 405-437.
- Kovacs, S. L. & Newcombe, N. (2006). Developments in source monitoring: The role of thinking of others. *Journal of Experimental Child Psychology*, *93*, 25-44.
- Krähenbühl, S., Blades, M. & Eiser, C. (2009). The effect of repeated questioning on children's accuracy and consistency in eyewitness testimony. *Legal and Criminological Psychology*, *14*(2), 263-278.
- Krebs, S. S. & Roebbers, C. M. (2010). Children's strategic regulation, metacognitive monitoring, and control processes during test taking. *British Journal of Educational Psychology*, *80*, 325-340.
- Krix, A. C. & Sauerland, M. (2013). Wie lassen sich Zeugenaussagen verbessern? – Neue Entwicklungen und Methoden. *Praxis der Rechtspsychologie – Themenschwerpunkt Aussage- und Zeugenpsychologie*. Berlin: DPV.
- Lamb, M. E. (1994). The investigation of child sexual abuse: An interdisciplinary consensus statement. *Child Abuse & Neglect*, *18*, 1021-1028.
- Lamb, M. E., Orbach, Y., Hershkowitz, I., Esplin, P. W. & Horowitz, D. (2007). A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: a review of research using the NICHD Investigative Interview Protocol. *Child abuse & neglect*, *31*(11-12), 1201-1231.
- Lamb, M. E., Sternberg, K. J. & Esplin, P. W. (1998). Conducting investigative interviews of alleged sexual abuse victims. *Child Abuse & Neglect*, *22*, 813-823.

- Lange, G., Guttentag, R. E. & Nida, R. E. (1990). Relationships between Study Organization, Retrieval Organization, and General and Strategy-Specific Memory Knowledge in Young Children. *Journal of Experimental Child Psychology*, 49, 126–146.
- Larsson, A. S. (2005). *Interviewing child witnesses*. Dissertation, Universität Göteborg, Schweden.
- Leichtman, M. D., Morse, M. B., Dixon, A. & Spiegel, R. (2000). Source monitoring and suggestibility: an individual differences approach. In K.P. Roberts & M. Blades (Eds.), *Children's source monitoring* (pp. 257–287). Mahwah, NJ: Erlbaum.
- Leslie, A. M. (2000). The Theory of Mind as a mechanism of selective attention. In M. S. Gazzaniga (Ed.), *The New Cognitive Neurosciences (2nd Edition)* (pp.1235-1247). Cambridge, Mass: MIT Press.
- Lewin, K. (1963). *Feldtheorie in den Sozialwissenschaften*. Bern: Huber.
- Lindsay, D. S., Johnson, M. K. & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology*, 52, 297-318.
- Loohs, S. (1996). *Die Verwendung spezifischer Explorationsmethoden zur Befragung kindlicher Zeugen im Hinblick auf Gedächtnisleistung, Suggestibilität und das Wiedererkennen von Gesichtern*. Dissertation, Universität Regensburg.
- Mantwill, M., Köhnken, G. & Aschermann, E. (1995). Effects of the cognitive interview on the recall of familiar and unfamiliar events. *Journal Of Applied Psychology*, 80(1), 68-78.
- Markowitsch, H. J. & Staniloiu, A. (2011). Amygdala in action: Relaying biological and social significance to autobiographic memory. *Neuropsychologia*, 49, 718–733.
- Mayring, P. (1991). Qualitative Inhaltsanalyse. In U. Flick (Hrsg.), *Handbuch qualitative Sozialforschung. Grundlagen, Konzepte, Methoden und Anwendungen* (209-213). München: Psychologie Verlags Union.

- Mayring, P. (2015). *Qualitative Inhaltsanalyse – Grundlagen und Techniken*. Weinheim: Beltz.
- Meiser, T. & Bröder, A. (2002). Memory for multidimensional source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(1), 116–137.
- Memon, A. & Bull, R. (1991). The cognitive interview: Its origin, empirical support, evaluation and practical implications. *Journal of Community & Applied Social Psychology*, 1(4), 291-307.
- Memon, A. & Vartoukian, R. (1996) The Effects of Repeated Questioning on Young Children's Eyewitness Testimony. *British Journal of Psychology*, 87, 403-415.
- Memon, A., Wark, L., Bull, R. & Köhnken, G. (1997). Isolating the effects of the cognitive interview techniques. *British Journal of Psychology*, 88(2), 179-198.
- Milne, R. & Bull, R. (2003). *Psychologie der Vernehmung*. Bern: Huber.
- Murnane, K. & Bayen, U. J. (1996). An evaluation of empirical measures of source identification. *Memory and Cognition* 24(4), 417-428.
- Neisser, U. (1982). Snapshots or benchmarks?. In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 43–48). San Francisco: Freeman.
- Nelson, K. (1986). *Event knowledge: A functional approach to cognitive development*. Hillsdale, NJ: Erlbaum.
- Nelson, K. (2000). Memory and belief in development. In D. L. Schacter & E. Scarry (Ed.), *Memory, brain and belief* (pp. 259-289). Cambridge: Harvard University Press.
- Nelson, K. & Fivush, R. (2004). The emergence of autobiographical memory: a social cultural developmental theory. *Psychological Review*, 111(2), 486-511.
- Nelson, T. O. & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Ed.), *Metacognition: Knowing about knowing* (pp. 1-25). Cambridge, MA, US: The MIT Press.

- Nesbitt, M. & Markham, R. (1999). Improving young children's accuracy of recall for an eyewitness event. *Journal of Applied Developmental Psychology, 20*(3), 449-459.
- Niehaus, S. (2008). Merkmalsorientierte Inhaltsanalyse. In R. Volbert (Hrsg.), *Handbuch der Rechtspsychologie* (S. 311-321). Göttingen: Hogrefe.
- Niehaus, S., Krause, A. & Schmidke, J. (2005). Täuschungsstrategien bei der Schilderung von Sexualstraftaten. *Zeitschrift für Sozialpsychologie, 36*(4), 175-187.
- Nolde, S. F., Johnson, M. K. & D'Esposito, M. (1998). Left prefrontal activation during episodic remembering: an event-related fMRI study. *Neuroreport, 9*, 3509–3514.
- Nyberg, L. (1999). Imaging episodic memory. Implications for cognitive theories and phenomena. *Memory, 7*(5), 585–597.
- O'Sullivan, J. T. (1993). Preschoolers' beliefs about effort, incentives, and recall. *Journal of Experimental Child Psychology, 55*, 396–414.
- Oberlader, V. A., Naefgen, C., Koppehele-Goseel, J., Quinten, L., Banse, R. & Schmidt, A. F. (2016). Validity of content-based techniques to distinguish true and fabricated statements: A meta-analysis. *Law and Human Behavior, 40*(4), 440-457.
- Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P.W. & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged abuse victims. *Child Abuse & Neglect, 24*, 733–752.
- Ornstein, P. A., Baker-Ward, L., Gordon, B. N. & Merritt, K. A. (1997). Children's memory for medical experiences: Implications for testimony. *Applied Cognitive Psychology, 11*, S87–S104.
- Ornstein, P. A., Naus, M. J. & Liberty, C. (1975). Rehearsal and organizational processes in children's memory. *Child Development, 46*, 818–830.
- Pipe, M.-E., Lamb, M. E., Orbach, Y. & Esplin, P. W. (2004). Recent research on children's testimony about experienced and witnessed events. *Developmental Review, 24*(4), 440-468.

- Poole, D. A. & Lamb, M. E. (1998). *Investigative interviews of children: A guide for helping professionals*. Washington, DC: American Psychological Association.
- Poole, M.-E. & Lindsay, D. S. (2001). Children's Eyewitness Reports after Exposure to Misinformation From Parents. *Journal of Experimental Psychology: Applied*, 7(1), 27-50.
- Poole, D. A., & Lindsay, D. S. (2002). Reducing child witnesses' false reports of misinformation from parents. *Journal of Experimental Child Psychology*, 81, 117-140.
- Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515-526.
doi:10.1017/S0140525X00076512
- Priestley, G., Roberts, S., & Pipe, M-E. (1999). Returning to the scene: Reminders and context reinstatement enhance children's recall. *Developmental Psychology*, 35, 1006–1019.
- Principe, G. F. & Schindewolf, E. (2012) Natural conversations as a source of false memories in children: Implications for the testimony of young witnesses. *Developmental Review*, 32, 205-223.
- Quas, J. A. & Schaaf, J. M. (2002). Children's memories of experienced and nonexperienced events following repeated interviews. *Journal of Experimental Child Psychology*, 83, 304–338.
- Ramer, S. L. (2005). Site-ation pearl growing: methods and librarianship history and theory. *Journal of the Medical Library Association*, 93, 3, 397-403.
- Regber, A. (2007). Glaubhaftigkeit und Suggestibilität kindlicher Zeugenaussagen unter Einbeziehung entwicklungspsychologischer Aspekte. In T. Ley (Hrsg.), *Schriftenreihe der Thüringer Fachhochschule für öffentliche Verwaltung, Fachbereich Polizei* (Band 5). Frankfurt: Verlag für Polizeiwissenschaft.
- Ricci, C. M., Beal, C. R. & Dekle, D. J. (1996). The Effect of Parent Versus Unfamiliar Interviewers on Children's Eyewitness Memory and Identification Accuracy, *Law and Human Behavior*, 20(5), 483-506.

- Riehle, B. M & Zwingmann, C. (2000). Effektstärkevarianten beim Eingruppen-Prä-Post-Design: Eine kritische Betrachtung. *Rehabilitation*, 39, 189–199.
- Roberts, K. P. & Blades, M. (2000). *Children's source monitoring*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Robinson, W. P. (1996). *Language and language behaviors, Vol. 6. Deceit, delusion, and detection*. Thousand Oaks, CA, US: Sage Publications, Inc.
- Robinson, E. J., Haigh, S. N. & Nurmsoo, E. (2008). Children's working understanding of knowledge sources: Confidence in knowledge gained from testimony. *Cognitive Development*, 23(1), 105–118.
- Roderer, T. & Roebbers, C. M. (2008). Children's strategic regulation of memory accuracy. In M. R. Kelley (Ed.), *Applied memory* (pp. 253-274). Hauppauge: Nova Science Publishers.
- Roebbers, C. M. & Fernandez, O. (2002). The Effects of Accuracy Motivation on Children's and Adult's Event Recall, Suggestibility and their Answers to unanswerable Questions. *Journal of Cognition and Development*, 3, 415-443.
- Roebbers, C. M., Howie, P. & Beuscher, E. (2007). Can private reports enhance children's event recall, lower their suggestibility and foster their metacognitive monitoring compared to face-to-face interviews? *Computers in Human Behavior* 23 (1), S. 749–769. doi: 10.1016/j.chb.2004.11.007.
- Roebbers, C. M., Schmid, C. & Roderer, T. (2009). Metacognitive monitoring and control processes involved in primary school children's test performance. *British Journal of Educational Psychology*, 62(79), 749-767.
- Roebbers, C. M. & Schneider, W. (2005a). The strategic regulation of children's memory performance and suggestibility. *Journal of Experimental Child Psychology*, 91, 24-44.
- Roebbers, C. M. (2010). Teil II Entwicklung des Gedächtnisses. Strategische Kontrolle von Gedächtnisleistungen bei Kindern. In H.-P. Trollenier, W. Lenhard & P. Marx (Hrsg.), *Brennpunkte der Gedächtnisforschung* (S. 213-233). Göttingen: Hogrefe.

- Roebers, C. M. & Elischberger, H. B. (2002). Autobiographische Erinnerung bei jungen Kindern: Möglichkeiten und Grenzen bei der Verbesserung ihrer freien Berichte. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie*, 34(1), 2–10.
- Roebers, C. M. & Schneider, W. (2005b). Individual differences in young children's suggestibility: Relations to event memory, language abilities, working memory, and executive functioning. *Cognitive Development*, 20(3), 427-447.
- Roebers, C. M., Moga, N. & Schneider, W. (2001). The role of accuracy motivation on children's and adults' event recall. *Journal of Experimental Child Psychology*, 78 (4), 313-329.
- Sader, M. (1957). *Instruktionsverständnis und Testleitung. Untersuchung über Vorphase und Hauptphase eines psychologischen Prüfversuchs*. Dissertation, Johann Wolfgang Goethe-Universität zu Frankfurt am Main.
- Saywitz, K. & Snyder, L. (1993). Improving children's testimony with preparation. In G. S. Goodman & B.L. Bottoms, *Child victims, child witnesses* (pp. 117-146). NY: Guilford.
- Saywitz, K. J., Geiselman, R. E. & Bornstein, G. K. (1992). Effects of cognitive interviewing and practice on children's recall performance. *Journal of Applied Psychology*, 77(5), 744–756. <https://doi.org/10.1037//0021-9010.77.5.744>
- Schacter, D. L., Norman, K. A. & Koutstaal, W. (1998). The cognitive neuroscience of constructive memory. *Annual review of psychology*, 49, 289–318. <https://doi.org/10.1146/annurev.psych.49.1.289>
- Schade, B. & Harschneck, M. (2000). Die BGH-Entscheidung im Rückblick auf die Wormser Missbrauchsprozesse. Konsequenzen für die Glaubhaftigkeitsbegutachtung aus der Sicht des psychologischen Gutachters und des Strafverteidigers. *Praxis der Rechtspsychologie*, 10, 1, 28-47.
- Schlosser, R. W., Wendt, O., Bhavnani, S. & Nail-Chiwetalu, B. (2006). Use of information-seeking strategies for developing systematic reviews and engaging in evidence-based practice: the application of traditional and comprehensive Pearl Growing. A review. *International Journal of Language & Communication Disorders*, 41, 567-582. doi:10.1080/13682820600742190

- Schneider, W., & Bjorklund, D. F. (1992). Expertise, aptitude, and strategic remembering. *Child Development*, *63*, 461-473.
- Schütz, J. (2011). Kontinuierliche versus diskrete Modelle der Rekognition und des Gedächtnisses. Dissertation, Rheinische Friedrich-Wilhelms-Universität Bonn.
- Schwarz, S. & Roebbers, C. M. (2006). Age differences in the effects of social influence on children's eyewitness performance and their metacognitive monitoring. *Journal of Experimental Child Psychology*, *94*(3), 229-248.
- Scullin, M. H. & Bonner, K. (2006). Theory of mind, inhibitory control, and preschool-age children's suggestibility in different interviewing contexts. *Journal of Experimental Child Psychology*, *93*(2), 120-138.
<https://doi.org/10.1016/j.jecp.2005.09.005>
- Shapiro, L. R. & Purdy, T. L. (2005). Suggestibility and source monitoring errors: blame the interview style, interviewer consistency, and the child's personality. *Applied Cognitive Psychology*, *19*(4), 489–506.
- Sluzenski, J., Newcombe, N. & Ottinger, W. (2004). Changes in reality monitoring and episodic memory in early childhood. *Developmental Science*, *7* (2), 225–245.
- Steffens, M. C. & Mecklenbräuer, S. (2007). False Memories. *Journal of Psychology*, *215*(1), 12-24.
- Steller, M. & Köhnken, G. (1989). Criteria-based statement analysis. In D. C. Raskin (Hrsg.), *Psychological methods in criminal investigation and evidence* (S. 217-245). New York: Springer.
- Steller, M. & Volbert, R. (1999). Wissenschaftliches Gutachten. Forensisch-aussagepsychologische Begutachtung (Glaubwürdigkeitsbegutachtung). *Praxis der Rechtspsychologie*, *9*(2), 46-101.
- Steller, M. (2008). Glaubhaftigkeitsbegutachtung. In R. Volbert und M. Steller (Hrsg.), *Handbuch der Rechtspsychologie* (S. 300-310). Göttingen: Hogrefe.
- Strack, F. (1994). *Zur Psychologie der standardisierten Befragung. Kognitive und kommunikative Prozesse*. Berlin: Springer.

- Strafgesetzbuch (StGB). §153 StGB. [WWW Dokument] URL <https://dejure.org/gesetze/StGB/153.html> [letzter Zugriff 01.02.2018]
- Sugimura, T. (2008). External Source Monitoring in a Real-life Event. Developmental Changes in Ability to Identify Source Persons. *Applied Cognitive Psychology*, 22, 527–539.
- Sutherland, R., Pipe, M.-E., Schick, K., Murray, J. & Gobbo, C. (2003). Knowing in advance: The impact of prior event information on memory and event knowledge. *Journal of Experimental Child Psychology*, 84, 244–263.
- Talwar, V. & Lee, K. (2008). Social and Cognitive Correlates of Children's Lying Behavior. *Child Development*, 79, 866–881.
- Templeton, L. M. & Wilcox, S. A. (2000). A tale of two representations: The misinformation effect and children's developing theory of mind. *Child Development*, 71(2), 402-416. <http://dx.doi.org/10.1111/1467-8624.00153>
- Thierry, K. L. & Pipe, M.-E. (2009). The susceptibility of young preschoolers to source similarity effects: confusing story or video events with reality. *Journal of experimental child psychology*, 102(4), 392–407.
- Thierry, K. L., Goh, C. L., Pipe, M.-E. & Murray, J. (2005). Source recall enhances children's discrimination of seen and heard events. *Journal of experimental psychology. Applied*, 11(1), 33–44.
- Thierry, K. L., Lamb, M. E., Pipe, M.-E. & Spence, M. J. (2010). The flexibility of source monitoring training: Reducing young children's source confusions. *Applied Cognitive Psychology*, 24(5), 626–644.
- Thierry, K. L., Spence, M. J. & Memon, A. (2001). Before misinformation is encountered: Source monitoring decreases child witness suggestibility. *Journal of Cognition and Development*, 2, 1–26.
- Tobey, A. E. & Goodman, G. S. (1992). Children's eyewitness memory: Effects of participation and forensic context. *Child abuse & neglect*, 16(6), 779-796. [https://doi.org/10.1016/0145-2134\(92\)90081-2](https://doi.org/10.1016/0145-2134(92)90081-2)
- Treisman, A. (1964). Selective attention in man. *British Medical Bulletin*, 12, 12–16.

- Tulving, E. (1972). *Organization of memory*. New York, NY: Academic Press.
- Tulving, E. (1983). *Elements of episodic memory*. Oxford: Clarendon Press.
- Tulving, E. & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352–373.
- Tulving, E. & Osler, S. (1968). Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology*, *77*(4), 593-601.
<http://dx.doi.org/10.1037/h0026069>
- Undeutsch, U. (1954). *Die Entwicklung der gerichtropsychologischen Gutachtertätigkeit*. Göttingen: Hogrefe.
- Undeutsch, U. (1989). The development of statement reality analysis. In J. C. Yuille (Hrsg.), *Credibility assessment* (S. 101-119). Dordrecht: Kluwer.
- Velden, M. (1982). *Die Signalentdeckungstheorie in der Psychologie*. Stuttgart: Kohlhammer.
- Volbert, R. (2004). *Beurteilung von Aussagen über Traumata. Erinnerungen und ihre psychologische Bewertung*. Bern: Huber.
- Volbert, R. & Dahle, K. P. (2010). *Forensisch-psychologische Diagnostik im Strafverfahren*. Göttingen: Hogrefe Verlag.
- Volbert, R. & Steller, M. (2014). Is this testimony truthful, fabricated, or based on false memory? *European Psychologist*, *19*(3), 207-220.
- Volbert, R. (2005). Teil 4 Entwicklungspsychologische Grundlagen der Aussagepsychologie. Die Entwicklung von Aussagefähigkeiten. In K.-P. Dahle & R. Volbert (Hrsg.), *Entwicklungspsychologische Aspekte der Rechtspsychologie* (241-257). Göttingen: Hogrefe.
- Volbert, R., Steller, M. & Galow, A. (2010). Das Glaubhaftigkeitsgutachten. In R. Volbert (Hrsg.), *Handbuch der forensischen Psychiatrie* (Vol. 2, S. 623-679). Göttingen: Hogrefe.
- Vrij, A. (2005). Criteria-Based Content Analysis: A Qualitative Review of the First 37 Studies. *Psychology, Public Policy, and Law*, *11*(1), 3–41.
<https://doi.org/10.1037/1076-8971.11.1.3>

- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester: Wiley.
- Vrij, A., Kneller, W. & Mann, S. (2000). The effect of informing liars about Criteria-based content analysis on their ability to deceive CBCA-raters. *Legal and Criminological Psychology, 5(1)*, 57-70.
- Warren, A. M. & McGough, L. S. (1996). Research on Children's Suggestibility. Implications for the investigative interview. *Criminal Justice and behaviour, 2(23)*, 296–303.
- Waterman, A. H., Blades, M. & Spencer, C. (2001). Interviewing Children and Adults: The Effect of Question Format on the Tendency to Speculate. *Applied Cognitive Psychology, 15*, 521-531.
- Wegener, H. & Steller, M. (1987). Psychologische Diagnostik vor Gericht. Methodische und ethische Probleme forensisch psychologischer Diagnostik. *Zeitschrift für Differentielle und Diagnostische Psychologie, 8(103)*, 126.
- Westhoff, K. & Kluck, M.-L. (2014). *Psychologische Gutachten schreiben und beurteilen* (6. Auflage). Heidelberg: Springer.
- Willis, G. B. (2005). *Cognitive interviewing: a tool for improving questionnaire design*. Thousand Oaks: Sage.
- Wrightsmann, L. S. (2001). *Forensic Psychology*. Stanford: Wadsworth.

Anhang

Anhang A Anschreiben Schulen/ Projektbeschreibung



Universität zu Köln

Universität zu Köln • Albertus-Magnus-Platz • 50923 Köln

Human-
wissenschaftliche
Fakultät

Department Psychologie
Pädagogische Psychologie

Liesel Heiermann (M.Sc.)

Telefon +49 221 470-5436
Telefax +49 221 470-6920
liesel.heiermann@uni-koeln.de

Berna Abegg (B.Sc.)

Telefon +4917620921422
bernaabegg@web.de

Sekretariat

Telefon +49 221 470-5438
Telefax +49 221 470-6920
marlies.hess@uni-koeln.de

Köln, 21.05.2014

Liebe Lehrerinnen und Lehrer,

im Rahmen unserer Doktor- bzw. Masterarbeit am Lehrstuhl für Pädagogische Psychologie der Universität zu Köln möchten wir die Gedächtnisentwicklung und -leistung von Grundschulern untersuchen.

Wir möchten anfragen, ob ihre Schule uns dabei unterstützen möchte.

In unserer Untersuchung wird den Schülerinnen und Schülern eine Zirkusgeschichte vorgelesen. Später sollen unterschiedliche Elemente in einem sogenannten Quellenidentifikationstest wieder erkannt werden. So möchten wir genauere Kenntnisse über das kindliche Gedächtnis erhalten.

Die Befragung soll mit Kindern der **4. Klasse im Juni 2014** durchgeführt werden. Zunächst werden die Kinder in 10er Gruppen aufgeteilt und ihnen wird eine Geschichte vorgelesen. Dies dauert ca. 15 Minuten. Eine Woche später werden die Schüler einzeln für ca. 15 Minuten befragt. Wir richten uns dabei selbstverständlich nach den Vorstellungen und Möglichkeiten Ihrer Schule. Uns ist ein großes Anliegen, den Schulalltag möglichst wenig zu stören und den Aufwand für Ihr Kollegium möglichst gering zu halten.

Bei dieser Studie werden keine persönlichen Daten erhoben. Alle Daten werden anonymisiert, so dass Rückschlüsse auf einzelne Schülerinnen und Schüler und individuelle Leistungen nicht möglich sind. Nach Abschluss der Auswertung erhält Ihre Schule gerne einen Ergebnisbericht.

Wir würden uns sehr freuen, wenn Sie sich an unserem Projekt beteiligen. Für Rückfragen oder weitere Informationen steht Ihnen Frau Heiermann gerne zur Verfügung.

Vielen herzlichen Dank und mit freundlichen Grüßen
Prof. Dr. Ellen Aschermann
Liesel Heiermann, M.Sc.
Berna Abegg, B.Sc.

Gronewaldstraße 2
50931 Köln

Anhang B Elterneinverständniserklärung mit Rückschein



Universität zu Köln

Universität zu Köln • Albertus-Magnus-Platz • 50923 Köln

**Human-
wissenschaftliche
Fakultät**

**Department Psychologie
Pädagogische Psychologie**

Liesel Heiermann (M.Sc.)

Telefon +49 221 470-5436
Telefax +49 221 470-6920
liesel.heiermann@uni-koeln.de

Köln, 05.06.2014

Berna Abegg (B.Sc.)

Telefon +4917620921422
bernaabegg@web.de

Liebe Eltern,

im Rahmen unserer Doktor- bzw. Masterarbeit am Lehrstuhl für Pädagogische Psychologie der Universität zu Köln möchten wir die Gedächtnisentwicklung und -leistung von Grundschulern untersuchen.

In unserer Untersuchung wird den Schülerinnen und Schülern eine Zirkusgeschichte vorgelesen. Später sollen unterschiedliche Elemente in einem sogenannten Quellenidentifikationstest wieder erkannt werden. So möchten wir genauere Kenntnisse über das kindliche Gedächtnis erhalten.

Bei dieser Studie werden keine persönlichen Daten erhoben. Alle Daten werden anonymisiert, so dass Rückschlüsse auf einzelne Schülerinnen und Schüler und individuelle Leistungen nicht möglich sind.

Die Schulleitung hat sich bereit erklärt uns zu unterstützen. Sollten Sie damit einverstanden sein, dass ihr Kind an der Untersuchung teilnimmt, geben Sie den unteren Abschnitt bitte spätestens bis zum 02.06.2014 an ihre Klassenlehrerin/ Klassenlehrer zurück.

Wir würden uns über die Teilnahme Ihres Kindes sehr freuen,

Für Rückfragen oder weitere Informationen stehen Ihnen Frau Heiermann und Frau Abegg gerne zur Verfügung.

Vielen herzlichen Dank und mit freundlichen Grüßen

Prof. Dr. Ellen Aschermann
Liesel Heiermann, M.Sc. Psychologie
und Berna Abegg, B.Sc. Psychologie

Sekretariat

Telefon +49 221 470-5438
Telefax +49 221 470-6920
marlies.hess@uni-koeln.de

Gronewaldstraße 2
50931 Köln



Universität zu Köln

Einverständnis Gedächtnistest

Ich bin damit einverstanden, dass mein Kind _____, der Klasse ____
im Juni 2014 an einem Gedächtnistest der Universität zu Köln teilnimmt. Ich bin
darüber informiert worden, dass keine personenbezogenen Daten erhoben werden.

Unterschrift der/des Erziehungsberechtigten

Ort und Datum

Hier abtrennen

Anhang C Stimulusgeschichten A und B

Zirkusgeschichte A

Hallo, ich bin _____. Ich war am letzten Sonntag/ Ich war letzte Woche auch im Zirkus. Aber ich war in einer ganz anderen Vorstellung. Da gab es ganz viel zu sehen. Als alle Leute auf ihren Plätzen saßen, wurde es ganz dunkel. Nur in der Manege schien noch Licht.

Der Zirkusdirektor, in seiner grünen Uniform, ritt auf einem Kamel in der Manege herum. Dabei schwenkte er seinen Hut und begrüßte alle kleinen und großen Leute.

Dann zeigten sechs Bodenturner ihr Können. Sie bauten alle zusammen einen großen Menschenturm und zogen unter Beifall aus der Manege.

Beim Ponybalett gingen die Ponies genau nach der Musik und liefen alle gleichzeitig auf den Hinterbeinen.

Darauf wurde es wieder ganz still in dem großen Zirkuszelt, denn ein Zauberer trat auf. Aus seinem silbernen Zylinderhut holte er auf einmal ein Kaninchen hervor, obwohl der Hut vorher ganz leer war.

Nun liefen zwei kleine Elefanten in die Manege. Die beiden Elefanten machten einen Handstand, so dass ihre Hinterbeine in der Luft waren.

Nach den Elefanten kam die Fahrradnummer. Ein Mädchen balancierte mit einem Fuß auf dem Sattel und mit dem anderen Fuß auf dem Lenker des Fahrrades und fuhr so immer im Kreis herum. Da haben alle geklatscht.

Nach der Pause holte ein Clown in einem gestreiften Kostüm und mit ganz großen Schuhen die Leute wieder auf ihre Plätze zurück. Der Clown spielte dabei auf einer sehr großen Tuba.

Jetzt war der Käfig für die Tiger aufgebaut. Die Tiger stürmten in die Manege und einer von ihnen balancierte auf einer Laufkugel. Das fand ich sehr aufregend.

Trommelwirbel kündigten den Auftritt der Trapezkünstler an. Ein Mann und eine Frau kletterten bis hoch unter das Dach des Zirkuszeltens. Mir wurde schon vom Zusehen ganz schwindelig. Die Frau flog hoch oben durch die Luft und der Mann an der Schaukel schnappte sie aus der Luft, gerade als es so aussah, als ob sie abstürzen würde.

Zum Schluss kam noch die Bärennummer. Das Tollste war, als ein Braunbär auf einem Dreirad in der Manege herumfuhr. Da haben alle geklatscht.

Wie schade, dass die Vorstellung so schnell zu Ende war. Ich hätte gerne noch viel länger zugehört. Nach einem langen Applaus gingen alle schließlich nach Hause.

Zirkusgeschichte B

Hallo, ich bin _____. Ich war am letzten Sonntag/ Ich war letzte Woche auch im Zirkus. Aber ich war in einer ganz anderen Vorstellung. Da gab es ganz viel zu sehen. Als alle Leute auf ihren Plätzen saßen, wurde es ganz dunkel. Nur in der Manege schien noch Licht.

Da fing die Musik an zu spielen. Ein großer brauner Bär und ein Mädchen kamen in die Manege. Sie begrüßten die Leute und tanzten zur Musik.

Plötzlich kam ein großes braunes Kamel in die Manege getrabt. Dabei wackelte es mit seinen Höckern. Nach einer Runde legte es sich hin und schaute wie ich beim Kunstreiten zu.

Beim Kunstreiten stand ein Mädchen mit ausgestreckten Armen auf dem Rücken eines geschmückten Pferdes, das schnell im Kreis galoppierte. Da haben die Zuschauer alle geklatscht.

Dann liefen zwei Turner herein. Sie fingen an, die tollsten Dinge zu machen. Ein Turner machte zuerst einen Handstand auf den Händen des anderen. Dann beugte er seinen Körper soweit, dass es schien, als ob er über dem anderen Turner schweben würde. Es sah so einfach aus, aber die beiden mussten bestimmt lange dafür üben.

Im Anschluss daran wurde ein Käfig für die Tiger aufgebaut. Die Tiger stürmten herein und einer von ihnen sprang von einem Hocker mitten durch einen Reifen, den der Dompteur hochhielt. Das war wirklich aufregend.

Nach der Pause rief ein Clown auf ganz hohen Stelzen die Leute zurück in das Zirkuszelt. Es sah so aus, als ob er ganz, ganz lange Beine hätte.

Im Zirkuszelt war hoch über uns ein Seil gespannt. Auf dem Seil balancierte ein Mann mit einer schwarzen Stange. Eine Frau machte auf seinem Kopf einen Handstand auf einer Hand. Das sah ganz schön gefährlich aus.

Danach trotteten zwei große Elefanten in die Manege. Die Elefanten liefen über ganz kleine Podeste, ohne auch nur einmal danebenzutreten.

Nachdem es einen Moment dunkel im Zelt war, stand plötzlich ein Zauberer mit einem großen blauen Umhang mitten in der Manege. Er konnte uns tolle Tricks mit Spielkarten zeigen.

Die Einradnummer, die zum Schluss kam, hat mir gut gefallen. Ein Junge fuhr auf einem hohen Einrad in großen Kreisen durch die Manege. Nicht einmal sein Hut verrutschte dabei.

Wie schade, dass die Vorstellung so schnell vorbei war, ich hätte gerne noch viel länger zugesehen. Nach dem langen Applaus gingen schließlich alle nach Hause.

Anhang D Interviewleitfäden für die vier Bedingungen

Interviewleitfaden: Itemrecognition/Quellenidentifikation Bedingung **Forcierung-/Strafe-**:

1. Kind abholen (Wenn Klassenliste vorhanden, Namen von Liste streichen): „Ich möchte jetzt gerne mit ... sprechen.“
2. Begrüßung: „Hallo ... ich bin die ...und Du bist...?“
3. Beziehung zum Kind aufbauen:

Mögliche Themen: Lieblingsspiel, Haustier, Schule

4. Anlass der Untersuchung erläutern:

NICHT Worte wie „Test“ oder „Experiment“ verwenden sondern: „Wir wollen untersuchen, wie gut sich Kinder in deinem Alter erinnern können.“

5. **Freier Bericht:** „Ich habe hier ein Tonbandgerät das mache ich jetzt kurz an und zeichne deine Stimme auf. Ich spreche jetzt was drauf...(Vormachen) und das kann ich mir auch wieder anhören. Die Stimme hört sich abgespielt immer anders an. Ich weiß, dass du letzte Woche 2 Zirkusgeschichten von zwei Frauen gehört hast. Ich habe die Geschichten nicht gehört und weiß nicht, was in den Geschichten vorkam. Woran kannst du dich noch erinnern? Fällt dir noch etwas ein?“

Erstes Item

6. („Du hast letzte Woche zwei Zirkusgeschichten von zwei Frauen gehört und da gab es viel Zirkusnummern. Du sollst dich jetzt erinnern, welche in den Geschichten vorkamen und welche nicht.) Ich habe hier eine Liste mit Zirkusnummern, diese lese ich dir vor und du sollst mir sagen, ob diese Nummer in einer der beiden Geschichten vorkam. Wenn du dir unsicher bist, kannst du auch „Ich weiß nicht“ sagen. Hast du verstanden was ich erklärt habe? Die erste Nummer auf meiner Liste heißt: XYZ ... Kam das in einer der beiden Geschichten vor?“

7. „Jetzt möchte ich gerne wissen, welche Erzählerin diese Zirkusnummer erzählt hat. Ich zeige dir zwei Fotos und du sollst mir sagen, wer von den beiden die Zirkusnummer erzählt hat. Wenn du dir unsicher bist, kannst du auch „Ich weiß nicht“ sagen. Hast du verstanden was ich erklärt habe? Dann sag mir bitte welche Frau die Nummer erzählt hat.“

Zweites und jedes weitere Item

8. „Die zweite/dritte... Zirkusnummer auf meiner Liste heißt XYZ. Kam das in einer der beiden Geschichten vor? Wenn du dir unsicher bist, kannst du auch „Ich weiß nicht“ sagen.“
9. „Jetzt möchte ich wieder wissen welche Erzählerin die Zirkusnummer erzählt hat. (Zeigen auf die Fotos). Du sollst mir jetzt sagen, wer von den beiden die Zirkusnummer erzählt hat. Wenn du dir unsicher bist, kannst du auch „Ich weiß nicht“ sagen.“

Abschluss

10. „Sag mal Welche Vorleserin fandest du eigentlich netter? Und welche Geschichte hat dir besser gefallen?“

11. „Das hast du wirklich super gemacht. Vielen lieben Dank. Wenn wir alle Kinder eurer Klasse befragt haben, bekommt jeder ein tolles Geschenk.“

Interviewleitfaden: Itemrecognition/Quellenidentifikation Bedingung **Forcierung-/Strafe+**

1. Kind abholen (Wenn Klassenliste vorhanden, Namen von Liste streichen): „Ich möchte jetzt gerne mit ... sprechen.“

2. Begrüßung: „Hallo ... ich bin die ...und Du bist...?“

3. Beziehung zum Kind aufbauen:

Mögliche Themen: Lieblingsspiel, Haustier, Schule

4. Anlass der Untersuchung erläutern:

NICHT Worte wie „Test“ oder „Experiment“ verwenden sondern: „Wir wollen untersuchen, wie gut sich Kinder in deinem Alter erinnern können.“

5. **Freier Bericht:** „Ich habe hier ein Tonbandgerät das mache ich jetzt kurz an und zeichne deine Stimme auf. Ich spreche jetzt was drauf...(Vormachen) und das kann ich mir auch wieder anhören. Die Stimme hört sich abgespielt immer anders an. Ich weiß, dass du letzte Woche 2 Zirkusgeschichten von zwei Frauen gehört hast. Ich habe die Geschichten nicht gehört und weiß nicht, was in den Geschichten vorkam. Woran kannst du dich noch erinnern? Fällt dir noch etwas ein?“

Erstes Item

6. („Du hast letzte Woche zwei Zirkusgeschichten von zwei Frauen gehört und da gab es viele Zirkusnummern. Du sollst dich jetzt erinnern, welche in den Geschichten vorkamen und welche nicht.) Ich habe hier eine Liste mit Zirkusnummern, diese lese ich dir vor und du sollst mir sagen, ob diese Nummer in einer der beiden Geschichten vorkam. Wenn du dir unsicher bist, kannst du auch „Ich weiß nicht“ sagen. Du bekommst nun 46 bunte Büroklammern und für eine falsche Antwort musst du eine davon abgeben. Wenn du „Ich weiß nicht“ sagst was passiert dann mit den Büroklammern? (richtige Antwort: Es passiert nichts) Hast du verstanden was ich erklärt habe? Die erste Nummer auf meiner Liste heißt: XYZ ... Kam das in einer der beiden Geschichten vor?“

7. „Jetzt möchte ich gerne wissen, welche Erzählerin diese Zirkusnummer erzählt hat. Ich zeige dir zwei Fotos und du sollst mir sagen, wer von den beiden die Zirkusnummer erzählt hat. Wenn du dir unsicher bist, kannst du auch „Ich weiß nicht“ sagen. Wie vorhin musst du für eine falsche Antwort eine Büroklammer abgeben. Hast du verstanden was ich erklärt habe? Dann sag mir bitte, welche Frau die Nummer erzählt hat.“

Zweites und jedes weitere Item

8. „Die zweite/dritte... Zirkusnummer auf meiner Liste heißt XYZ. Kam das in einer der beiden Geschichten vor? Wenn du dir unsicher bist, kannst du auch „Ich weiß nicht“ sagen. Du musst aber für eine falsche Antwort eine der Büroklammern abgeben.“

9. „Jetzt möchte ich wieder wissen welche Erzählerin die Zirkusnummer erzählt hat. (Zeigen auf die Fotos) Du sollst mir jetzt sagen, wer von den beiden die Zirkusnummer erzählt hat. Wenn du dir

unsicher bist, kannst du auch „Ich weiß nicht“ sagen. Du musst aber für eine falsche Antwort eine der Büroklammern abgeben.“

Abschluss

10. „Sag mal Welche Vorleserin fandest du eigentlich netter? Und welche Geschichte hat dir besser gefallen?“

11. „Das hast du wirklich super gemacht. Vielen lieben Dank. Wenn wir alle Kinder eurer Klasse befragt haben, bekommt jeder ein tolles Geschenk.“

Interviewleitfaden: Itemerkennung /Quellenidentifikation Bedingung **Forcierung+/Strafe-**

1. Kind abholen (Wenn Klassenliste vorhanden, Namen von Liste streichen): „Ich möchte jetzt gerne mit ... sprechen.“

2. Begrüßung: „Hallo ... ich bin die ...und Du bist...?“

3. Beziehung zum Kind aufbauen:

Mögliche Themen: Lieblingsspiel, Haustier, Schule

4. Anlass der Untersuchung erläutern:

NICHT Worte wie „Test“ oder „Experiment“ verwenden sondern: „Wir wollen untersuchen, wie gut sich Kinder in deinem Alter erinnern können.“

5. **Freier Bericht:** „Ich habe hier ein Tonbandgerät das mache ich jetzt kurz an und zeichne deine Stimme auf. Ich spreche jetzt was drauf...(Vormachen) und das kann ich mir auch wieder anhören. Die Stimme hört sich abgespielt immer anders an. Ich weiß, dass du letzte Woche 2 Zirkusgeschichten von zwei Frauen gehört hast. Ich habe die Geschichten nicht gehört und weiß nicht, was in den Geschichten vorkam. Woran kannst du dich noch erinnern? Fällt dir noch etwas ein?“

Erstes Item

6. („Du hast letzte Woche zwei Zirkusgeschichten von zwei Frauen gehört und da gab es viele Zirkusnummern. Du sollst dich jetzt erinnern, welche in den Geschichten vorkamen und welche nicht.) Ich habe hier eine Liste mit Zirkusnummern, diese lese ich dir vor und du sollst mir sagen, ob diese Nummer in einer der beiden Geschichten vorkam. Wenn du dir unsicher bist, sollst du raten. Hast du verstanden was ich erklärt habe? Die erste Nummer auf meiner Liste heißt: XYZ ... Kam das in einer der beiden Geschichten vor?“

Hinweise/ mögliche verbale Ergänzung vom VL:

Kind zögert, sagt „Ich weiß es nicht“: „Versuche dich bitte genau zu erinnern und rate, ob die Zirkusnummer vorkam bzw. in welcher Geschichte die Zirkusnummer vorkam.“

7. „Jetzt möchte ich gerne wissen, welche Erzählerin diese Zirkusnummer erzählt hat. Ich zeige dir zwei Fotos und du sollst mir sagen, wer von den beiden die Zirkusnummer erzählt hat. Wenn du dir unsicher bist, sollst du raten. Hast du verstanden was ich erklärt habe? Dann sag mir bitte welche Frau die Nummer erzählt hat.“

Hinweis/ mögliche verbale Ergänzung vom VL:

- Kind zögert, sagt ich weiß es nicht: „Versuche dich bitte genau zu erinnern und rate, welche Erzählerin diese Zirkusnummer erzählt hat.“

Zweites und jedes weitere Item

8. „Die zweite/dritte... Zirkusnummer auf meiner Liste heißt XYZ. Du sollst mir jetzt sagen ob diese Zirkusnummer in einer der beiden Geschichten vorkam. Wenn du dir unsicher bist, sollst du raten.“

9. „Jetzt möchte ich wieder wissen welche Erzählerin die Zirkusnummer erzählt hat. (Zeigen auf die Fotos) Du sollst mir jetzt sagen, wer von den beiden die Zirkusnummer erzählt hat. Wenn du dir unsicher bist, sollst du raten.“

Abschluss

10. „Sag mal, welche Vorleserin fandest du eigentlich netter? Und welche Geschichte hat dir besser gefallen?“

11. „Das hast du wirklich super gemacht. Vielen lieben Dank. Die Büroklammern brauchen wir leider noch. Wenn wir alle Kinder eurer Klasse befragt haben, bekommt jeder ein tolles Geschenk.“

Interviewleitfaden: Itemerkennung /Quellenidentifikation Bedingung **Forcierung+/Strafe+**

1. Kind abholen (Wenn Klassenliste vorhanden, Namen von Liste streichen): „Ich möchte jetzt gerne mit ... sprechen.“

2. Begrüßung: „Hallo ... ich bin die ...und Du bist...?“

3. Beziehung zum Kind aufbauen:

Mögliche Themen: Lieblingsspiel, Haustier, Schule

4. Anlass der Untersuchung erläutern:

NICHT Worte wie „Test“ oder „Experiment“ verwenden sondern: „Wir wollen untersuchen, wie gut sich Kinder in deinem Alter erinnern können.“

5. **Freier Bericht:** „Ich habe hier ein Tonbandgerät das mache ich jetzt kurz an und zeichne deine Stimme auf. Ich spreche jetzt was drauf...(Vormachen) und das kann ich mir auch wieder anhören. Die Stimme hört sich abgespielt immer anders an. Ich weiß, dass du letzte Woche 2 Zirkusgeschichten von zwei Frauen gehört hast. Ich habe die Geschichten nicht gehört und weiß nicht, was in den Geschichten vorkam. Woran kannst du dich noch erinnern? Fällt dir noch etwas ein?“

Erstes Item

6. („Du hast letzte Woche zwei Zirkusgeschichten von zwei Frauen gehört und da gab es viele Zirkusnummern. Du sollst dich jetzt erinnern, welche in den Geschichten vorkamen und welche nicht.) Ich habe hier eine Liste mit Zirkusnummern, diese lese ich dir vor und du sollst mir sagen, ob diese Nummer in einer der beiden Geschichten vorkam. Wenn du dir unsicher bist, sollst du raten. Hast du verstanden was ich erklärt habe? Die erste Nummer auf meiner Liste heißt: XYZ ... Kam das in einer der beiden Geschichten vor?“

Hinweise/ mögliche verbale Ergänzung vom VL:

Kind zögert, sagt „Ich weiß es nicht“: „Versuche dich bitte genau zu erinnern und rate, ob die Zirkusnummer vorkam bzw. in welcher Geschichte die Zirkusnummer vorkam.“

7. „Jetzt möchte ich gerne wissen, welche Erzählerin diese Zirkusnummer erzählt hat. Ich zeige dir zwei Fotos und du sollst mir sagen, wer von den beiden die Zirkusnummer erzählt hat. Wenn du dir unsicher bist, sollst du raten. Hast du verstanden was ich erklärt habe? Dann sag mir bitte welche Frau die Nummer erzählt hat.“

Hinweis/ mögliche verbale Ergänzung vom VL:

- Kind zögert, sagt ich weiß es nicht: „Versuche dich bitte genau zu erinnern und rate, welche Erzählerin diese Zirkusnummer erzählt hat.“

Zweites und jedes weitere Item

8. „Die zweite/dritte... Zirkusnummer auf meiner Liste heißt XYZ. Du sollst mir jetzt sagen ob diese Zirkusnummer in einer der beiden Geschichten vorkam. Wenn du dir unsicher bist, sollst du raten. Du musst aber für eine falsche Antwort eine der Büroklammern abgeben,“

9. „Jetzt möchte ich wieder wissen welche Erzählerin die Zirkusnummer erzählt hat. (Zeigen auf die Fotos). Du sollst mir jetzt sagen, wer von den beiden die Zirkusnummer erzählt hat. Wenn du dir unsicher bist, sollst du raten. Du musst aber für eine falsche Antwort eine der Büroklammern abgeben.“

Abschluss

10. „Sag mal, welche Vorleserin fandest du eigentlich netter? Und welche Geschichte hat dir besser gefallen?“

11. „Das hast du wirklich super gemacht. Vielen lieben Dank. Die Büroklammern brauchen wir leider noch. Wenn wir alle Kinder eurer Klasse befragt haben, bekommt jeder ein tolles Geschenk.“

Anhang E Standardisierte Interviewantworten

Weitere Ergänzungen zum Interviewleitfaden

- Zum freien Bericht: „Ich möchte gerne, dass du mir alles erzählst woran du dich noch erinnern kannst. Auf die Vollständigkeit kommt es nicht an.“
- Beim ersten Item einführen: „Manche Nummern kamen vor und manche auch nicht. Die waren in ganz anderen Geschichten.“
- Beim ersten Item einführen: „Versuche so viele Büroklammern wie nur möglich zu behalten.“
- Zum Motivieren in den *Strafe* Bedingungen: „Schau mal, so viele Büroklammern hast du schon.“
- Wenn sich das Kind beschwert, dass es Klammern abgeben muss: „Hmm, das ist jetzt wirklich blöd.“
- Kind fragt warum das Diktaphon benutzt wurde: „Das brauche ich, weil ich nicht alles mitschreiben kann.“
- Beim ersten benutzten des Diktaphons, vorführen: „Sag doch mal deinen Namen und deine Klasse hinein, dann hören wir uns das an...“

Anhang F QI-Fragebögen

Forcierung-/Strafe- Version 1

Datum: _____ Name: _____
 Klasse: _____ Alter: _____

Schule: _____
 Namen Vorleser: 1=Maren / 2=Anna, Lisa

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Ich weiß nicht“	Vorleser 1	Vorleser 2	„Ich weiß nicht“
6=Tiger balancierte auf Laufkugel						
18=Schlangenbeschwörer mit Flöte						
10=Tiger sprang durch Reifen						
21= Schwertschlucker schluckt Schwert						
14= Zauberer mit Kartentricks						
17=Feuerspucker spuckt Flamme						
9= Bodenturner machte Handstand						
3= Zauberer holte Kaninchen						
28= Löwe brüllte vor Sprung auf Hocker						
5= Clown spielte Tuba						
8= Mädchen stand auf Pferd						
27= Mann auf Seil jonglierte mit Bällen						
15= Seehund mit Ball						
1= Bodenturner bauten Menschenturm						
4= Elefanten machten Handstand						
26= Mann auf Wippe, Frau in der Luft						
12= Mann auf Seil, Frau Handstand						
19= Affen spielten mit Ball						
11= Clown mit Stelzen rief Leute						
2= Ponyballett mit Musik						
25= Clown machte Seifenblasen						
7= Mann schnappte Frau aus Luft						
22= Elefanten liefen in langen Reihen						
20=Entfesselungskünstler befreite sich						
16= Tauben auf Treppe						
13= Elefanten auf Podesten						
24= Ponies über Wassergraben						
23= Zauberer und Blumenstrauß						

1-7 A / 8-14 B / 15-28 Distraktoritens

Welche Vorleserin fandest Du netter? _____

Forcierung-/Strafe- Version 2

Datum:
Klasse:
Interviewer:

Name:
Alter:

Schule:

Namen Vorleser: 1=Maren / 2=Anna, Lisa

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Ich weiß nicht“	Vorleser 1	Vorleser 2	„Ich weiß nicht“
23= Zauberer und Blumenstrauß						
18=Schlangen- beschwörerin mit Flöte						
9= Bodenturner machte Handstand						
13= Elefanten auf Podesten						
19= Affen spielten mit Ball						
17=Feuerspucker spuckt Flamme						
28= Löwe brüllte vor Sprung auf Hocker						
12= Mann auf Seil, Frau Handstand						
25= Clown machte Seifenblasen						
3= Zauberer holte Kaninchen						
22= Elefanten liefen in langen Reihen						
1= Bodenturner bauten Menschenturm						
16= Tauben auf Treppe						
26= Mann auf Wippe, Frau in der Luft						
24= Ponies über Wassergraben						
8= Mädchen stand auf Pferd						
20=Entfesselungs- künstler befreite sich						
14= Zauberer mit Kartentricks						
27= Mann auf Seil jonglierte mit Bällen						
6=Tiger balancierte auf Laufkugel						
7= Mann schnappte Frau aus Luft						
2= Ponyballett mit Musik						
4= Elefanten machten Handstand						
11= Clown mit Stelzen rief Leute						
15= Seehund mit Ball						
5= Clown spielte Tuba						
10=Tiger sprang durch Reifen						
21= Schwertschlucker schluckt Schwert						

1-7 A / 8-14 B / 15-28 Distraktoritems

Welche Vorleserin fandest Du netter? _____

Forcierung-/Strafe- Version 3

Datum:
Klasse:
Interviewer:

Name:
Alter:

Schule:

Namen Vorleser: 1=Maren / 2=Anna, Lisa

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Ich weiß nicht“	Vorleser 1	Vorleser 2	„Ich weiß nicht“
19= Affen spielten mit Ball						
16= Tauben auf Treppe						
14= Zauberer mit Kartentricks						
9= Bodenturner machte Handstand						
23= Zauberer und Blumenstrauß						
20=Entfesselungskünstler befreite sich						
27= Mann auf Seil jonglierte mit Bällen						
7= Mann schnappte Frau aus Luft						
2= Ponyballett mit Musik						
8= Mädchen stand auf Pferd						
4= Elefanten machten Handstand						
18=Schlangenbeschwörerin mit Flöte						
21= Schwertschlucker schluckt Schwert						
11= Clown mit Stelzen rief Leute						
25= Clown machte Seifenblasen						
12= Mann auf Seil, Frau Handstand						
10=Tiger sprang durch Reifen						
26= Mann auf Wippe, Frau in der Luft						
22= Elefanten liefen in langen Reihen						
13= Elefanten auf Podesten						
15= Seehund mit Ball						
28= Löwe brüllte vor Sprung auf Hocker						
6=Tiger balancierte auf Laufkugel						
5= Clown spielte Tuba						
1= Bodenturner bauten Menschenturm						
3= Zauberer holte Kaninchen						
24= Ponies über Wassergraben						
17=Feuerspucker spuckt Flamme						

1-7 A / 8-14 B / 15-28 Distraktoritems

Welche Vorleserin fandest Du netter? _____

Forcierung-/Strafe- Version 4

Datum:
Klasse:
Interviewer:

Name:
Alter:

Schule:

Namen Vorleser: 1=Maren / 2=Anna, Lisa

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Ich weiß nicht“	Vorleser 1	Vorleser 2	„Ich weiß nicht“
1= Bodenturner bauten Menschenturm						
9= Bodenturner machte Handstand						
16= Tauben auf Treppe						
28= Löwe brüllte vor Sprung auf Hocker						
20=Entfesselungskünstler befreite sich						
27= Mann auf Seil jonglierte mit Bällen						
18=Schlangenbeschwörerin mit Flöte						
13= Elefanten auf Podesten						
25= Clown machte Seifenblasen						
15= Seehund mit Ball						
8= Mädchen stand auf Pferd						
21= Schwertschlucker schluckt Schwert						
24= Ponies über Wassergraben						
3= Zauberer holte Kaninchen						
14= Zauberer mit Kartentricks						
7= Mann schnappte Frau aus Luft						
4= Elefanten machten Handstand						
6=Tiger balancierte auf Laufkugel						
26= Mann auf Wippe, Frau in der Luft						
2= Ponyballett mit Musik						
10=Tiger sprang durch Reifen						
12= Mann auf Seil, Frau Handstand						
19= Affen spielten mit Ball						
23= Zauberer und Blumenstrauß						
17=Feuerspucker spuckt Flamme						
11= Clown mit Stelzen rief Leute						
5= Clown spielte Tuba						
22= Elefanten liefen in langen Reihen						

1-7 A / 8-14 B / 15-28 Distraktoritems

Welche Vorleserin fandest Du netter? _____

Forcierung-/Strafe+ Version 1

Datum: _____ Name: _____
 Klasse: _____ Alter: _____
 Interviewer: _____

Schule: _____
 Namen Vorleser: 1=Maren / 2=Anna, Lisa

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Ich weiß nicht“	Vorleser 1	Vorleser 2	„Ich weiß nicht“
6=Tiger balancierte auf Laufkugel						
18=Schlangen- beschwörerin mit Flöte						
10=Tiger sprang durch Reifen						
21= Schwertschlucker schluckt Schwert						
14= Zauberer mit Kartentricks						
17=Feuerspucker spuckt Flamme						
9= Bodenturner machte Handstand						
3= Zauberer holte Kaninchen						
28= Löwe brüllte vor Sprung auf Hocker						
5= Clown spielte Tuba						
8= Mädchen stand auf Pferd						
27= Mann auf Seil jonglierte mit Bällen						
15= Seehund mit Ball						
1= Bodenturner bauten Menschenturm						
4= Elefanten machten Handstand						
26= Mann auf Wippe, Frau in der Luft						
12= Mann auf Seil, Frau Handstand						
19= Affen spielten mit Ball						
11= Clown mit Stelzen rief Leute						
2= Ponyballett mit Musik						
25= Clown machte Seifenblasen						
7= Mann schnappte Frau aus Luft						
22= Elefanten liefen in langen Reihen						
20=Entfesselungs-künstler befreite sich						
16= Tauben auf Treppe						
13= Elefanten auf Podesten						
24= Ponies über Wassergraben						
23= Zauberer und Blumenstrauß						

1-7 A / 8-14 B / 15-28 Distraktoritems

Welche Vorleserin fandest Du netter? _____

Forcierung-/Strafe+ Version 2

Datum: _____ Name: _____
 Klasse: _____ Alter: _____
 Interviewer: _____

Schule: _____
 Namen Vorleser: 1=Maren / 2=Anna, Lisa

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Ich weiß nicht“	Vorleser 1	Vorleser 2	„Ich weiß nicht“
23= Zauberer und Blumenstrauß						
18=Schlangen- beschwörerin mit Flöte						
9= Bodenturner machte Handstand						
13= Elefanten auf Podesten						
19= Affen spielten mit Ball						
17=Feuerspucker spuckt Flamme						
28= Löwe brüllte vor Sprung auf Hocker						
12= Mann auf Seil, Frau Handstand						
25= Clown machte Seifenblasen						
3= Zauberer holte Kaninchen						
22= Elefanten liefen in langen Reihen						
1= Bodenturner bauten Menschenturm						
16= Tauben auf Treppe						
26= Mann auf Wippe, Frau in der Luft						
24= Ponies über Wassergraben						
8= Mädchen stand auf Pferd						
20=Entfesselungs- künstler befreite sich						
14= Zauberer mit Kartentricks						
27= Mann auf Seil jonglierte mit Bällen						
6=Tiger balancierte auf Laufkugel						
7= Mann schnappte Frau aus Luft						
2= Ponyballett mit Musik						
4= Elefanten machten Handstand						
11= Clown mit Stelzen rief Leute						
15= Seehund mit Ball						
5= Clown spielte Tuba						
10=Tiger sprang durch Reifen						
21= Schwertschlucker schluckt Schwert						

1-7 A / 8-14 B / 15-28 Distraktoritems

Welche Vorleserin fandest Du netter? _____

Forcierung-/Strafe+ Version 3

Datum: _____ Name: _____
 Klasse: _____ Alter: _____
 Interviewer: _____

Schule: _____
 Namen Vorleser: 1=Maren / 2=Anna, Lisa

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Ich weiß nicht“	Vorleser 1	Vorleser 2	„Ich weiß nicht“
19= Affen spielten mit Ball						
16= Tauben auf Treppe						
14= Zauberer mit Kartentricks						
9= Bodenturner machte Handstand						
23= Zauberer und Blumenstrauß						
20=Entfesselungskünstler befreite sich						
27= Mann auf Seil jonglierte mit Bällen						
7= Mann schnappte Frau aus Luft						
2= Ponyballett mit Musik						
8= Mädchen stand auf Pferd						
4= Elefanten machten Handstand						
18=Schlangenbeschwörerin mit Flöte						
21= Schwertschlucker schluckt Schwert						
11= Clown mit Stelzen rief Leute						
25= Clown machte Seifenblasen						
12= Mann auf Seil, Frau Handstand						
10=Tiger sprang durch Reifen						
26= Mann auf Wippe, Frau in der Luft						
22= Elefanten liefen in langen Reihen						
13= Elefanten auf Podesten						
15= Seehund mit Ball						
28= Löwe brüllte vor Sprung auf Hocker						
6=Tiger balancierte auf Laufkugel						
5= Clown spielte Tuba						
1= Bodenturner bauten Menschenturm						
3= Zauberer holte Kaninchen						
24= Ponies über Wassergraben						
17=Feuerspucker spuckt Flamme						

1-7 A / 8-14 B / 15-28 Distraktoritems

Welche Vorleserin fandest Du netter? _____

Forcierung-/Strafe+ Version 4

Datum:
Klasse:
Interviewer:

Name:
Alter:

Schule:

Namen Vorleser: 1=Maren / 2=Anna, Lisa

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Ich weiß nicht“	Vorleser 1	Vorleser 2	„Ich weiß nicht“
1= Bodenturner bauten Menschenturm						
9= Bodenturner machte Handstand						
16= Tauben auf Treppe						
28= Löwe brüllte vor Sprung auf Hocker						
20=Entfesselungskünstler befreite sich						
27= Mann auf Seil jonglierte mit Bällen						
18=Schlangenbeschwörerin mit Flöte						
13= Elefanten auf Podesten						
25= Clown machte Seifenblasen						
15= Seehund mit Ball						
8= Mädchen stand auf Pferd						
21= Schwertschlucker schluckt Schwert						
24= Ponies über Wassergraben						
3= Zauberer holte Kaninchen						
14= Zauberer mit Kartentricks						
7= Mann schnappte Frau aus Luft						
4= Elefanten machten Handstand						
6=Tiger balancierte auf Laufkugel						
26= Mann auf Wippe, Frau in der Luft						
2= Ponyballett mit Musik						
10=Tiger sprang durch Reifen						
12= Mann auf Seil, Frau Handstand						
19= Affen spielten mit Ball						
23= Zauberer und Blumenstrauß						
17=Feuerspucker spuckt Flamme						
11= Clown mit Stelzen rief Leute						
5= Clown spielte Tuba						
22= Elefanten liefen in langen Reihen						

1-7 A / 8-14 B / 15-28 Distraktoritems

Welche Vorleserin fandest Du netter? _____

Forcierung+/Strafe- Version 1

Datum: _____ Name: _____

Klasse: _____ Alter: _____

Interviewer: _____

Schule: _____

Namen Vorleser: _____

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Geraten“	Vorleser 1	Vorleser 2	„Geraten“
6=Tiger balancierte auf Laufkugel						
18=Schlangen- beschwörerin mit Flöte						
10=Tiger sprang durch Reifen						
21= Schwertschlucker schluckt Schwert						
14= Zauberer mit Kartentricks						
17=Feuerspucker spuckt Flamme						
9= Bodenturner machte Handstand						
3= Zauberer holte Kaninchen						
28= Löwe brüllte vor Sprung auf Hocker						
5= Clown spielte Tuba						
8= Mädchen stand auf Pferd						
27= Mann auf Seil jonglierte mit Bällen						
15= Seehund mit Ball						
1= Bodenturner bauten Menschenturm						
4= Elefanten machten Handstand						
26= Mann auf Wippe, Frau in der Luft						
12= Mann auf Seil, Frau Handstand						
19= Affen spielten mit Ball						
11= Clown mit Stelzen rief Leute						
2= Ponyballett mit Musik						
25= Clown machte Seifenblasen						
7= Mann schnappte Frau aus Luft						
22= Elefanten liefen in langen Reihen						
20=Entfesselungs- künstler befreite sich						
16= Tauben auf Treppe						
13= Elefanten auf Podesten						
24= Ponies über Wassergraben						
23= Zauberer und Blumenstrauß						

1-7 A / 8-14 B / 15-28 Distraktoritens

Welche Vorleserin fandest Du netter? _____

Forcierung+/Strafe- Version 2

Datum: _____ Name: _____

Klasse: _____ Alter: _____

Interviewer: _____

Schule: _____

Namen Vorleser: _____

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Geraten“	Vorleser 1	Vorleser 2	„Geraten“
23= Zauberer und Blumenstrauß						
18=Schlangen- beschwörerin mit Flöte						
9= Bodenturner machte Handstand						
13= Elefanten auf Podesten						
19= Affen spielten mit Ball						
17=Feuerspucker spuckt Flamme						
28= Löwe brüllte vor Sprung auf Hocker						
12= Mann auf Seil, Frau Handstand						
25= Clown machte Seifenblasen						
3= Zauberer holte Kaninchen						
22= Elefanten liefen in langen Reihen						
1= Bodenturner bauten Menschenturm						
16= Tauben auf Treppe						
26= Mann auf Wippe, Frau in der Luft						
24= Ponies über Wassergraben						
8= Mädchen stand auf Pferd						
20=Entfesselungs- künstler befreite sich						
14= Zauberer mit Kartentricks						
27= Mann auf Seil jonglierte mit Bällen						
6=Tiger balancierte auf Laufkugel						
7= Mann schnappte Frau aus Luft						
2= Ponyballett mit Musik						
4= Elefanten machten Handstand						
11= Clown mit Stelzen rief Leute						
15= Seehund mit Ball						
5= Clown spielte Tuba						
10=Tiger sprang durch Reifen						
21= Schwertschlucker schluckt Schwert						

1-7 A / 8-14 B / 15-28 Distraktoritens

Welche Vorleserin fandest Du netter? _____

Forcierung+/Strafe- Version 3

Datum: _____ Name: _____

Klasse: _____ Alter: _____

Interviewer: _____

Schule: _____

Namen Vorleser: _____

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Geraten“	Vorleser 1	Vorleser 2	„Geraten“
19= Affen spielten mit Ball						
16= Tauben auf Treppe						
14= Zauberer mit Kartentricks						
9= Bodenturner machte Handstand						
23= Zauberer und Blumenstrauß						
20=Entfesselungskünstler befreite sich						
27= Mann auf Seil jonglierte mit Bällen						
7= Mann schnappte Frau aus Luft						
2= Ponyballett mit Musik						
8= Mädchen stand auf Pferd						
4= Elefanten machten Handstand						
18=Schlangenbeschwörerin mit Flöte						
21= Schwertschlucker schluckt Schwert						
11= Clown mit Stelzen rief Leute						
25= Clown machte Seifenblasen						
12= Mann auf Seil, Frau Handstand						
10=Tiger sprang durch Reifen						
26= Mann auf Wippe, Frau in der Luft						
22= Elefanten liefen in langen Reihen						
13= Elefanten auf Podesten						
15= Seehund mit Ball						
28= Löwe brüllte vor Sprung auf Hocker						
6=Tiger balancierte auf Laufkugel						
5= Clown spielte Tuba						
1= Bodenturner bauten Menschenturm						
3= Zauberer holte Kaninchen						
24= Ponies über Wassergraben						
17=Feuerspucker spuckt Flamme						

1-7 A / 8-14 B / 15-28 Distraktoritens

Welche Vorleserin fandest Du netter? _____

Forcierung+/Strafe- Version 4

Datum: _____ Name: _____

Klasse: _____ Alter: _____

Interviewer: _____

Schule: _____

Namen Vorleser: _____

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Geraten“	Vorleser 1	Vorleser 2	„Geraten“
1= Bodenturner bauten Menschenturm						
9= Bodenturner machte Handstand						
16= Tauben auf Treppe						
28= Löwe brüllte vor Sprung auf Hocker						
20=Entfesselungskünstler befreite sich						
27= Mann auf Seil jonglierte mit Bällen						
18=Schlangenbeschwörer mit Flöte						
13= Elefanten auf Podesten						
25= Clown machte Seifenblasen						
15= Seehund mit Ball						
8= Mädchen stand auf Pferd						
21= Schwertschlucker schluckt Schwert						
24= Ponies über Wassergraben						
3= Zauberer holte Kaninchen						
14= Zauberer mit Kartentricks						
7= Mann schnappte Frau aus Luft						
4= Elefanten machten Handstand						
6=Tiger balancierte auf Laufkugel						
26= Mann auf Wippe, Frau in der Luft						
2= Ponyballett mit Musik						
10=Tiger sprang durch Reifen						
12= Mann auf Seil, Frau Handstand						
19= Affen spielten mit Ball						
23= Zauberer und Blumenstrauß						
17=Feuerspucker spuckt Flamme						
11= Clown mit Stelzen rief Leute						
5= Clown spielte Tuba						
22= Elefanten liefen in langen Reihen						

1-7 A / 8-14 B / 15-28 Distraktoritens

Welche Vorleserin fandest Du netter? _____

Forcierung+/Strafe+ Version 1

Datum: _____ Name: _____

Klasse: _____ Alter: _____

Interviewer: _____

Schule: _____

Namen Vorleser: _____

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Geraten“	Vorleser 1	Vorleser 2	„Geraten“
6=Tiger balancierte auf Laufkugel						
18=Schlangen- beschwörerin mit Flöte						
10=Tiger sprang durch Reifen						
21= Schwertschlucker schluckt Schwert						
14= Zauberer mit Kartentricks						
17=Feuerspucker spuckt Flamme						
9= Bodenturner machte Handstand						
3= Zauberer holte Kaninchen						
28= Löwe brüllte vor Sprung auf Hocker						
5= Clown spielte Tuba						
8= Mädchen stand auf Pferd						
27= Mann auf Seil jonglierte mit Bällen						
15= Seehund mit Ball						
1= Bodenturner bauten Menschenturm						
4= Elefanten machten Handstand						
26= Mann auf Wippe, Frau in der Luft						
12= Mann auf Seil, Frau Handstand						
19= Affen spielten mit Ball						
11= Clown mit Stelzen rief Leute						
2= Ponyballett mit Musik						
25= Clown machte Seifenblasen						
7= Mann schnappte Frau aus Luft						
22= Elefanten liefen in langen Reihen						
20=Entfesselungs-künstler befreite sich						
16= Tauben auf Treppe						
13= Elefanten auf Podesten						
24= Ponies über Wassergraben						
23= Zauberer und Blumenstrauß						

1-7 A / 8-14 B / 15-28 Distraktoritens

Welche Vorleserin fandest Du netter? _____

Forcierung+/Strafe+ Version 2

Datum: _____ Name: _____
 Klasse: _____ Alter: _____
 Interviewer: _____
 Schon mal im Zirkus gewesen: ja / nein

Schule: _____
 Namen Vorleser: _____

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Geraten“	Vorleser 1	Vorleser 2	„Geraten“
23= Zauberer und Blumenstrauß						
18=Schlangen- beschwörerin mit Flöte						
9= Bodenturner machte Handstand						
13= Elefanten auf Podesten						
19= Affen spielten mit Ball						
17=Feuerspucker spuckt Flamme						
28= Löwe brüllte vor Sprung auf Hocker						
12= Mann auf Seil, Frau Handstand						
25= Clown machte Seifenblasen						
3= Zauberer holte Kaninchen						
22= Elefanten liefen in langen Reihen						
1= Bodenturner bauten Menschenturm						
16= Tauben auf Treppe						
26= Mann auf Wippe, Frau in der Luft						
24= Ponies über Wassergraben						
8= Mädchen stand auf Pferd						
20=Entfesselungs- künstler befreite sich						
14= Zauberer mit Kartentricks						
27= Mann auf Seil jonglierte mit Bällen						
6=Tiger balancierte auf Laufkugel						
7= Mann schnappte Frau aus Luft						
2= Ponyballett mit Musik						
4= Elefanten machten Handstand						
11= Clown mit Stelzen rief Leute						
15= Seehund mit Ball						
5= Clown spielte Tuba						
10=Tiger sprang durch Reifen						
21= Schwertschlucker schluckt Schwert						

1-7 A / 8-14 B / 15-28 Distraktoritens

Welche Vorleserin fandest Du netter? _____

Forcierung+/Strafe+ Version 3

Datum: _____ Name: _____

Klasse: _____ Alter: _____

Interviewer: _____

Schule: _____

Namen Vorleser: _____

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Geraten“	Vorleser 1	Vorleser 2	„Geraten“
19= Affen spielten mit Ball						
16= Tauben auf Treppe						
14= Zauberer mit Kartentricks						
9= Bodenturner machte Handstand						
23= Zauberer und Blumenstrauß						
20=Entfesselungskünstler befreite sich						
27= Mann auf Seil jonglierte mit Bällen						
7= Mann schnappte Frau aus Luft						
2= Ponyballett mit Musik						
8= Mädchen stand auf Pferd						
4= Elefanten machten Handstand						
18=Schlangenbeschwörerin mit Flöte						
21= Schwertschlucker schluckt Schwert						
11= Clown mit Stelzen rief Leute						
25= Clown machte Seifenblasen						
12= Mann auf Seil, Frau Handstand						
10=Tiger sprang durch Reifen						
26= Mann auf Wippe, Frau in der Luft						
22= Elefanten liefen in langen Reihen						
13= Elefanten auf Podesten						
15= Seehund mit Ball						
28= Löwe brüllte vor Sprung auf Hocker						
6=Tiger balancierte auf Laufkugel						
5= Clown spielte Tuba						
1= Bodenturner bauten Menschenturm						
3= Zauberer holte Kaninchen						
24= Ponies über Wassergraben						
17=Feuerspucker spuckt Flamme						

1-7 A / 8-14 B / 15-28 Distraktoritens

Welche Vorleserin fandest Du netter? _____

Forcierung+/Strafe+ Version 4

Datum: _____ Name: _____

Klasse: _____ Alter: _____

Interviewer: _____

Schule: _____

Namen Vorleser: _____

Schon mal im Zirkus gewesen: ja / nein

	Alt/Neu= Kam das in einer der Geschichten vor?			QI= Wer hat das vorgelesen?		
	Ja	Nein	„Geraten“	Vorleser 1	Vorleser 2	„Geraten“
1= Bodenturner bauten Menschenturm						
9= Bodenturner machte Handstand						
16= Tauben auf Treppe						
28= Löwe brüllte vor Sprung auf Hocker						
20=Entfesselungskünstler befreite sich						
27= Mann auf Seil jonglierte mit Bällen						
18=Schlangenbeschwörerin mit Flöte						
13= Elefanten auf Podesten						
25= Clown machte Seifenblasen						
15= Seehund mit Ball						
8= Mädchen stand auf Pferd						
21= Schwertschlucker schluckt Schwert						
24= Ponies über Wassergraben						
3= Zauberer holte Kaninchen						
14= Zauberer mit Kartentricks						
7= Mann schnappte Frau aus Luft						
4= Elefanten machten Handstand						
6=Tiger balancierte auf Laufkugel						
26= Mann auf Wippe, Frau in der Luft						
2= Ponyballett mit Musik						
10=Tiger sprang durch Reifen						
12= Mann auf Seil, Frau Handstand						
19= Affen spielten mit Ball						
23= Zauberer und Blumenstrauß						
17=Feuerspucker spuckt Flamme						
11= Clown mit Stelzen rief Leute						
5= Clown spielte Tuba						
22= Elefanten liefen in langen Reihen						

1-7 A / 8-14 B / 15-28 Distraktoritens

Welche Vorleserin fandest Du netter? _____

Anhang G QI Auswertungsbögen für die vier Bedingungen

1. Forcierung-/Strafe-

		Quelle (tatsächlich)			
		Quelle 1 (Maren) A	Quelle 2 (Anna) B	Neu N	Ich weiß nicht IW
Antwort	Quelle 1 (Maren) A	Treffer X_{AA1}	Fehler X_{BA1}	False Alarm X_{NA1}	Treffer X_{IWA1}
	Quelle 2 (Anna) B	Fehler X_{AB1}	Treffer X_{BB1}	False Alarm X_{NB1}	Treffer X_{IWB1}
	Neu N	Miss X_{AN1}	Miss X_{BN1}	Treffer X_{NN1}	IRfalsch; QI? X_{IWN1}
	Ich weiß nicht IW	Miss X_{AIW1}	Miss X_{BIW1}	Miss X_{NIW1}	
SUMME					

2. Forcierung-/Strafe+

		Quelle (tatsächlich)			
		Quelle 1 (Maren) A	Quelle 2 (Anna) B	Neu N	Ich weiß nicht IW
Antwort	Quelle 1 (Maren) A	Treffer X_{AA2}	Fehler X_{BA2}	False Alarm X_{NA2}	Treffer X_{IWA2}
	Quelle 2 (Anna) B	Fehler X_{AB2}	Treffer X_{BB2}	False Alarm X_{NB2}	Treffer X_{IWB2}
	Neu N	Miss X_{AN2}	Miss X_{BN2}	Treffer X_{NN2}	X_{IWIW2}
	Ich weiß nicht IW	Miss X_{AIW2}	Miss X_{BIW2}	Miss X_{NIW2}	
SUMME					


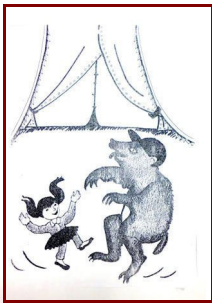
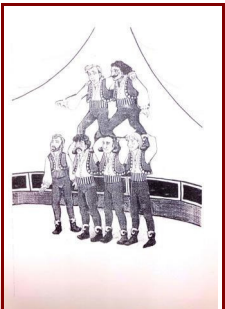
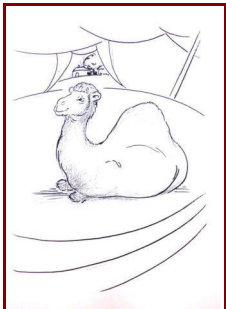
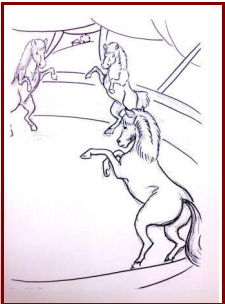
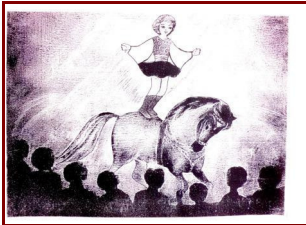
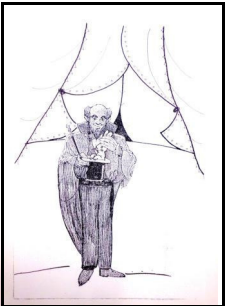

3. Forcierung+/Strafe-

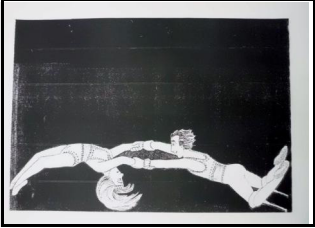


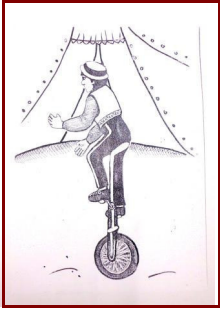
		Quelle (tatsächlich)					
		Quelle 1 (Maren) A	Quelle 2 (Anna) B	Neu N	Geraten G_IR	Geraten G_QI	
Antwort	Quelle 1 (Maren) A	Treffer Y _{AA3}	Fehler Y _{BA3}	False Alarm Y _{NA3}			
	Quelle 2 (Anna) B	Fehler Y _{AB3}	Treffer Y _{BB3}	False Alarm Y _{NB3}	G_IRa3	G_QIa3	
	Neu N	Miss Y _{AN3}	Miss Y _{BN3}	Treffer Y _{NN3}	G_IRb3	G_QIb3	
	SUMME						

4. Forcierung+/Strafe+

		Quelle (tatsächlich)					
		Quelle 1 (Maren) A	Quelle 2 (Anna) B	Neu N	Geraten G_IR	Geraten G_QI	
Antwort	Quelle 1 (Maren) A	Treffer Y _{AA4}	Fehler Y _{BA4}	False Alarm Y _{NA4}		G _{QIa4}	
	Quelle 2 (Anna) B	Fehler Y _{AB4}	Treffer Y _{BB4}	False Alarm Y _{NB4}	G _{IRa4}	G _{QIb4}	
	Neu N	Miss Y _{AN4}	Miss Y _{BN4}	Treffer Y _{NN4}	G _{IRn4}		
	SUMME						

Anhang H Illustrationen der 10 Inhaltssequenzen der Geschichten sowie die dazugehörigen Items zur Bestimmung der Erinnerungsleistung und ihrer Detailliertheit

Zirkusgeschichte A		Zirkusgeschichte B	
<i>Illustration</i>	<i>zugeordnete Items</i>	<i>Illustration</i>	<i>zugeordnete Items</i>
<p>Attraktion 1</p> 	<p><i>Zusatzitems:</i> Z1_A_DirekGrüUnif Z2_A_DirekKamel Z3_A_DirekHut</p>	<p>Attraktion 1</p> 	<p><i>Zusatzitems:</i> Z19_B_Musik Z20_B_BärMäd</p>
<p>Attraktion 2</p> 	<p><i>Item-QI-Bogen:</i> I1_A_BoTurnTurm</p> <p><i>Zusatzitems:</i> Z4_A_ApplBoTurn</p>	<p>Attraktion 2</p> 	<p><i>Zusatzitems:</i> Z21_B_KamHöck Z22_B_KamHinleg</p>
<p>Attraktion 3</p> 	<p><i>Item-QI-Bogen:</i> I2_A_PoBaMusi</p> <p><i>Zusatzitems:</i> Z5_A_PoBaHinBei Z6_A_StillNachPoBa</p>	<p>Attraktion 3</p> 	<p><i>Item-QI-Bogen:</i> I8_B_MädPferd</p> <p><i>Zusatzitems:</i> Z23_B_ApplReit</p>
<p>Attraktion 4</p> 	<p><i>Item-QI-Bogen:</i> I3_A_ZaubKanin</p> <p><i>Zusatzitems:</i> Z7_A_ZaubHutFarb Z8_A_ZaubHutLee</p>	<p>Attraktion 4</p> 	<p><i>Item-QI-Bogen:</i> I9_B_BoTurnHand</p> <p><i>Zusatzitems:</i> Z24_B_TurnHand Z25_B_BodTurnÜb</p>

Zirkusgeschichte A		Zirkusgeschichte B	
<i>Illustration</i>	<i>zugeordnete Items</i>	<i>Illustration</i>	<i>zugeordnete Items</i>
Attraktion 9		Attraktion 9	
	<i>Item-QI-Bogen:</i> I7_A_MannFrauLuft		<i>Item-QI-Bogen:</i> I14_B_ZaubKarten
	<i>Zusatzitems:</i> Z15_A_TromWir Z16_A_KlettDach		<i>Zusatzitems:</i> Z33_B_ZeltDu Z34_B_ZaubUm
	<i>Emotions-Item:</i> E2_A_Trapez		
Attraktion 10		Attraktion 10	
	<i>Zusatzitems:</i> Z17_A_BärRad Z18_A_ApplBär		<i>Zusatzitems:</i> Z35_B_EinRad Z36_B_HutRad

Anhang I Kodierleitfaden QI-Bogen Items

Item QI-Bogen	Geschichte - Attraktion	Textteil	Definition	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
11_A_BoTurmTurm	A-2	[...] Sie bauten alle zusammen einen großen Menschenturm [...]	Es muss hervorgehen, dass die Bodenturner gemeinsam einen menschlichen Turm zeigen.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Bodenturner, Artisten, Akrobaten, Männer - Turm, Menschenturm, Pyramide, Menschenpyramide - übereinandergestellt, aufeinander	„Also, da waren noch Artisten, die haben eine Pyramide gemacht.“
12_A_PoBaMusi	A-3	[...] Beim Ponyballett gingen die Ponys genau nach der Musik [...]	Es muss hervorgehen, dass Ponys oder Pferde tanzten.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Ponys, Pferde, Ponyballett, Pferdeballett - tanzen, rumtanzen, Ponytanz - Musik	„Dann kamen Pferde, die haben da rumgetanz.“
13_A_ZaubKanin	A-4	[...] Zauberer [...] holte er auf einmal ein Kaninchen hervor [...]	Es muss hervorgehen, dass ein Zauberer plötzlich ein Kaninchen vorzeigte.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Zauberer - Kaninchen, Hase - hervorholen, hervorzaubern, holen	„Ein Zauberer, der hat aus seinem Hut dann ein Kaninchen rausgeholt.“
14_A_ElefHand	A-5	[...] Elefanten machten einen Handstand, so dass ihre Hinterbeine in der Luft waren [...]	Es muss hervorgehen, dass Elefanten einen Handstand zeigten.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Elefanten, Elefant - Handstand, Kopfstand - Beine, Hinterbeine, Vorderbeine - oben, in Luft - Kunstturnen machen	„Die beiden Elefanten haben dann auf ihren Vorderbeinen gestanden und einen Handstand gemacht.“
15_A_ClownTuba	A-7	[...] Clown spielte dabei auf einer sehr großen Tuba [...]	Es muss hervorgehen, dass ein Clown auf einer Tuba oder einem ähnlichen Blasinstrument spielte.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Clown - Tuba, Trompete, Posaune	„Der eine Clown, der hat auf einer Trompete geblasen.“
16_A_TigLaufkug	A-8	[...] Tiger stürmten in die Manege und einer von ihnen balancierte auf einer Laufkugel [...]	Es muss hervorgehen, dass Tiger auftraten und einer der Tiger auf einem runden Gegenstand balancierte.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Tiger, Tigernummer, Tigerdressur - Kugel, Ball - gelaufen, balanciert	„Dann waren da auch Tiger und der eine von denen, der war auf einem Ball.“
17_A_MannFrauLuft	A-9	[...] Frau flog hoch oben durch die Luft und der Mann an der Schaukel schnappte sie aus der Luft [...]	Es muss hervorgehen, dass eine weibliche und eine männliche Person an einem Trapez hingen und ein Kunststück in der Luft zeigten.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Mann/Frau; Junge/Mädchen - Trapez; Schaukel - gehangen, (fest)gehalten, schwingen, fallen	„Da waren ein Mann und eine Frau an einem Trapez und die Frau ließ sich zu dem Mann rüberschwingen.“

Item QI-Bogen	Geschichte - Attraktion	Textteil	Definition	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
I8_B_MädPferd	B-3	[...] stand ein Mädchen mit ausgestreckten Armen auf dem Rücken eines geschmückten Pferdes [...]	Es muss hervorgehen, dass eine weibliche Person stehend auf einem Pferd auftrat.	bei Nennung von mind. zwei Begriffseinheiten: - Frau, Mädchen, Tänzerin, Ballerina, Artist - Pferd, Pferde - Kunststücke, Kunstturnen, stehend reiten	„Eine Frau auf einem Pferd machte Kunststücke.“
I9_Bo_TurnHand	B-4	[...] Ein Turner machte zuerst einen Handstand auf den Händen des anderen [...]	Es muss hervorgehen, dass zwei Personen auftraten und eine einen Handstand auf den Händen der anderen zeigte.	bei Nennung von mind. zwei Begriffseinheiten: - Bodenturner, Turner - Handstand, auf den Händen - auf dem anderen, übereinander	„Ein Turner machte so einen Handstand auf dem anderen.“
I10_B_TigReif	B-5	[...] Tiger stürmten herein und einer von ihnen sprang von einem Hocker mitten durch einen Reifen [...]	Es muss hervorgehen, dass Tiger auftraten und einer der Tiger durch einen Reifen sprang.	bei Nennung von mind. zwei Begriffseinheiten: - Tiger, Tigernummer, Tigerdressur - Reifen, Ring, Kreis - springen, hüpfen, durch gehen	„Und bei der anderen Tigernummer, da ist einer durch so einen Ring gesprungen.“
I11_B_ClownStelzRuf	B-6	[...] rief ein Clown auf ganz hohen Stelzen die Leute zurück in das Zirkuszeit [...]	Es muss hervorgehen, dass ein Clown auf Stelzen vorkam, der die Leute rief.	bei Nennung von mind. zwei Begriffseinheiten: - Clown, Mann - Stelzen, Holzdingler, Latten - rufen, zurückrufen, zurückholen	„In der Pause kam ein Clown mit hohen Stelzen. Der hat die Leute zurückgeholt.“
I12_B_MaSeilFHa	B-7	[...] Auf dem Seil balancierte ein Mann [...] [...] Eine Frau machte auf seinem Kopf einen Handstand auf einer Hand [...]	Es muss hervorgehen, dass eine männliche und eine weibliche Person auf einem Seil balancierten und die weibliche Person dabei einen Handstand auf dem Kopf der männlichen Person zeigt.	bei Nennung von mind. zwei Begriffseinheiten: - Frau, Mann, Junge, Mädchen, Seiltänzerin, Seiltänzer - Seil - balancieren - Handstand, Kopf	„Da war eine Balletttänzerin auf'm Hochseil. Und da war dann noch so ein Mann da auf'm Hochseil und die Frau stand auf dem in einem Spagat, im Handstand.“
I13_B_ElefPod	B-8	[...] Die Elefanten liefen über ganz kleine Podeste [...]	Es muss hervorgehen, dass Elefanten über Podeste liefen.	bei Nennung von mind. zwei Begriffseinheiten: - Elefanten, zwei Elefanten - Podeste, Hocker, „Hügel“, „Dinger“, „runde Teile“, „Bühnen“ - gelaufen, balanciert, drauf gehen	„Dann kamen Elefanten. Die sind auf so kleinen Bühnen gegangen.“
I14_B_ZaubKarten	B-9	[...] ein Zauberer [...] konnte uns tolle Tricks mit Spielkarten zeigen [...]	Es muss hervorgehen, dass ein Zauberer Kartentricks zeigte.	bei Nennung von mind. zwei Begriffseinheiten: - Zauberer, Mann - Zaubertricks, Tricks - Karten, Spielkarten, Kartenspiele	„Ein Mann hat mit Karten Zaubertricks gemacht.“

Anhang J Kodierleitfaden Zusatz-Items

Zusatz - Item	Geschichte - Attraktion	Textteil Zirkusgeschichte	Definition	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
Z1_A_DirekGrüUnif	A-1	[...] Der Zirkusdirektor in seiner grünen Uniform [...]	Es muss hervorgehen, dass eine männliche Person grüne Kleidung trug.	bei Nennung von mindestens zwei Begriffen und/oder Begriffseinheiten: - Zirkusdirektor, Zirkusbesitzer, Dompteur, Regisseur, Mann - grüne Uniform, grünes Jacket, grüner Anzug	„Der Zirkusdirektor hatte einen grünen Anzug an.“
Z2_A_DirekKamel	A-1	[...] Zirkusdirektor [...] ritt auf einem Kamel in der Manege herum [...]	Es muss hervorgehen, dass eine männliche Person auf einem Kamel ritt.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Zirkusdirektor, Zirkusbesitzer, Dompteur, Regisseur, Mann - Kamel, Lama - stand, ritt, saß	„Ein Mann ist auf einem Kamel geritten.“
Z3_A_DirekHut	A-1	[...] Dabei schwenkte er seinen Hut und begrüßte alle kleinen und großen Leute [...]	Es muss hervorgehen, dass der Zirkusdirektor einen Hut trug und das Publikum begrüßt.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Zirkusdirektor, Zirkusbesitzer, Mann - Leute, Publikum, alle - begrüßen, grüßen - Hut	„Der Direktor begrüßte die Leute“
Z4_A_ApplBoTurn	A-2	[...] Bodenturner [...] zogen unter Beifall aus der Manege [...]	Es muss hervorgehen, dass die Bodenturner Applaus erhielten.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Applaus, Klatschen, Jubel, Beifall - applaudieren, klatschen, jubeln	keine Nennung erfolgt
Z5_A_PoBaHinBei	A-3	[...] Ponys [...] und liefen alle gleichzeitig auf den Hinterbeinen [...]	Es muss hervorgehen, dass Ponys oder Pferde sich auf ihre Hinterbeine stellen.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Ponys, Pferde - Hinterbeine, zwei Beine, zwei Hufe	„Da kamen dann Pferde, die konnten auf den Hinterbeinen stehen.“
Z6_A_StilNachPoBa	A-3	[...] Darauf wurde es wieder ganz still in dem großen Zirkuszeit [...]	Es muss hervorgehen, dass Stille im Zirkus eintrat.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Stille - still, ruhig - Publikum, Leute	keine Nennung erfolgt

Zusatz - Item	Geschichte - Attraktion	Textteil Zirkusgeschichte	Definition	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
Z7_A_ZaubHutFarb	A-4	[...] Zauberer [...] aus seinem silbernen Zylinderhut [...]	Es muss hervorgehen, dass ein Zauberer einen silbernen Zylinderhut trug.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Zauberer, Mann - Zylinder, Hut, Zylinderhut - silber, silbern, grau	keine Nennung erfolgt
Z8_A_ZaubHutLee	A-4	[...] obwohl der Hut vorher ganz leer war [...]	Es muss hervorgehen, dass der Hut des Zauberers zunächst leer war.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Zauberer, Mann - Hut, Zylinder - leer, nichts drin - vorher, zuerst	„Der holte dann einen Hasen heraus. Dabei war der Hut zuerst ganz leer.“
Z9_A_ElefZweiKlein	A-5	[...] nun liefen zwei kleine Elefanten in die Manege [...]	Es muss hervorgehen, dass zwei kleine Elefanten auftraten.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - kleine Elefanten - zwei	keine Nennung erfolgt
Z10_A_FahrradMäd	A-6	[...] kam die Fahrradnummer. Ein Mädchen balancierte mit einem Fuß auf dem Sattel und mit dem anderen Fuß auf dem Lenker des Fahrrades und fuhr so immer im Kreis [...]	Es muss hervorgehen, dass ein Mädchen auf einem Fahrrad auftrat und darauf balancierte.	bei Nennung von mind. drei Begriffen und/oder Begriffseinheiten: - Mädchen, Frau, Artistin, Junge - eine, einer, weiche, jemand - Fahrrad, Fahrradnummer - Fuß, Füße - Lenker - Sattel - Kreis - fahren, gefahren, am stehen	„Jemand ist auf dem Sattel und dem Lenker Fahrrad gefahren.“
Z11_A_ApplFahr	A-6	[...] Da haben alle geklatscht [...]	Es muss hervorgehen, dass das Mädchen auf dem Fahrrad Applaus erhielt.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Mädchen, Frau - Fahrrad, Fahrradnummer - Applaus, Klatschen, Jubel, Beifall - applaudieren, klatschen, jubeln	keine Nennung erfolgt

Zusatz - Item	Geschichte - Attraktion	Textteil Zirkusgeschichte	Definition	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
Z12_A_Pause	A-7	[...] Nach der Pause [...]	Es muss hervorgehen, dass es eine Pause in der Vorstellung gab.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Vorstellung, Nummern - Pause	keine Nennung erfolgt
Z13_A_ClownKostü	A-7	[...] ein Clown in einem gestreiften Kostüm und mit ganz großen Schuhen [...]	Es muss hervorgehen, dass ein Clown ein gestreiftes Kostüm und große Schuhe trug.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Clown, Mann - Schuhe, Riesenschuhe - Füße, Riesenfüße - Anzug, Kostüm, Jacke - gestreift, Streifen - groß, riesig	„Da ist noch ein Clown mit riesengroßen Schuhen aufgetreten.“
Z14_A_TigKäf	A-8	[...] Käfig für die Tiger aufgebaut [...]	Es muss hervorgehen, dass ein Käfig für die Tiger aufgebaut war. Die Aussage muss in direkt vor oder nach einer Aussage mit „Tiger balanciert auf Laufkugel“ erfolgen.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Tiger, Tigernummer - Käfig, Gehege - aufgebaut, war da, stand, aufgestellt, umgewandelt	„Für die Tigernummer war dann da ein Käfig aufgestellt.“
Z15_A_TromWir	A-9	[...] Trommelwirbel kündigten den Auftritt [...] an [...]	Es muss hervorgehen, dass es einen Trommelwirbel gab.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Trommel, Trommeln, Schlagzeug - Trommelwirbel - aufspielen, trommeln	keine Nennung erfolgt
Z16_A_KlettDach	A-9	[...] Ein Mann und eine Frau kletterten bis hoch unter das Dach des Zirkuszeltes [...]	Es muss hervorgehen, dass zwei Personen unter das Zirkuszeltdach kletterten.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Frau, Mädchen, Trapezkünstlerin - Mann, Junge, Trapezkünstler - Zirkuszelt, Zirkusdach, Zeltdach - klettern, hochgehen	keine Nennung erfolgt

Zusatz - Item	Geschichte - Attraktion	Textteil Zirkusgeschichte	Definition	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
Z17_A_BärRad	A-10	[...] die Bärennummer. Das Tollste war, als ein Braunbär auf einem Dreirad in der Manege herumfuhr [...]	Es muss hervorgehen, dass ein Bär auf einem Dreirad oder ähnlichem Gefährt fuhr.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Bär, Braunbär - braun - Dreirad, Fahrrad, Rad, Einrad - fahren, herumfahren	„Ein Bär ist auf einem Dreirad gefahren.“
Z18_A_AppiBär	A-10	[...] Da haben alle geklatscht [...]	Es muss hervorgehen, dass der Bär auf dem Dreirad Applaus erhielt.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Bär, Braunbär, braun - Applaus, Klatschen, Jubel, Beifall - applaudieren, klatschen, jubeln	„Da haben alle für den Bär geklatscht.“
Z19_B_Musik	B-1	[...] fing die Musik an zu spielen [...]	Es muss hervorgehen, dass Musik einsetzte und anschließend Bär und Mädchen tanzten.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Musik, Lied - anfangen, beginnen	keine Nennung erfolgt
Z20_B_BärMäd	B-1	[...] Ein großer brauner Bär und ein Mädchen kamen in die Manege [...] und tanzten zur Musik [...]	Es muss hervorgehen, dass ein Bär und ein Mädchen auftraten und tanzten.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Bär, Braunbär - groß, braun - Mädchen, Frau, Tänzerin - tanzen, spielen	„Da waren ein Bär und ein Mädchen. Und dann haben die miteinander getanzt.“
Z21_B_KamHöck	B-2	[...] großes braunes Kamel in die Manege gefrabt. Dabei wackelte es mit seinen Höckern [...]	Es muss hervorgehen, dass ein Kamel auftrat.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Kamel, Dromedar, Lama	„Ein Kamel ist im Kreis gegangen.“

Zusatz - Item	Geschichte - Attraktion	Textteil Zirkusgeschichte	Definition	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
Z22_B_KamHinleg	B-2	[...] legte es sich hin und schaute wie ich beim Kunstreiten zu [...]	Es muss hervorgehen, dass das Kamel sich hinlegte und der Vorstellung zusah.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Kamel, Dromedar, Lama - hinlegen, sitzen, liegen - zusehen, zuschauen, gucken	„Das Kamel hat sich dann hingelegt und zugeschaut.“
Z23_B_ApplReit	B-3	[...] Da Haben die Zuschauer alle geklatscht [...]	Es muss hervorgehen, dass das Mädchen auf dem Pferd Applaus erhielt.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Frau, Mädchen, Tänzerin, Ballerina, Artist - Pferd, Pferde - Applaus, Klatschen, Jubel, Beifall - applaudieren, klatschen, jubeln	keine Nennung erfolgt
Z24_B_TurnHand	B-4	[...] Dann beugte er seinen Körper soweit, dass es schien, als ob er über dem anderen Turner schweben würde [...]	Es muss hervorgehen, dass eine Person im Handstand über der anderen schwebte.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Bodenturner, Turner - Handstand, auf den Händen - auf dem anderen, übereinander - schweben - gebogen	„Zwei Menschen waren da, der eine hat einen Handstand auf dem anderen gemacht. Und der war so über dem gebogen.“
Z25_B_BodTurnÜb	B-4	[...] Es sah so einfach aus, aber die beiden mussten bestimmt lange dafür üben [...]	Es muss hervorgehen, dass die Handstandattraktion der Bodenturner viel Übung erfordert.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Bodenturner, Turner - Handstand, auf den Händen - Übung, Training - trainieren, üben, oft machen	keine Nennung erfolgt
Z26_B_TigKäF	B-5	[...] wurde ein Käfig für die Tiger aufgebaut [...]	Es muss hervorgehen, dass ein Käfig für die Tigerattraktion aufgebaut wurde. Die Aussage muss in direkt vor oder nach einer Aussage mit „Tiger sprang durch Reifen“ erfolgen.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Tiger, Tigernummer - Käfig, Gehege - aufgebaut, war da, stand, umgewandelt	„Da wurde dann das Gehege aufgebaut.“

Zusatz - Item	Geschichte - Attraktion	Textteil Zirkusgeschichte	Definition	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
Z27_B_DomReif	B-5	[...] durch einen Reifen, den der Dompteur hochhielt [...]	Es muss hervorgehen, dass der Dompteur einen Reifen hochhielt.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Dompteur, Direktor, Mann - Reifen, Ring, Kreis - hochhalten, festhalten	„Der Dompteur hat einen Ring für die Tiger hoch gehalten.“
Z28_B_ClowBei	B-6	[...] sah so aus, als ob er ganz lange Beine hätte [...]	Es muss hervorgehen, dass ein Clown aussah, als ob er sehr lange Beine habe.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Clown, Mann, Typ - Stelzen, Stäbe - Beine - lang, groß, hoch	„Es gab einen ganz langen Mann, also der Clown, der mit den Stelzen.“
Z29_B_Seil	B-7	[...] war hoch über uns ein Seil gespannt [...]	Es muss hervorgehen, dass hoch über dem Publikum ein Seil gespannt war.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Seil, Band - Zirkusdach, Dach, unterm Zelt - gespannt, festgemacht, angebracht	„Und dann war da ein Seil oben unterm Zirkuszelt.“
Z30_B_SeilBalan	B-7	[...] balancierte ein Mann [...] mit einer schwarzen Stange [...]	Es muss hervorgehen, dass eine Person mit einer schwarzen Stange balancierte.	bei Nennung von mind. drei Begriffen und/oder Begriffseinheiten: - Mann, Mensch, Person - balancieren, gehen, laufen - Stange, Stock - schwarz	„Der ist mit einer Stange auf dem Seil gelaufen.“
Z31_B_GroElef	B-8	[...] trotteten zwei große Elefanten [...]	Es muss hervorgehen, dass zwei große Elefanten auftraten.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - große Elefanten - zwei	keine Nennung erfolgt
Z32_B_ElefFehl	B-8	[...] ohne auch nur einmal daneben zu treten [...]	Es muss hervorgehen, dass die Elefanten keine Fehler machten.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Elefanten - keine(n) Fehler - ohne zu stolpern - ohne daneben zu treten / gehen	„Und die Elefanten sind dabei nicht gestolpert.“

Zusatz - Item	Geschichte - Attraktion	Textteil Zirkusgeschichte	Definition	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
Z33_B_ZeitDu	B-9	[...] einen Moment dunkel im Zelt [...]	Es muss hervorgehen, dass es kurz dunkel im Zirkuszelt wurde.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Zirkus, Zirkuszelt, Zelt, Vorstellung - dunkel - Licht aus - kurz, Moment	keine Nennung erfolgt
Z34_B_ZaubUm	B-9	[...] mit einem großen blauen Umhang [...]	Es muss hervorgehen, dass der Zauberer einen blauen Umhang trug.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Zauberer - Umhang - blau	keine Nennung erfolgt
Z35_B_EinRad	B-10	[...] Einradnummer, die zum Schluss kam [...] Ein Junge fuhr auf einem hohen Einrad in großen Kreisen durch die Manege.[...]	Es muss hervorgehen, dass eine Person auf einem Einrad fuhr.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Junge, Mädchen, Mann, Frau, Person, Mensch - Einrad, Fahrrad - hoch, groß - fahren, durch Manege	„Ein Mann ist auf einem ganz großen Einrad gefahren.“
Z36_B_HutRad	B-10	[...] sein Hut [...]	Es muss hervorgehen, dass die Person auf dem Einrad einen Hut trug.	bei Nennung von mind. drei Begriffen und/oder Begriffseinheiten: - Junge, Mann, Person - Einrad, Dreirad, Fahrrad - Hut, Mütze	keine Nennung erfolgt

Anhang K Kodierleitfaden Rahmentext-Items

Item	Geschichte / Attraktion	Textteil	Kodierregel	Ankerbeispiele
Rahm1_AB_Ein	A + B / Einleitungstext	[...] Ich war am letzten Sonntag (Ich war letzte Woche auch) im Zirkus. (Aber ich war in einer ganz anderen Vorstellung.) Da gab es ganz viel zu sehen. Als alle Leute auf ihren Plätzen saßen, wurde es dunkel. Nur in der Manege schien noch Licht. [...]	Es muss hervorgehen, dass damit die Zirkusvorstellung begonnen hat; Nennung von mindestens einem Begriffen und /oder Begriffseinheiten: - zwei verschiedene Vorstellungen - viel zusehen, viele Sachen - Beginn, losgehen, Anfang - es wurde dunkel, Licht ging aus - in Manege noch Licht, hell, Lampen	„Die waren in zwei verschiedenen Zirkussen.“ „Am Anfang, da wurde es dunkel. Nur in der Manege war noch was Licht.“
Rahm2_AB_End	A + B / Schlusstext	[...] Wie schade, dass die Vorstellung so schnell zu Ende war. Ich hätte gerne noch viel länger zusehen. Nach einem langen Applaus gingen schließlich alle nach Hause. [...]	Es muss hervorgehen, dass damit die Zirkusvorstellung geendet hat; Nennung von mindestens einem Begriffen und /oder Begriffseinheiten: - Ende der Vorstellung, fertig, Schluss - gerne länger zusehen, länger bleiben - Beifall, Applaus, geklatscht - alle gingen nach Hause, heim gehen	„Zum Schluss war Applaus.“ „Zum Schluss gingen alle nach Hause.“ „Die waren auch traurig, wie es zu Ende war.“

Anhang L Kodierleitfaden Schema-Items

Schema - Item	Geschichte - Attraktion	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
S1_Clown	A-7, B-6	bei reiner Singular- und/oder Pluralnennung des Begriffes „Clown“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Da waren so Clowns.“
S2_Bär	A-10, B-1	bei reiner Singular- und/oder Pluralnennung des Begriffes „Bär“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Ein Bär.“
S3_Elef	A-5, B-8	bei reiner Singular- und/oder Pluralnennung des Begriffes „Elefant“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Es gab Elefanten.“
S4_Pfe	A-3, B-3	bei reiner Singular- und/oder Pluralnennung des Begriffes „Pferd“ oder „Pony“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Pferde waren da.“ „Es gab Ponys.“
S5_Tig	A-8, B-5	bei reiner Singular- und/oder Pluralnennung des Begriffes „Tiger“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Da kamen Tiger.“
S6_Löw	nicht vorhanden	bei reiner Singular- und/oder Pluralnennung des Begriffes „Löwe“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Und da waren noch Löwen.“
S7-Kam	A-1, B-2	bei reiner Singular- und/oder Pluralnennung des Begriffes „Kamel“ oder „Lama“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Kamele.“
S8_Seiltanz	B-7	bei reiner Singular- und/oder Pluralnennung des Begriffes „Seiltanz“, „Seiltänzer“ oder „Seiltanznummer“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Und bei der Seiltanznummer, da ging ein Mann über das Seil.“
S9_Trapez	A-9	bei reiner Singular- und/oder Pluralnennung des Begriffes „Trapez“, „Trapezkünstler“ oder „Trapeznummer“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Es gab eine Nummer am Trapez.“
S10_Zaub	A-4, B-9	bei reiner Singular- und/oder Pluralnennung des Begriffes „Zauberer“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Da waren Zauberer.“
S11_Turn	A-2, B-4	bei reiner Singular- und/oder Pluralnennung des Begriffes „Turner“ oder „Bodenturner“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Da haben welche geturnt.“
S12_Akrob	A-2, A-9, B-4, B-7	bei reiner Singular- und/oder Pluralnennung des Begriffes „Akrobat“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Und Akrobaten.“
S13_FahrRad	A-6, A-10, B-10	bei reiner Singular- und/oder Pluralnennung des Begriffes „Fahrrad“ oder „Fahrradnummer“ oder „Dreirad“ oder „Einrad“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Da war was mit einem Fahrrad.“
E6_Schema_Aufreg	A-8, B-7	bei reiner Singular- und/oder Pluralnennung des Begriffes „Vorstellung“ und „aufregend“ oder „spannend“ ohne weitere Beschreibung eines inhaltlichen Geschehens	„Die Frau fand die Vorstellung war spannend.“

Anhang M Kodierleitfaden Vorleser-Items

Vorleser – Item (Quellenidentifikation)	Definition	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
VL1_Glatte Haare	Haarstruktur Anna: glatt	bei Nennung: - glatt, gerade - Haare, Frisur (Nennung des Namens ist keine Bedingung)	keine Nennung erfolgt
VL2_LockigeHaare	Haarstruktur Maren: lockig	bei Nennung: - Locken, lockig, Wellen, gewellt - Haare, Frisur (Nennung des Namens ist keine Bedingung)	„Also die erste, die hatte lockige Haare.“
VL3_DunkleHaare	Haarfarbe Anna: dunkelbraun	bei Nennung: - Haarfarbe, Haare, Farbe - dunkel, braun, schwarz, dunkelbraun, dunkler (Nennung des Namens ist keine Bedingung)	„Eine Frau, die war braun – schwarzhaarig.“
VL4_BlondeHaare	Haarfarbe Maren: dunkelblond	bei Nennung: - Haarfarbe, Haare, Farbe - hell, blond, dunkelblond, heller (Nennung des Namens ist keine Bedingung)	keine Nennung erfolgt
VL5_Zusamm	Frisur Anna: zusammengebundene, hochgesteckte Haare	bei Nennung: - zusammen, gebunden, Zopf, am Kopf zusammen gemacht (Nennung des Namens ist keine Bedingung)	keine Nennung erfolgt
VL6_Offen	Frisur Maren: offene Haare	bei Nennung: - offen, auf (Nennung des Namens ist keine Bedingung)	keine Nennung erfolgt
VL7_NameAn	spontane Nennung des Namens „Anna“ oder ähnlicher Varianten	bei spontaner Nennung des Namens „Anna“ oder „Anne“	„Als erstes war da die Anna.“
VL8_NameMar	spontane Nennung des Namens „Maren“ oder ähnlicher Varianten	bei spontaner Nennung des Namens „Maren“ oder „Mara“ oder „Mareen“	„Also, da war ein Mädchen, Maren hieß die, glaube ich.“

Anmerkung:

Die Items VL7_Name Anna und VL8_Name Maren wurden nicht weiter ausgewertet, da bei einigen Interviews entgegen der Instruktion die Namen der Vorleserinnen vom Interviewer genannt wurden und somit eine Beeinflussung erfolgte.

Anhang N Kodierleitfaden Emotions-Items

Emotions - Item	Geschichte - Attraktion	Textteil	Definition	Kodierregel Eine Aussage wird als Treffer kodiert:	Ankerbeispiel
E1_A_Tiger	A-8	[...] Das fand ich sehr aufregend [...]	Es muss hervorgehen, dass die Tigerattraktion auf der Laufkugel als aufregend empfunden wurde.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Tiger, Tigernummer - Laufkugel, Ball - aufregend, spannend	keine Nennung erfolgt
E2_A_Trapez	A-9	[...] Mir wurde schon vom Zusehen ganz schwindelig [...]	Es muss hervorgehen, dass bei Trapeznummer Schwindel empfunden wurde.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Trapez, Trapeznummer - schwindelig, schummrig	„Da konnte einem ganz schwindelig von werden.“
E3_B-Tiger	B-5	[...] Das war wirklich aufregend [...]	Es muss hervorgehen, dass die Tigerattraktion mit Sprung durch Reifen als aufregend empfunden wurde.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Tiger, Tigernummer - Reifen, Sprung - aufregend, spannend	keine Nennung erfolgt
E4_B_Seitanz	B-7	[...] Das sah ganz schön gefährlich aus [...]	Es muss hervorgehen, dass die Seiltanzattraktion mit Balanceakt als gefährlich empfunden wurde.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten:	keine Nennung erfolgt
E5_B_Einrad	B-10	[...] hat mir gut gefallen [...]	Es muss hervorgehen, dass die Einradnummer gut gefallen hat.	bei Nennung von mind. zwei Begriffen und/oder Begriffseinheiten: - Einrad, Einradnummer - gut gefallen, war toll	keine Nennung erfolgt

Anhang O Fehlerrate - Beispiele für Konfabulationen

Beispiel 1 (VP-Nr. 90):

„Da war ein Mann, der hat mit drei Bällen jongliert.“

Beispiel 2 (VP-Nr. 12):

„Ein Clown saß auf einem Stuhl, ist kurz aufgestanden um was zu holen und da hat der andere Clown ihm den Stuhl weggezogen.“

Beispiel 3 (VP-Nr. 12):

„Der Zirkusdirektor kam als Clown angezogen.“

Beispiel 4 (VP-Nr. 46):

„Ein Elefant war angebunden, saß in einem Wagen und wurde herumgeschoben.“

Beispiel 5 (VP-Nr. 49):

„Es gab einen Seehund, mmh, auf jeden Fall ein Tier das Kunststücke gemacht hat.“

Beispiel 6 (VP-Nr. 78):

„Ein Mensch ist auf einem Elefanten geritten.“

Beispiel 7 (VP-Nr. 86; VP-Nr. 46):

„Es gab Giraffen.“