

Wissenschaftliches Gutachten

Forensisch-aussagepsychologische Begutachtung (Glaubwürdigkeitsbegutachtung)

Max Steller und Renate Volbert

Entsprechend dem Auftrag des 1. Strafsenats des Bundesgerichtshofs vom 25.02.1999 wird ein wissenschaftliches Gutachten „zur Methodik und den Kriterien psychologischer Glaubwürdigkeitsgutachten“ erstattet. Gemäß Auftrag wird auch auf den Fragenkatalog des BGH eingegangen. Dieser enthält allgemeine Fragen zur Methodik und Praxis der forensisch-psychologischen Glaubwürdigkeitsbegutachtung sowie Fragen zur sachlichen Richtigkeit eines in der Revisionsache 1 StR 618/98 vorliegenden schriftlichen Gutachtens einer Diplompsychologin.

Als fallspezifische Unterlagen standen zur Verfügung:

1. schriftliches aussagepsychologisches Fachgutachten von Frau Diplompsychologin P. vom 10.04.1998,
2. schriftliche psychologische Stellungnahme zu diesem Gutachten durch Herrn Prof. Dr. Sch. vom 09.07.1998,
3. Urteil des Landgerichts Ansbach in der Sache KLS 10 Js 7266/97 jug.

Mit Schreiben vom 25.03.1999 übersandte der BGH entsprechende Unterlagen aus einem anderen Verfahren (1 StR 520/98) zur eventuellen Berücksichtigung bei der Gutachtenerstellung.

Auf die verwendete wissenschaftliche Literatur wird im Gutachten an den entsprechenden Stellen Bezug genommen.

Das vorliegende Gutachten gliedert sich wie folgt:

Seite

1. Methodik forensisch-aussagepsychologischer Begutachtungen

47

- | | |
|---|----|
| 1.1 Aussagepsychologische Fragestellungen | 47 |
| 1.2 Merkmalsorientierte Aussageanalyse | 50 |
| 1.2.1 Zugrundeliegende Überlegungen | 50 |
| 1.2.2 Systeme für merkmalsorientierte Aussageanalysen | 53 |
| 1.3 Integrierende Glaubhaftigkeitsbeurteilung | 57 |

2. Praxis forensisch-aussagepsychologischer Begutachtungen	61
2.1 Erhebungsbereiche und -methoden	61
2.1.1 Allgemeine Überlegungen	61
2.1.2 Exploration zur Sache	62
2.1.3 Leistungs- und Persönlichkeitsdiagnostik	63
2.2 Grenzen forensisch-aussagepsychologischer Befunderhebungen	65
2.3 Das schriftliche (vorbereitende) Gutachten	68
2.4 Standards forensisch-aussagepsychologischer Begutachtungen.....	75
3. Wissenschaftliche Grundlagen der Aussageanalyse	77
3.1 Feldstudien	78
3.2 Simulationsstudien	80
3.3 Bewertung der empirischen Studien.....	82
3.4 Neue Forschungsrichtungen	87
3.5 Inhaltliche Qualität in suggerierten Aussagen	90
4. Zur Kritik forensisch-aussagepsychologischer Begutachtungen in Strafverfahren	95
5. Literatur	99
Zum Glaubwürdigkeitsgutachten in vorliegender Revisionssache*	
6. Zusammenfassende Beantwortung des Fragenkatalogs des BGH	107

1 Methodik forensisch-aussagepsychologischer Begutachtungen

1.1 Aussagepsychologische Fragestellungen

Die im Titel dieser Ausarbeitung verwendete Formulierung „forensisch-aussagepsychologische Begutachtung“ verweist darauf, daß in der juristischen Praxis mit dem Begriff Glaubwürdigkeitsgutachten bzw. Glaubwürdigkeitsbegutachtung unterschiedliche Fragestellungen zusammengefaßt werden, deren gemeinsames Element in der Verwendung psychologischer, speziell aussagepsychologischer Erkenntnisse besteht. Begutachtung wurde als Begriff gewählt, da mit ihm auch die Durchführung der Befunderhebung (Datensammlung) durch Sachverständige und nicht nur das (abschließende) Gutachten erfaßt wird.

* Dieses Kapitel über das primäre Glaubwürdigkeitsgutachten und die darauf bezogene Methodenkritik durch Prof. Sch. wird hier nicht veröffentlicht, da diesbezügliche Vorbehalte der Erstgutachterin nicht ausgeräumt werden konnten.

In der forensischen Praxis hat sich in den vergangenen Jahrzehnten Konsens dahin gehend entwickelt, daß von Sachverständigen eine aussagebezogene Glaubhaftigkeitseinschätzung erwartet wird, auch wenn der gutachterliche Auftrag personenbezogene Begriffe enthält. Dieser Konsens erscheint schon deswegen angemessen, da eine hinreichende definitorische Übereinstimmung hinsichtlich des Konzepts einer personalen Glaubwürdigkeit wahrscheinlich nicht zu erreichen ist und Feststellungen von Sachverständigen über die (allgemeine) Glaubwürdigkeit von Zeugen im Vergleich zu Feststellungen über die Glaubhaftigkeit und Genauigkeit (d. h. über den Realitätsgehalt) ihrer Aussagen bei der Rekonstruktion von Sachverhalten durch Gerichte nur geringen indiziellen Wert haben.¹

Die Fragestellungen bei aussagepsychologischen Begutachtungen betreffen die Unterscheidung zwischen wahrheitsgemäßen Darstellungen und Lügen (Beurteilung der Glaubhaftigkeit einer Aussage), zuweilen geht es aber auch um die Unterscheidung von realitätsgerechten Darstellungen und Irrtümern (Beurteilung der Richtigkeit bzw. Genauigkeit einer Aussage). Es ist trivial, daß Irrtümer auch in intentional wahren (also glaubhaften) Aussagen vorkommen können. Weniger trivial ist allerdings, daß unter suggestiven Bedingungen Irrtümer bis zu einem Ausmaß entstehen können, daß (vermeintliche) „Erinnerungs“-Vorstellungen tatsächlich keinerlei Entsprechung in einer vorausgegangenen Realität haben. Die Bezeichnung solcher – subjektiv wahren, aber objektiv falschen – Aussagen als Lüge erscheint nicht sinnvoll.²

¹ Teilweise werden mit identischen Begriffen allerdings unterschiedliche Sachverhalte angesprochen. In einigen älteren Arbeiten wird allgemeine Glaubwürdigkeit als ein eigenschaftsorientiertes, situationsübergreifendes Konzept im Sinne von Aufrichtigkeit verstanden (Lefrenz, 1972). Selbst wenn man von definitorischen Schwierigkeiten bei einem solchen Konzept absieht, entspricht es bereits der Alltagserfahrung, daß Feststellungen über einen allgemein anerkannt positiven Leumund keine hinreichend eindeutigen Beziehungen zu der Verlässlichkeit von spezifischen Aussagen dieser Person zulassen, ebenso wie aus einem schlechten Leumund nicht zwingend auf die Unglaubhaftigkeit konkreter Aussagen zu schließen ist. Andere Autoren (z. B. Lempp, 1989, S. 615) verstehen unter allgemeiner Glaubwürdigkeit Aspekte der Zeugentüchtigkeit (Vorliegen von Voraussetzungen, um entsprechende Beobachtungen zu machen, sie zu behalten und wiederzugeben) bzw. spezifische Kompetenzen und Verhaltenstendenzen des Zeugen (Tendenz zur Konfabulation, Phantasiebegabung, besonderes Geltungsbedürfnis sowie Suggestibilität). Die Benutzung der Begriffe „allgemeine Glaubwürdigkeit“ und „spezielle Glaubwürdigkeit“ bei diesem Verständnis kann aber zu einer Verwechslung mit einem eher eigenschaftsorientierten Ansatz führen; außerdem wird eine Hierarchie der beiden Konzepte signalisiert, die in Wirklichkeit nicht besteht.

² Die an sich trivialen begrifflichen Klärungen gewinnen an Bedeutung, wenn man sich vergegenwärtigt, daß ein in der Vergangenheit zu beobachtender fehlerhafter Umgang mit dem Verdacht auf sexuellen Kindesmißbrauch durchaus mit konzeptionellen, d.h. auch begrifflichen Unklarheiten verbunden war. So kann zum Beispiel für die

In einer früheren Arbeit (Steller, Volbert & Wellershaus, 1993) wurden unterschiedliche Formen nicht erlebnisentsprechender Aussagen bei kindlichen Zeugen³ systematisiert (vgl. auch Steller & Volbert, 1997, S. 23), die prinzipiell auch für erwachsene Zeugen gelten. Diese Systematisierung soll hier wiedergegeben werden:

Tabelle 1: Kategorien nicht erlebnisentsprechender Kinderaussagen

Absichtliche Falschaussagen

1. Intentionale Falschaussage
2. Intentionaler Transfer
 - a) eines eigenen Erlebnisses
 - b) einer sonstigen Wahrnehmung auf den Beschuldigten

Fremdbeeinflussungen

3. Intentionale Induktion einer Falschaussage durch einen Dritten, die vom Kind subjektiv als unwahr erkannt, aber übernommen wird
4. Intentionale Induktion einer Falschaussage durch einen Dritten, die vom Kind subjektiv als wahre Aussage übernommen wird
5. Irrtümliche Induktion einer Falschaussage durch einen Dritten, die vom Kind subjektiv als unwahr erkannt, aber übernommen wird
6. Irrtümliche Induktion einer Falschaussage durch einen Dritten, die vom Kind subjektiv als wahre Aussage übernommen wird

Autosuggestionen

7. Unabsichtlicher falscher Transfer
 - a) eines Erlebnisses
 - b) einer sonstigen Wahrnehmung

Wormser Massenmißbrauchsprozesse vor dem Landgericht Mainz gezeigt werden, daß die lange Fokussierung auf die Problemstellung „Lüge oder Wahrheit“ in bezug auf die Kinderaussagen dem eigentlichen Sachverhalt nicht gerecht wurde. Angesichts der Aussageentstehungen und weiteren Entwicklungen bei den kindlichen Zeugen dieser Verfahren war von vornherein die Problemstellung „Suggestion oder Realitätsgehalt“ vorrangig, angesichts des geringen Alters einiger kindlicher Zeugen war die Frage nach einer möglichen Lüge über sexuelle Mißbrauchserfahrungen sogar völlig inadäquat.

³ Da überwiegend Mädchen betroffen sind und auch zahlreiche Diplompsychologinnen gutachterlich tätig sind, wird der zeitgeistige Hinweis herausgefordert, daß in der vorliegenden Ausarbeitung die männliche Form als grammatikalische Neutralität verstanden wird und keine anderen Implikationen enthält.

auf den Beschuldigten

8. Eingeschränkte bzw. aufgehobene Fähigkeit, zwischen eigenem Phantasieprodukt und Realität zu unterscheiden

Aus der Übersicht ergibt sich, daß den unterschiedlichen Möglichkeiten nicht erlebnisentsprechender Schilderungen unterschiedliche Vorbedingungen auf seiten des aussagenden Kindes entsprechen. So sind z.B. für absichtliche Falschbezeichnungen kognitive Fähigkeiten und Wissensbestände (z.B. über sexuelle Interaktionen) nötig, die nicht in jedem Alter ohne weiteres vorausgesetzt werden können. Bei dem Verdacht auf Fremdbeeinflussungen ist die Analyse potentieller suggestionsfördernder Bedingungen von besonderer Bedeutung, während bei der Hypothese von Autosuggestionen die psychopathologische Diagnosestellung zentral ist. Die globale Auftragserteilung von Gerichten zur Glaubwürdigkeitsbegutachtung von Zeugen kann also sehr unterschiedliche Problemkonstellationen betreffen.

In einem aussagepsychologischen Glaubhaftigkeitsgutachten müssen nicht regelmäßig alle denkbaren Unwahr-Kategorien fallbezogen abgehandelt werden. Vielmehr ist die Spezifizierung der zutreffenden Problemstellung bzw. der zu prüfenden Alternativhypothesen zur Wahrheitsannahme im konkreten Einzelfall bereits ein wesentlicher Teil des Begutachtungsprozesses.

1.2 Merkmalsorientierte Aussageanalyse

1.2.1 Zugrundeliegende Überlegungen

Prinzipiell sind verschiedene Zugänge zur Einschätzung der Glaubhaftigkeit von Aussagen denkbar. Köhnken (1990) unterschied u.a. verhaltensorientierte und inhaltsorientierte Ansätze. Bei allen geht es um die Frage, welche zu beobachtenden Begleiterscheinungen mit Täuschungen (Lügen) assoziiert sind und ob (und gegebenenfalls wie eindeutig) aus diesen auf Täuschung geschlossen werden kann. Verhaltensorientiert ist die Verwendung von Merkmalen des Sprechverhaltens (z. B. Sprechgeschwindigkeit, Sprechfehler), von nonverbalen Verhaltensmerkmalen (z. B. Mimik, Gestik) oder von psychophysiologischen Phänomenen (z. B. Hautleitfähigkeits- oder Blutdruckänderungen) zur Einschätzung der Glaubhaftigkeit. Bei inhaltsorientierten Ansätzen wird die Aussage selbst einer Analyse anhand bestimmter Merkmale unterzogen. Für Merkmale des Sprechverhaltens und für frei zu beobachtende Ausdruckserscheinungen (z. B. Mimik, Gestik) ist festzustellen, daß trotz umfangreicher Forschung ein diagnostisch verwertbarer Zusammenhang zur Glaubhaftigkeit einer Aussage nicht hinreichend gesichert ist. Die psychophysiologische Aussagebeurteilung wurde vom I. Strafsenat des BGH in zwei Urteilen vom 17.12.1998 für gerichtliche Verfahren als Beweismittel generell ausgeschlossen (1 StR 156/98 und 1 StR 258/98).⁴

⁴ Vgl. NJW 9, 1999, S. 657 ff. („Anmerkung der Schriftleitung: Die psychologischen Gutachten, sowie das vollständige BGH-Urteil sind in dem Sonderheft Juli 1999 der Praxis der Rechtspsychologie erschienen.“)

Der inhaltsanalytische Ansatz beruht auf der Konzeptualisierung einer Aussage als geistige Leistung, die bereits von Stern (1904) vorgenommen worden ist. Es stellt eine schwierige Aufgabe mit hoher Anforderung an die kognitive Leistungsfähigkeit eines Zeugen dar, eine Aussage über ein komplexes Handlungsgeschehen ohne eigene Wahrnehmungsgrundlage zu erfinden und gegebenenfalls über verschiedene Befragungen, d.h. auch über längere Zeiträume, relativ konsistent aufrechtzuerhalten bzw. zu reproduzieren.

Ein grundlegender Unterschied zwischen einer wahren und einer gelogenen Darstellung besteht darin, daß der aufrichtige Kommunikator seinen Bericht aus dem Gedächtnis rekonstruiert, während der lügende Zeuge seine Aussage aus dem gespeicherten Allgemeinwissen konstruieren muß. Ein theoretisches Modell zur Erklärung der Unterschiede zwischen wahrheitsgemäßen und erfundenen Aussagen muß von daher berücksichtigen, wie derartiges Allgemeinwissen gespeichert ist. Dieser Frage wird mit Hilfe sogenannter Schematheorien nachgegangen. Kognitive Schemata sind abstrakte Wissensstrukturen, die Aufmerksamkeit, Erwartungen, Interpretationen und Inferenzen bei der Wahrnehmung, Verarbeitung und Rekonstruktion von Informationen leiten. Sie enthalten eine Zusammenfassung der Eigenschaften, die typischerweise in einem Exemplar des jeweiligen Gegenstandsreichs vorkommen. Ein Ereignisschema enthält also nicht spezifische, sondern für dieses Ereignis typische Informationen. Wenn jemand einen nicht selbst wahrgenommenen Sachverhalt schildert, steht ihm als Grundlage das abstrakte Schemawissen zur Verfügung. Inkonsistente und irrelevante Details, die bei einer realen Wahrnehmung als „Anhang“ zum Schema gespeichert werden, finden sich hier nicht. In Schilderungen, die aus vorhandenem Schemawissen konstruiert werden, sind von daher vor allem elementare, direkt zum Handlungsziel hinführende Handlungssequenzen zu erwarten. Die Wahrscheinlichkeit, daß auch nebensächliche Details, abgebrochene Handlungsketten, unerwartete Komplikationen oder gar phänomengemäße Schilderungen unverstandener Handlungselemente beschrieben werden, ist dagegen ausgesprochen gering (Köhnken, 1990).

Ein zweiter Unterschied zwischen einem aufrichtigen und einem lügenden Kommunikator betrifft die Selbstpräsentation. Ein lügender Kommunikator verfolgt das Ziel, bei den Rezipienten den Eindruck eines glaubwürdigen Kommunikators, also einen falschen Eindruck zu erzeugen, um so die Wirksamkeit der falschen Aussage zu unterstützen. Hierzu greift der Kommunikator auf Vorstellungen darüber zurück, welche Verhaltensweisen, Äußerungen etc. einen solchen Eindruck bei dem Rezipienten bewirken und welche umgekehrt zum Verdacht der Unglaubwürdigkeit führen. Von daher ist anzunehmen, daß falsche Aussagen in der Regel nur in geringem Ausmaß Selbstkorrekturen, Zugeben von Erinnerungslücken, Selbstbelastungen o.ä. enthalten (Köhnken, 1990).

Erfundene Handlungsschilderungen werden also – je nach gegebener Leistungsfähigkeit des Aussagenden – inhaltlich relativ wenig elaboriert ausfallen, da der lügende Zeuge ein erhebliches Ausmaß seiner kognitiven Energie auf kreative Prozesse und auf Kontrollprozesse verwenden muß. Daraus ergibt sich, daß eine erfundene Handlungsschilderung im *intraindividuellen* Vergleich wahrscheinlich eine geringere inhaltliche Qualität aufweist als eine wahre Bekundung über ein Erlebnis.⁵

In erlebnisbegründeten Schilderungen ist häufig ein hohes Ausmaß an Detaillierung und individueller Durchzeichnung festzustellen. Inhaltliche Besonderheiten, deren Erfindung eine gewisse Kompetenz voraussetzen würde, treten in erlebnisbegründeten Aussagen sozusagen von allein auf. Solche inhaltlichen Besonderheiten können in der Schilderung von Begleitgefühlen zu dem Erlebnis oder in ausgefallenen Details, dem Erwähnen von Nebensächlichkeiten oder von Komplikationen bzw. Handlungsabbrüchen bestehen. Weitere inhaltliche Besonderheiten können in Gesprächswiedergaben oder in Schilderungen von Mißverständnissen während solcher Gespräche deutlich werden. Mit anderen Worten: Bei der Inhaltsanalyse geht es um die Identifizierung von Aussagequalitäten, die über die reine Handlungsschilderung hinausgehen, welche ja auch in einer belastenden Lüge enthalten sein muß.

In der Literatur wurden diverse Merkmale mit Indikatorwert für den Erlebnisgehalt einer Schilderung beschrieben (sogenannte Glaubwürdigkeitsmerkmale bzw. Realkennzeichen). Diese wurden im letzten Jahrzehnt auch einer empirischen Überprüfung unterzogen. Grundsätzlich konnte die Annahme über eine unterschiedliche Qualität erlebnisbegründeter und erfundener Handlungsschilderungen durch empirische Forschung nicht nur bei Kindern bestätigt werden (vgl. Kapitel 3).⁶

Für die Praxis ist folgendes zu betonen: Die Aussageanalyse führt in einem ersten Schritt lediglich zu einer Beurteilung der Qualität einer Aussage. Diese Qualitätsanalyse gewinnt ihre Aussagekraft für die Glaubhaftigkeitsbeurteilung erst durch ihren Bezug auf die spezifischen Kompetenzen und Erfahrungen des Aussagenden. Je nach Alter, geistiger Leistungsfähigkeit und be-

⁵ Diese Arbeitshypothese der inhaltsorientierten Glaubhaftigkeitsbeurteilung wurde von Undeutsch (1967, S. 126) herausgearbeitet und daher von Steller (1989) als Undeutsch-Hypothese bezeichnet, was in der internationalen Literatur aufgegriffen wurde.

⁶ Aus den bisherigen Erörterungen ergibt sich, daß eine Grundvoraussetzung für die Anwendung der Methodik in einem gewissen Mindestumfang der Aussage liegt. Einfache Negationen und sehr knappe Sachverhaltsdarstellungen eignen sich nicht für dieses Verfahren. Damit hängt die Anwendbarkeit des Verfahrens nicht zuletzt auch von der Komplexität des in Frage stehenden Geschehens ab.

reichsspezifischer Erfahrung des Aussagenden ist der Indikatorwert der Qualitätseinschätzung für die Glaubhaftigkeitsbeurteilung unterschiedlich zu bewerten. Jede schematische Verwendung von inhaltsorientierten oder anderen Realkennzeichen im Sinne einer Checklisten-diagnostik und jede Quantifizierung im Sinne der Festlegung von allgemeingültigen Schwellenwerten (Cut-off-Scores) ist fehlerhaft und der aussagepsychologischen Begutachtungsmethodik nicht angemessen. Der Indikatorwert einer Qualitätsanalyse für die Beurteilung der Erlebnisbegründetheit einer Schilderung ergibt sich in der Praxis nur in der intraindividuellen Perspektive (vgl. Arntzen, 1993; Greuel et al., 1998).

1.2.2 Systeme für merkmalsorientierte Aussageanalysen

Umfassende Beschreibungen von sogenannten Glaubwürdigkeitsmerkmalen mit erläuternden Fallbeispielen wurden von Undeutsch (1967) erstellt, von Arntzen (1. Auflage 1970, 3. Auflage 1993) stammt eine Systematisierung mit der Unterscheidung von Merkmalen, die auf den Inhalt, die Aussageweise und die Aussagemotivation bezogen sind. Im Lehrbuch über forensische Psychologie von Dettenborn, Fröhlich und Szewczyk (1984) aus der ehemaligen DDR ist ebenfalls eine Beschreibung von Glaubwürdigkeitsmerkmalen enthalten, wobei besonders auf Szewczyk und Littmann (1982) Bezug genommen wird. In Schweden hat Trankell (1971) Glaubwürdigkeitsmerkmale beschrieben. Die derzeit häufig zitierte Aufstellung von Steller und Köhnken (1989) erfolgte explizit ohne Anspruch, eigene Glaubwürdigkeitsmerkmale zu begründen, sondern mit dem Ziel, ausschließlich die vorgefundenen Merkmale zur Erfassung inhaltlicher Qualitäten darzustellen, um ein System zur Verfügung zu haben, das in empirischer Forschung verwendet werden kann.⁷ Da die Merkmalsliste von Steller und Köhnken tatsächlich in internationaler Forschung häufige Verwendung fand und zum Teil (wohl wegen ihrer Transparenz) auch in forensisch-aussagepsychologischer Begutachtungspraxis benutzt wird, soll sie auch hier referiert werden. Gleichzeitig wird auf die entsprechenden ausführlichen Darstellungen bei Greuel et al. (1998, S. 89 ff.) verwiesen.

Tabelle 2: Realkennzeichen in der Kategorisierung von Steller und Köhnken (1989)

ALLGEMEINE MERKMALE

1. Logische Konsistenz
2. Ungeordnet sprunghafte Darstellung
3. Quantitativer Detailreichtum

⁷ Aus den Ausführungen in Abschnitt 1.2.1 ergibt sich, daß es abwegig ist, diese Merkmalsliste als eine Art Test zur Prüfung der Glaubhaftigkeit von Aussagen zu verstehen. Derartige Mißverständnisse sind aber in Praxis und Forschung festzustellen.

SPEZIELLE INHALTE

4. Raum-zeitliche Verknüpfungen
5. Interaktionsschilderung
6. Wiedergabe von Gesprächen
7. Schilderung von Komplikationen im Handlungsverlauf

INHALTLICHE BESONDERHEITEN

8. Schilderung ausgefallener Einzelheiten
9. Schilderung nebensächlicher Einzelheiten
10. Phänomengemäße Schilderung unverständener Handlungselemente
11. Indirekt handlungsbezogene Schilderungen
12. Schilderung eigener psychischer Vorgänge
13. Schilderung psychischer Vorgänge des Angeschuldigten

MOTIVATIONSBEZOGENE INHALTE

14. Spontane Verbesserungen der eigenen Aussage
15. Eingeständnis von Erinnerungslücken
16. Einwände gegen die Richtigkeit der eigenen Aussage
17. Selbstbelastungen
18. Entlastung des Angeschuldigten

DELIKTSPEZIFISCHE INHALTE

19. Deliktspezifische Aussageelemente

Tabelle 2 zeigt 19 inhaltliche Merkmale für die Qualitätsanalyse. Es handelt sich um Merkmale zur Analyse *einer* Aussage. Indikatoren, die sich aus dem Vergleich von wiederholten Aussagen (über denselben Sachverhalt) erschließen lassen (Konstanz, Widersprüche, Auslassungen und Ergänzungen), sind daher nicht in dieser Kategorisierung enthalten, sie sind aber bei der praktischen Begutachtung zu berücksichtigen (vgl. Abschnitt 1.3).

Die merkmalsorientierte Inhaltsanalyse als Teil der Qualitätsanalyse umfaßt fünf Kategorien von Realkennzeichen. In dem Komplex „Allgemeine Merkmale“ werden Merkmale zusammengefaßt, die sich auf eine Zeugenaussage in ihrer Gesamtheit beziehen. Sie können ohne Rückgriff auf Einzelheiten des Aussageinhalts beurteilt werden. „Logische Konsistenz“ (Arntzen, 1993, S. 50; Trankell, 1971, S. 126; Undeutsch, 1967, S. 138) und „Quantitativer Detailreichtum“ (Arntzen, 1993, S. 27 f.; Dettenborn et al., 1984, S. 313) als Realkennzeichen sind ohne weiteres verständlich, während das Merkmal „Unstrukturierte Darstellung“ einer weiteren Definition bedarf: Es wird angenommen, daß Falschaussagen eher durch eine kontinuierliche, strukturierte, meist chronologische Darstellungsweise gekennzeichnet sind. Eine unstrukturierte Darstellung wird als Realkennzeichen dann angesehen, wenn die unstrukturierte Schilderung dennoch zu einem geschlossenen, logisch konsistenten Bild rekonstruiert werden kann (Arntzen, 1993, S. 78). Bei Verstößen

gegen die logische Konsistenz wird eine unstrukturierte Darstellung nicht im Sinne eines Realkennzeichens gewertet.

In zwei weiteren Kategorien sind Merkmale zusammengefaßt, die sich auf einzelne Inhalte einer Aussage beziehen. Für diese zehn Aussagemerkmale ist das Ausmaß der Konkretheit und Anschaulichkeit einer Schilderung entscheidend. In der Kategorie „Spezielle Inhalte“ sind die vier Merkmale „Raum-zeitliche Verknüpfungen“ (Arntzen, 1993, S. 38; Dettenborn et al., 1984, S. 313; Undeutsch, 1967, S. 139 f.), „Interaktionsschilderungen“ (Arntzen, 1993, S. 35), „Wiedergabe von Gesprächen“ (Arntzen, 1993, S. 32; Dettenborn et al., 1984, S. 314) und „Schilderung von Komplikationen im Handlungsverlauf“ (Arntzen, 1993, S. 36; Dettenborn et al., 1984, S. 313; Undeutsch, 1967, S. 153) zusammengefaßt. „Inhaltliche Besonderheiten“ bilden eine dritte Kategorie von Aussagemerkmalen. Sie enthält die „Schilderung ausgefallener Einzelheiten“ (Arntzen, 1993, S. 32; Dettenborn et al., 1984, S. 313; Trankell, 1971, S. 126; Undeutsch, 1967, S. 138) sowie die „Schilderung nebensächlicher Einzelheiten“ (Arntzen, 1993, S. 27; Dettenborn et al., 1984, S. 313). Das Merkmal „Phänomengemäße Schilderung unverstandener Handlungselemente“ ist dann erfüllt, wenn ein Sachverhalt zutreffend geschildert wird und gleichzeitig deutlich wird, daß der Aussagende eben diesen Sachverhalt nicht in seiner Bedeutung erkannt hat. Beispiele aus dem Bereich der Glaubhaftigkeitsbegutachtung kindlicher Zeugenaussagen in Fällen des sexuellen Mißbrauchs sind die Schilderung von Ejakulationen oder Orgasmuszuständen durch junge Kinder (Arntzen, 1993, S. 31 f.; Dettenborn et al., S. 314; Trankell, 1971, S. 125; Undeutsch, 1967, S. 141). Als „Indirekt handlungsbezogene Schilderungen“ im Sinne eines Realkennzeichens gelten Inhalte einer Aussage, die den aktuell berichteten Handlungen ähnlich sind oder sich auf sie beziehen, aber zu anderer Zeit und mit anderen Personen stattgefunden haben (Arntzen, 1993, S. 37). Die „Schilderung eigener psychischer Vorgänge“ (Arntzen, 1993, S. 29; Dettenborn et al., 1984, S. 314; Trankell, 1971, S. 124 f.; Undeutsch, 1967, S. 145) und die „Schilderung psychischer Vorgänge des Angeschuldigten“ (Dettenborn et al., 1984, S. 314; Undeutsch, 1967, S. 143) sind weitere inhaltliche Besonderheiten, deren Vorkommen als Hinweise auf reale Erlebnisgrundlagen einer Schilderung gewertet werden.

Eine vierte Gruppe von Merkmalen besteht aus Aussageinhalten, die die Selbstpräsentation des Aussagenden betreffen und somit Rückschlüsse auf seine Motivation zulassen. Es handelt sich um inhaltliche Realkennzeichen, da sie aus dem Inhalt einer Aussage erkannt werden können, ohne daß direkt im Hinblick auf die Aussagemotivation exploriert wird. Während bei den Gruppen „Spezielle Inhalte“ und „Inhaltliche Besonderheiten“ der kognitive Aspekt im Vordergrund steht (die untersuchungsleitende Frage ist, ob sich ein falschaussagender Zeuge Inhalte der beschriebenen Qualität ausgedacht haben *kann*), dominiert bei der vierten Gruppe der motivationale Aspekt. Die

Frage lautet hier, wie wahrscheinlich es ist, daß ein falschaussagender Zeuge Inhalte der beschriebenen Art in seine Aussage aufnehmen würde.

Neben „Spontanen Verbesserungen der eigenen Aussage“ (Dettenborn et al., 1984, S. 313; Undeutsch, 1967, S. 152 f.), dem „Eingeständnis von Erinnerungslücken“ (Littmann & Szewczyk, 1983, S. 64) sowie „Einwendungen gegen die Richtigkeit der eigenen Aussage“ (Dettenborn et al., 1984, S. 314; Undeutsch, 1967, S. 153) ist die Erwähnung selbstbelastender Details (Dettenborn et al., 1984, S. 314; Undeutsch, 1967, S. 153) beziehungsweise solcher, die den Beschuldigten entlasten, oder mindestens der Verzicht auf naheliegende Mehrbelastungen des Beschuldigten (Dettenborn et al., 1984, S. 314; Undeutsch, 1967, S. 153) zu berücksichtigen.

Aussageelemente, die in typischer Weise mit dem behaupteten Delikt in Verbindung stehen, werden in einer gesonderten Kategorie betrachtet. Der Sonderstatus des Kriteriums „Deliktstypische Aussageelemente“⁸ ergibt sich dadurch, daß zur Beurteilung delikttypischer Aussageinhalte empirisch-kriminologisches Wissen über „typische“ Begehungsformen von Sexualdelikten (an Kindern) nötig ist (Dettenborn et al., 1984, S. 314; Undeutsch, 1967, S. 137).

Die beschriebenen Qualitätsmerkmale können in einer Aussage in unterschiedlicher Weise vorhanden sein. Arntzen (1993) unterscheidet zwischen „Aussageeigenarten“ und „Glaubwürdigkeitsmerkmalen oder -kriterien“. Bei den Aussageeigenarten handelt es sich um sehr einfache Ausprägungen der dargestellten Merkmale, die für die Beurteilung irrelevant sind, da sie bis zu einem gewissen Grade auch in Falschbekundungen zu finden sind. „Erst wenn die Aussageeigenart eine bestimmte Steigerung ihrer Qualität aufweist [...], wird sie zum Glaubwürdigkeitsmerkmal“ (ebd., S. 20). Bei der Feststellung der Einzelmerkmale muß also beurteilt werden, ob eine Aussageeigenart quantitativ und/oder qualitativ so ausgeprägt ist, daß sie zu einem Qualitätsmerkmal wird. Dieser Beurteilungsprozeß erfordert bereits eine Berücksichtigung der spezifischen Kompetenzen und Vorerfahrungen des Zeugen, der Komplexität des vorgebrachten Geschehens sowie der Befragungsbedingungen (freier Bericht oder gezielte Fragen).

Eine Fokussierung allein auf die Anzahl erfüllter Qualitätsmerkmale ist irreführend und entspricht nicht der zugrundeliegenden Auswertungslogik. Im Einzelfall können auch einzelne Qualitätsmerkmale ausreichen, um den Er-

⁸ Es handelt sich um ein inhaltliches Merkmal einer Kinderaussage, da es sich auf solche delikttypischen Schilderungen (z. B. Schweigegebot, Progression der Handlungsintensität) bezieht, die von kindlichen Zeugen berichtet werden. Bei Anwendung des Gesichtspunktes „Delikttypizität“ auf das Gesamt aller Tatsachen eines Falles handelt es sich nicht mehr um ein inhaltliches Merkmal *einer* Aussage (so Raskin & Esplin, 1991, und Horowitz 1991).

lebnisbezug einer Aussage anzunehmen, beispielsweise ist das Merkmal „Phänomengemäße Schilderung unverstandener Handlungselemente“ in prägnanter Ausprägung von einem jungen Kind nicht simulierbar (wenn ausgeschlossen werden kann, daß ein entsprechender Aussageinhalt im Rahmen von suggestiven Vorbefragungen erlernt wurde). Dagegen reicht das Vorliegen einer Reihe von wenig prägnanten Qualitätsmerkmalen bei einer gut begabten jugendlichen Zeugin oft nicht aus, um einen Erlebnisbezug zu belegen.

Greuel et al. (1998) unterscheiden außerdem zwischen Ausschluß- und Qualifizierungsmerkmalen, wobei sich die von ihnen genannten inhaltlichen Merkmale nicht alle in der Systematisierung von Steller und Köhnken (1989) finden: Nach dieser Unterscheidung kann der Erlebnisbezug der Aussage nicht mehr mit der im forensischen Kontext gebotenen Zuverlässigkeit bestätigt werden, wenn folgende Merkmale „im Rahmen des individuums- und sachverhaltsbezogenen Erwartbaren“ (ebd., S. 162) nicht erfüllt sind: Logische Konsistenz, Detaillierungsgrad und Konstanz der Aussage. Bei diesen Merkmalen handelt es sich demnach um Ausschlußmerkmale, also um notwendige, aber nicht hinreichende Bedingungen, die für sich genommen zum Beleg eines Erlebnisbezugs nicht herangezogen werden können. Als besonders bedeutsame Qualifizierungsmerkmale von hoher diagnostischer Valenz nennen die Autoren: Schilderungen von Handlungskomplikationen, Phänomenorientierte Schilderungen unverstandener Handlungselemente, Schilderung des Erlebens phänomenaler Kausalität, Schilderung origineller Details, Schilderung von Interaktionsketten, Schilderung von Wirklichkeitskontrollen, *Spontane* Präzisierbarkeit der Aussage.

Die Bezeichnung der inhaltlichen Qualitätsmerkmale als Glaubwürdigkeitsmerkmale bzw. Realkennzeichen ist nicht zufällig erfolgt. Es handelt sich durchweg um Merkmale, deren Vorhandensein (unter den beschriebenen Voraussetzungen) auf den Erlebnisgehalt einer Schilderung hinweist. Bei Fehlen der Merkmale ist der Umkehrschluß auf eine Lüge nicht gerechtfertigt. Das Fehlen von Glaubhaftigkeitsmerkmalen kann zwar durch eine Lüge (fehlende Erlebnisgrundlage) bedingt sein, kann aber auch durch andere Faktoren (z. B. Hemmungen, Angst, Gedächtnismängel) verursacht sein. Der beschriebene inhaltsorientierte Ansatz ist also eine Methode zur Substantiierung des Erlebnisgehalts von Aussagen, er ist keine Methode zur „Lügendetektion“.

1.3 Integrierende Glaubhaftigkeitsbeurteilung

Es wurde bereits betont: Schlußfolgerungen über den Erlebnisbezug (Wahrheitsgehalt) einer Aussage erfolgen im Rahmen einer aussagepsychologischen Begutachtung in Form eines komplexen Urteilsvorgangs. Inhaltliche Qualitätsmerkmale einer Aussage dürfen nicht im Sinne eines Checklistenverfahrens mißverstanden werden, bei dem allein eine bestimmte Anzahl

von Merkmalen für die Glaubhaftigkeit einer Aussage spricht. Vielmehr erfolgt mit der merkmalsorientierten Inhaltsanalyse nur eine Einschätzung eines Aspekts der Qualität einer Aussage, zur Glaubhaftigkeitsbeurteilung ist diese Aussagequalität dann auf die personalen Voraussetzungen des Zeugen sowie auf die Entstehungs- und weitere Entwicklungsgeschichte der Aussage zu beziehen. Außerdem sind weitere Analyseergebnisse zu beachten. Neben der inhaltlichen Qualitätseinschätzung, die für *eine* Aussage vorgenommen werden kann und die bisher behandelt wurde (aussageimmanente Qualitätsmerkmale, Greuel et al., 1998), geht es um das *Aussageverhalten* eines Zeugen sowie um aussageübergreifende Qualitätsmerkmale, die sich aus dem Vergleich von Aussagen über denselben Sachverhalt zu unterschiedlichen Zeitpunkten ergeben (*Konstanzanalyse*). Zur abschließenden Beurteilung der Glaubhaftigkeit sind jedoch auch externe und interne Rahmenbedingungen der Aussage zu berücksichtigen, um festzustellen, ob eine qualitativ hochwertige und infolgedessen einen Erlebnisbezug indizierende Aussage aus aussagepsychologischer Sicht als zuverlässig bewertet werden kann. Diesbezügliche Einschränkungen können sich insbesondere ergeben, wenn die Aussage kein originäres „Produkt“ des Aussagenden ist, sondern durch *suggestive Einflüsse* ganz oder teilweise determiniert ist. Eine weitere Fehlerquelle kann in *psychologischen Besonderheiten* des Zeugen liegen. Zur Fehlerquellenanalyse kann auch die *Motivationsanalyse* gerechnet werden, durch die Hypothesen für eine intentionale Falschbezeichnung aufgestellt werden.

Im folgenden werden die über die Inhaltsanalyse hinausgehenden Analysebereiche näher erläutert:

Trotz der prinzipiellen Uneindeutigkeit *begleitender Ausdruckserscheinungen* im Hinblick auf die Glaubhaftigkeit der Schilderungen ist das Ausdrucksgeschehen des Zeugen Gegenstand der Beobachtung und Analyse im Rahmen von Glaubhaftigkeitsbegutachtungen. Zum Aussageverhalten gehört insbesondere die Gefühlsbeteiligung bei der Schilderung der in Frage stehenden Geschehnisse. Zutreffend weist Michaelis-Arntzen (in Arntzen, 1993, S. 70 ff.) darauf hin, daß Äußerungen von Angst, Scham und Peinlichkeit für sich allein nicht als Glaubhaftigkeitshinweise gewertet werden können. Gerade relativ einfache und durchgehende Gefühlszustände können durchaus simuliert werden. Auch das Fehlen von Begleitaffekten ist nicht eindeutig zu interpretieren: Außer dem Bedürfnis, als Zeuge gut zu „funktionieren“ (und daher eher sachlich zu berichten), können zeitlicher Abstand und wiederholte Befragungen das emotionale Nacherleben auch tatsächlich verringern. Dagegen kann das Auftreten verschiedener Gefühlsqualitäten, ein bei der Reproduktion des Geschehens gefühlsmäßiges Nacherleben, das dem geschilderten Geschehensablauf mit den dabei aufgetretenen wechselnden Gefühlen (z.B. eine Abfolge von Überraschung, Hilflosigkeit, Angst, Entrüstung) entspricht, einen Hinweis für einen Erlebnisbezug darstellen. Wie bei den inhaltlichen Glaubhaftigkeitsmerkmalen sind bei der Interpretation von Gefühlsäußerungen während einer Aussage die Zeugenpersönlichkeit und die spezifische

Situation zu berücksichtigen. Die Analyse dieser Interaktionen erlaubt zuweilen Interpretationen, die indiziellen Wert für die Glaubhaftigkeitseinschätzung einer konkreten Bekundung haben können.

Liegen Aussagen eines Zeugen von verschiedenen Zeitpunkten vor, wie es im strafrechtlichen Kontext in der Regel der Fall ist (polizeiliche Erstvernehmung, gegebenenfalls staatsanwaltschaftliche oder richterliche Vernehmung, Exploration durch den psychologischen Sachverständigen), so ist die *Konstanzprüfung* neben der Inhaltsanalyse ein weiteres wesentliches methodisches Element der Aussageanalyse. Im Rahmen der Konstanzprüfung ist ein Aussagevergleich im Hinblick auf Übereinstimmungen, Widersprüche, Ergänzungen und Auslassungen anzustellen. Erst in einem nächsten Schritt sind diese Feststellungen zu bewerten. Wichtig ist, daß nicht jede Inkonstanz einen Hinweis auf mangelnde Glaubhaftigkeit darstellt, sondern daß Gedächtnisunsicherheiten häufig eine hinreichende Erklärung für festzustellende Varianz des Aussageinhalts darstellen können. Spontane Ergänzungen und nachträgliche Präzisierungen können im Einzelfall einen wesentlichen Hinweis auf den Realitätsgehalt einer Bekundung darstellen (vgl. Arntzen, 1993, S. 41 ff.). Die Konstanzanalyse wird als Bestandteil der Qualitätsanalyse verstanden.

Der Schluß von einer hohen inhaltlichen Qualität auf eine positive Einschätzung der Glaubhaftigkeit einer Aussage ist nur gerechtfertigt, wenn keine moderierenden Einflüsse in einer Weise gewirkt haben, die die hohe Aussagequalität auch anders als durch einen Erlebnisbezug erklären ließen. Daher ist eine Fehlerquellenanalyse sozusagen als Zwischenstufe zwischen Personen- und Aussageanalyse einzufügen. Ist die Frage der *suggestiven Beeinflussung* zu prüfen, muß die Rekonstruktion der Aussageentstehung und ihrer weiteren Entwicklung im Vordergrund stehen. Die Rekonstruktion der Aussagegenese erstreckt sich auf Umstände und Inhalte von Befragungen, die in die qualitätsanalytische Bewertung einbezogen werden müssen. Damit ergibt sich der enge Bezug der Verwertbarkeit von Ergebnissen der Qualitätsanalyse mit der Adäquatheit von Befragungen (Explorationen, Vernehmungen), die zu der zu bewertenden Aussage geführt haben oder ihr vorausgegangen sind.

Zur Fehlerquellenanalyse kann auch die *Motivationsanalyse* gerechnet werden. Wesentliche Anhaltspunkte für potentielle Belastungsmotive können die Analyse der Beziehung zwischen Zeugen und Beschuldigten und insbesondere die Analyse der Konsequenzen der Anschuldigung für den Zeugen bzw. für den Beschuldigten oder beteiligte Drittpersonen sein. Überlegungen, wer Vor- und wer Nachteile durch die Beschuldigung bzw. eine Verurteilung des Beschuldigten haben könnte, können Hinweise für die Hypothesenbildung über mögliche Belastungsmotive sein.

Zur Fehlerquellenanalyse gehört auch die personenbezogene Fragestellung, ob im Einzelfall vorliegende *Besonderheiten des Erlebens und Verhaltens* (z. B. Selbstwertprobleme und Geltungsbedürfnis, Neurotizismus) oder vorliegende Persönlichkeitsstörungen einen bedingenden Faktor für eine mögliche Falschaussage darstellen können.

Erkennbare Belastungsmotivationen oder problematische Dispositionen auf seiten der aussagenden Person sind natürlich keine zwingenden Hinweise auf das Vorliegen einer Falschaussage, sie sind aber bei der abwägenden Bewertung der Interaktion von Kompetenz- und Qualitätsanalyse als mögliche moderierende Einflüsse zu berücksichtigen zum Beispiel dahin gehend, daß bei gegebener Belastungsmotivation von einer hohen Aktivierung vorhandener Fähigkeiten ausgegangen wird. Auch Greuel et al. (1998, S. 174 ff.) weisen darauf hin, daß Schlußfolgerungen auf den Wirklichkeitsgehalt einer Aussage aufgrund der Motivationsanalyse allein höchst ungewiß sind. Sie fordern aber übereinstimmend mit der hiesigen Auffassung eine Motivationsanalyse in aussagepsychologischen Gutachten, da sie die Interpretationsbasis für Glaubhaftigkeitseinschätzungen im Verbund mit Kompetenz- und Qualitätsanalyse trotz ihrer Unspezifität erweitern kann.

Der unauflöslche interaktive Bezug von Aussagequalität (inkl. Konstanz und Aussageweise), personaler Kompetenz und Aussagegeschichte (Bedingungen der Erstaussage und der weiteren Aussageentwicklung) determiniert das praktische Vorgehen der Datenerhebung und das diagnostische Schlußfolgern bei der Glaubhaftigkeitsbeurteilung von Zeugenaussagen. Die Gewichtung der einzelnen Analysebereiche ist immer einzelfallspezifisch bestimmt. Die adäquate Spezifizierung, d.h. die Formulierung der zutreffenden Fragestellungen und der zu prüfenden Hypothesen, stellt bereits einen wesentlichen Teil des Begutachtungsprozesses dar. Denn neben der Hypothese, daß die Aussage mit hoher Wahrscheinlichkeit auf einem eigenen Erleben basiert, sind immer Alternativhypothesen zu spezifizieren und zu prüfen, wobei die Datenerhebung durch die aufgestellten und im Laufe der Untersuchung aktualisierten Hypothesen determiniert wird.

Hilfreich für die Integration der Ergebnisse aller Analysebereiche für die Glaubhaftigkeitsbeurteilung einer konkreten Aussage ist dabei die von uns so genannte Leitfrage der Glaubhaftigkeitsbeurteilung:

Könnte dieser Zeuge mit den gegebenen *individuellen Voraussetzungen* unter den *gegebenen Befragungsumständen* und unter Berücksichtigung der im konkreten Fall *möglichen Einflüsse von Dritten* diese *spezifische Aussage* machen, ohne daß sie auf einem realen Erlebnishintergrund basiert?

Die kursiv gesetzten Elemente der Leitfrage verweisen auf die Wechselwirkung von Kompetenz eines Aussagenden (*individuelle Voraussetzungen*) und Qualität einer Aussage (*spezifische Aussage*), deren Indikatorwert für die Glaubhaftigkeit einer Aussage modifiziert (z. B. beeinträchtigt bzw. aufge-

hoben) werden kann durch Bedingungen der Aussageentstehung und ihrer weiteren Entwicklung. Die Notwendigkeit der möglichst vollständigen Rekonstruktion der Aussageentwicklung für die Glaubhaftigkeitsbeurteilung kann insofern gar nicht überbetont werden.

Das methodische Prinzip, das durch die Leitfrage der Aussagebeurteilung verdeutlicht wird, ist in den empirischen Wissenschaften unbestritten. Es besteht darin, daß ein zu überprüfender Sachverhalt (hier Glaubhaftigkeit der spezifischen Aussage) so lange negiert wird, bis diese Negation mit gesammelten Fakten nicht mehr vereinbar ist. Der wissenschaftlich ausgebildete psychologische Sachverständige arbeitet (gedanklich) also zunächst mit der Unwahrannahme als sogenannter Nullhypothese. Ergeben seine Prüfstrategien, daß die Unwahrhypothese mit den vorliegenden Fakten nicht kompatibel ist, wird sie verworfen, und es gilt dann die Alternativhypothese, die Wahrheitsannahme. Der skizzierte methodische Ansatz korrespondiert mit dem heutigen Verständnis von psychologischer Begutachtung als hypothesengeleitete problemorientierte Entscheidungsstrategie (vgl. z. B. Jäger, 1983; Steller, 1994). Das Prinzip wissenschaftlichen Denkens (im Sinne der Beibehaltung der Nullhypothese bis zu ihrer Falsifikation) ist auch mit dem juristischen Prinzip der Unschuldsvermutung kompatibel.

2 Praxis forensisch-aussagepsychologischer Begutachtungen

2.1 Erhebungsbereiche und -methoden

2.1.1 Allgemeine Überlegungen

Aus den im ersten Kapitel dargestellten methodischen Prinzipien von aussagepsychologischen Begutachtungen läßt sich das praktische Vorgehen im Einzelfall ableiten. Der inhaltsanalytische Zugang zur Glaubhaftigkeitsbeurteilung erfordert zwingend eine Analyse aussagerelevanter Kompetenzen des Zeugen zur Erhebung eines individuellen Vergleichsstandards. Neben der Feststellung seiner allgemeinen und sprachlichen intellektuellen Leistungsfähigkeit unter Einschluß kreativer Komponenten ist bei Sexualdelikten eine Einschätzung sexualbezogener Kenntnisse und Erfahrungen nötig. Allgemein geht es um die Feststellung bereichsspezifischer Kompetenzen und Erfahrungen des Zeugen zur Hypothesenbildung, ob die im Einzelfall vorfindbare Aussagequalität durch sogenannte Parallelerlebnisse oder reine Erfindung erklärbar sein könnte. Die Beurteilung der persönlichen Kompetenzen eines Zeugen und etwaiger aussagerelevanter Besonderheiten seiner Persönlichkeitsentwicklung (z. B. Selbstwertprobleme, Geltungsbedürfnis, Neurotizismus) erfolgt mit den allgemeinen Methoden der psychologischen Diagnostik: Befragung, Beobachtung, (standardisierte) Tests und Fragebögen sowie (freie bzw. wenig standardisierte) Proben (z. B. Phantasie- und Suggestionstests), unter Umständen projektive Verfahren. Da diese Erhebungsmethoden Bestandteile der allgemeinen Ausbildung von Diplompsychologen in Psychodiagnostik sind, werden in der aussagepsychologischen Literatur dazu in der Regel keine umfangreichen Ausführungen gemacht. Auch hier kann

keine methodische Anleitung zur Leistungs- und Persönlichkeitsdiagnostik erfolgen, auf einige praxisrelevante Fragen wird in Abschnitt 2.1.3 gesondert eingegangen.

Die Kompetenzanalyse bildet einerseits den Bezugspunkt für die Qualitätsanalyse, deren Logik und Methodik in Kapitel 1 dargestellt wurden. Andererseits dient die Kompetenzanalyse auch zur Beurteilung der sogenannten Ausagetüchtigkeit (Synonym: Zeugentüchtigkeit), also der prinzipiellen Fähigkeit des Zeugen zur Wahrnehmung, Speicherung und Reproduktion von Sachverhalten der in Frage stehenden Art und Komplexität und zur Realitätskontrolle.

Das Material zur Analyse der Aussagequalität wird mit Hilfe einer Exploration zum im Frage stehenden Sachverhalt erhoben (zu Einzelheiten vgl. Abschnitt 2.1.2). Relevante Anknüpfungstatsachen für die Rekonstruktion der Aussageentstehung und ihrer weiteren Entwicklung sowie für die Motivationsanalyse liegen häufig bereits in den Ermittlungsakten vor, zusätzliche Informationen kann der Sachverständige durch Befragung des zu begutachtenden Zeugen gewinnen. Inwieweit der aussagepsychologische Sachverständige auch Befragungen von weiteren Personen durchführen sollte oder (aus rechtlicher Sicht) durchführen darf, soll wegen der besonderen Bedeutung in einem gesonderten Abschnitt diskutiert werden (vgl. Abschnitt 2.2).

Das in Abschnitt 1.3 dargestellte Konzept von aussagepsychologischer Begutachtung als hypothesengeleiteter, integrierender Entscheidungsprozeß beinhaltet, daß eine standardisierte Routine zur Durchführung praktischer Fallbearbeitungen nicht festgelegt werden kann. Vielmehr richten sich Methodenauswahl und Analyseschwerpunkte nach den Gegebenheiten des Einzelfalls.

2.1.2 Exploration zur Sache

Die wesentliche aussagepsychologische Erhebungstechnologie besteht in der Exploration zur Sache. Dabei sollte zunächst immer versucht werden, durch eine entsprechende Aufforderung einen zusammenhängenden Bericht des Zeugen zu erhalten. Anschließende Fragen sollten zunächst so offen wie möglich sein und erst allmählich spezifischer werden (Trichtertechnik). Bei multiplen Sachverhalten ist die hierarchische Fragensequenz eventuell mehrfach zu durchlaufen. Wird einem Zeugen durch geschlossene Fragen nur die Möglichkeit gegeben, diese zu bejahen oder zu verneinen, können die oben beschriebenen inhaltlichen Qualitätsmerkmale nicht produziert werden. Befragungen ohne offene Erzählaufforderungen vermindern also die diagnostische Kraft der inhaltsorientierten Aussageanalyse bzw. können sie vollständig invalidieren. Durch inadäquate Explorationsstrategien wird damit die positive Feststellung des Realitätsgehalts von Aussagen erschwert, was zu einer Schwächung der Position von Opfer-Zeugen führen kann. Eine Reduzie-

rung des indiziellen Wertes der merkmalsorientierten Aussageanalyse ist auch dann gegeben, wenn in den gestellten Fragen Hinweise auf die Produktion solcher Beschreibungen enthalten sind, die als Glaubwürdigkeitsmerkmale gewertet werden. Daß sich in einer aussagepsychologischen Exploration zur Glaubhaftigkeitsbeurteilung die Anwendung von Techniken der (suggestiven) Aufdeckungsarbeit verbietet, wie sie zum Beispiel von Fürniss (1991) mit der Technik der „Geschichte von dem anderen Kind“⁹ vorgeschlagen wurde, versteht sich von selbst.

Für die anschließende Durchführung der merkmalsorientierten Inhaltsanalyse ist der Wortlaut der Exploration von Bedeutung; die Exploration zur Sache¹⁰ sollte von daher auf Ton- oder Videoband aufgezeichnet werden. Nur durch die genaue Dokumentation der in der Exploration verwendeten Berichtsanstöße und Fragen (bzw. an geeigneter Stelle auch Vorhalte) kann eine Abschätzung erfolgen, welche Aussagequalitäten bei den Schlußfolgerungen zur Glaubhaftigkeitseinschätzung verwertet werden können. Die Durchführung einer Aussageanalyse ohne Tonaufzeichnung – also ausschließlich aufgrund von Mit- oder Nachschriften – erscheint besonders bei komplexen Sachverhalten problematisch bis unmöglich. Die Notwendigkeit der Tonaufzeichnung ist nicht gleichzusetzen mit der Anfertigung eines wörtlichen Transkripts der aussagepsychologischen Exploration für die Analysetätigkeit des Sachverständigen bzw. mit der Beifügung dieses wörtlichen Transkripts zum aussagepsychologischen Gutachten. Dazu werden in Abschnitt 2.3 Ausführungen gemacht.

2.1.3 Leistungs- und Persönlichkeitsdiagnostik

Aus den voranstehenden Ausführungen über Logik und Methodik der aussagepsychologischen Begutachtung sowie über die notwendigen Erhebungsbereiche wurde bereits deutlich, daß persönlichkeitsbezogene Untersuchungen des Zeugen besonders im Hinblick auf die Erfassung aussagerelevanter kognitiver Fähigkeiten erfolgen. Neben Daten aus der biographischen Analyse (z. B. Schulerfolg) können hier auch standardisierte (psychometrische) psychologische Testverfahren zur Anwendung kommen. Nicht selten ist aus

⁹ Diese vermeintliche Aufdeckungstechnik beinhaltet, daß bei Verdacht auf sexuellen Kindesmißbrauch dem betreffenden Kind in explizit sexualisierter Sprache der sexuelle Mißbrauch eines anderen Kindes detailliert geschildert wird. Hintergrund ist, dem bis dahin schweigenden Kind durch die Vorgabe „zur Sprache zu verhelfen“. Der Anregungsgehalt für eine Falschaussage bzw. der suggestive Charakter einer solchen Methodik insbesondere bei sehr jungen Kindern wird von den Propagandisten der Aufdeckungsphilosophie nicht in Erwägung gezogen.

¹⁰ In der Praxis wird man auch die biographische Exploration aufzeichnen, schon weil Kinder häufig bereits in diesem Zusammenhang von sich aus Aussagen zur Sache tätigen und eine besondere Akzentuierung der Exploration zur Sache durch Einführung des Tonbandes erst zu diesem Zeitpunkt zu vermeiden ist.

Gründen der Ökonomie eine Auswahl von relevanten Untertests aus allgemeinen Leistungsprüfverfahren ausreichend. Sie können auch mit anderen Verfahren (z. B. sprachfreien Intelligenztests) kombiniert werden. In Übereinstimmung mit Greuel et al. (1998, S. 46) ist festzustellen, daß häufig für die Erhebung des individualtypischen Berichtsverhaltens eines Zeugen eher (unstandardisierte) Verhaltensproben in der aussagepsychologischen Begutachtung geeignet erscheinen als die Verwendung standardisierter Testverfahren.¹¹ Tests und Proben können sich aber auch ergänzen.

Zur Abklärung etwaiger Besonderheiten von Erlebens- bzw. Verhaltensdispositionen des zu begutachtenden Zeugen können neben biographischer Analyse und problembezogener Exploration auch standardisierte Fragebögen zur Anwendung kommen. Im diagnostischen Kontext der Begutachtung von Verhaltensauffälligkeiten bzw. Persönlichkeitsstörungen kann im Einzelfall auch der Einsatz projektiver Verfahren gerechtfertigt sein. Auch die gelegentlich zu beobachtende Verwendung des Stimulusmaterials von projektiven Verfahren zur Evozierung einer Erzähl- bzw. auch Phantasieprobe erscheint gerechtfertigt, wenn der Stellenwert solcher Proben im Kontext des hypothesengeleiteten Prüfprozesses deutlich wird.

Die Erhebung von Fremdanamnesen zur Persönlichkeitsbeurteilung – natürlich mit Einwilligung des Probanden – stellt in der psychologischen Praxis ein weiteres methodisches Element dar. Ob und inwieweit diese Methodik in der forensisch-aussagepsychologischen Begutachtung Anwendung finden kann bzw. sollte, wird im folgenden Abschnitt erörtert.

Die Ausdeutung von Kinderzeichnungen oder die Deutung der Interaktion von Kindern mit sogenannten anatomisch korrekten Puppen haben keinen Stellenwert in wissenschaftlich begründeten forensisch-aussagepsychologischen Gutachten (so auch Scholz & Endres, 1995). Das gleiche gilt für die Ausdeutung von allgemeinen Störungssymptomen, zum Beispiel psychosomatischen Erkrankungen, die als Folge sexueller Mißbrauchserfahrungen zwar möglich, aber prinzipiell hinsichtlich ihrer Verursachung unspezifisch sind (vgl. u.a. Kendall-Tackett, Williams & Finkelhor, 1993). Davon unberührt ist die Tatsache, daß Auswirkungsbeschreibungen des Miß-

¹¹ Allerdings sind dabei die Grenzen solcher Proben zu beachten. Eine Suggestibilitätsprobe reicht beispielsweise nicht aus, um die Möglichkeit einer durch Dritte induzierten Aussage zu prüfen. Bei Suggestibilität handelt es sich nicht um ein situationsübergreifendes, persönlichkeitspezifisches Konstrukt, sondern um ein multidimensionales Phänomen, bei dem nach heutigem Erkenntnisstand eine Reihe von kognitiven und sozialpsychologischen Mechanismen die Haupteinflußgrößen darstellen (vgl. Kap. 3). Je nach spezifischer Konstellation kann sich schon von daher dasselbe Kind in einer Situation als suggestionsresistent, in einer anderen Situation als suggestionsanfällig erweisen.

brauchserlebens durch den zu begutachtenden Zeugen – zum Beispiel der Folgen für das Erleben von körperlichen Berührungen oder auch der Schilderung von (Schul-)Leistungsbeeinträchtigungen u.a. – innerhalb der Inhaltsanalyse ein wichtiges Qualitätsmerkmal darstellen können, und zwar im Sinne des Merkmals „Delikttypizität“ oder einer weiten Fassung des Merkmals „Schilderung eigener psychischer Vorgänge“ und zuweilen des Merkmals „indirekt handlungsbezogene Schilderungen“.

Abschließend wird betont, daß in der psychologisch-diagnostischen Praxis multimethodal vorgegangen wird (vgl. u.a. Stieglitz & Baumann, 1994). Eine Einengung psychologischer Diagnostik auf standardisierte oder gar psychometrische Verfahren würde der Komplexität der Anforderungen an psychologische Diagnostik nicht gerecht werden.

2.2 Grenzen forensisch-aussagepsychologischer Befunderhebungen

Bei den nachfolgenden Überlegungen wird von dem Verständnis der Aufgaben und Befugnisse von gerichtlich¹² bestellten Sachverständigen ausgegangen. Der Sachverständige ist bekanntlich Helfer des Richters. Er soll das Allgemeinwissen und die Lebenserfahrung des Richters durch 1.) Mitteilung von Erfahrungssätzen seines Faches, 2.) Tatsachenfeststellung mit Hilfe seiner spezifischen Sachkunde bzw. 3.) Beurteilung von Tatsachen mit Hilfe seiner spezifischen Sachkunde ergänzen. Sachverständige sind also Beweismittel, sie haben sich der Würdigung und Bewertung des gesamten Beweisergebnisses eines Verfahrens zu enthalten. „Der Sachverständige muß sich auf die Beurteilung der Umstände beschränken, auf die sich sein Fachwissen bezieht. Trotz Beachtung der Beweisfragen verletzt er die ihm gezogenen Grenzen dann, wenn er das Gutachten nicht unter Beschränkung auf die Anwendung seines Fachwissens, sondern unter Mitverwertung allgemeiner sein Fachgebiet nicht berührender Umstände erstattet“ (Jessnitzer & Frieling, 1992, Rdnr. 11). Derselbe Grundsatz gilt auch für die Erhebung von Befundtatsachen in Abgrenzung von der Erhebung von Zusatztatsachen.

Die übliche Aufgabenstellung bei forensisch-aussagepsychologischen Glaubwürdigkeitsbegutachtungen besteht in einer Kombination von Tatsachenfeststellung und Tatsachenbeurteilung durch den Sachverständigen aufgrund seiner spezifischen Fachkenntnisse. Im folgenden wird auf solche Datenerhebungsbereiche und Datenerhebungsmethoden eingegangen, die bei der forensisch-aussagepsychologischen Begutachtung eine Rolle spielen und bei denen sich Fragen der möglichen Grenzüberschreitung von der sachver-

¹² Im Strafverfahren wird die Beauftragung durch die Staatsanwaltschaft als dem gerichtlichen Auftrag äquivalent angesehen (Jessnitzer & Frieling, 1992, Rdnr. 2 und 175 ff.).

ständigen Erhebung von Befundtatsachen zur nicht sachverständigen Erhebung von Zusatztatsachen ergeben.

Zur Exploration zur Sache:

In Übereinstimmung mit Greuel et al. (1998, S. 268) ist festzustellen, daß eine Exploration zur Sache nicht gleichzusetzen ist mit einer Vernehmung – weder in Zielsetzung noch in Methodik. Im Vordergrund der aussagepsychologischen Exploration zur Sache steht nicht die Rekonstruktion des in Frage stehenden Sachverhalts, sondern die Gewinnung von Indikatoren für die Einschätzung innerpsychischer Vorgänge bei dem Zeugen wie zum Beispiel Erlebnisbezug vs. gedankliche Konstruktion (Lüge) als Grundlage der Sachverhaltsdarstellung, d.h., die Exploration dient der Gewinnung von Material für die aussagepsychologische Qualitätsanalyse. Der prinzipielle Unterschied zwischen Vernehmung und Exploration wird durch partielle Gemeinsamkeiten¹³ nicht in Frage gestellt. Die Exploration zur Sache als wesentliche aussagepsychologische Methodik ist aus fachlicher Sicht unverzichtbar,¹⁴ ohne sie ist eine aussagepsychologische Begutachtung nicht denkbar.

Zur Persönlichkeitsdiagnostik:

Die meisten in diesem Begutachtungsbereich verwendeten Erhebungsmethoden (Tests, Fragebögen, Verhaltensproben, gegebenenfalls projektive Verfahren) stellen ein spezifisches psychologisches Methodenarsenal dar, insoweit bedarf es keiner besonderen Problematisierung. Zweifelsfälle könnten sich in bezug auf fremdanamnestiche Befragungen ergeben. Die biographische Rekonstruktion mit Hilfe von Fremdanamnesen stellt eine spezifische psychologische Methodik dar (und eben keine Vernehmung von Zeugen). Sie ist bei forensisch-psychologischen Begutachtungen insbesondere von Kindern unverzichtbar zur diagnostischen Einschätzung der Kompetenz und möglicher aussagerelevanter Besonderheiten der kindlichen Zeugen, da von Kindern in einer Eigenanamnese meistens keine ausreichenden Aufschlüsse erhalten werden können und auch die anderen Erhebungsmethoden möglicherweise im Einzelfall nicht ausreichen.

¹³ Aufgrund dieser Gemeinsamkeiten erscheinen Überlegungen zur Reduktion von Mehrfachbefragungen sinnvoll: So könnte bei kindlichen Zeugen zum Beispiel eine Verbindung von frühen Befragungen durch Ermittlungsinstanzen und forensisch-psychologischen Explorationen erfolgen. Dabei erscheint es angebracht, der aussagepsychologischen Explorationstechnik mit offener Befragung den (zeitlichen) Vorrang zu geben und spezifische ermittlungstechnische Fragen später zu stellen.

¹⁴ Diese Feststellung beinhaltet nicht, daß im Einzelfall eine forensisch-aussagepsychologische Gutachtenfrage der Art „Beurteilung von Tatsachen“ nicht auch verfahrensangemessen sein kann. Nach den eigenen Beobachtungen wird von dieser Art der Beauftragung von forensisch-psychologischen Sachverständigen von Gerichten möglicherweise zuwenig Gebrauch gemacht.

Die fremdanamnestische Befragung von Drittpersonen erscheint also aus aussagepsychologischer Sicht in vielen Fällen notwendig, mindestens bei Kindern im Vorschul- und Grundschulalter. Inwieweit bei älteren Kindern oder Jugendlichen eine Fremdanamnese erforderlich ist, richtet sich nach den Gegebenheiten des Einzelfalles und sollte somit der Entscheidung des Sachverständigen überlassen werden. Natürlich sind die Befragungen von Drittpersonen über aussagepsychologisch relevante Details der Biographie des Zeugen im Gutachten mitzuteilen. Inwieweit es der Einführung dieser biographischen Informationen durch Befragung der Drittpersonen als Zeugen in der Hauptverhandlung bedarf oder ob diese Informationen als Befundtatsachen des Sachverständigen von diesem selbst eingeführt werden können, ist keine fachpsychologische, sondern eine rechtliche Frage. Hier bestehen in der gerichtlichen Praxis durchaus Unsicherheiten mit Klärungsbedarf.

Zur Rekonstruktion der Aussageentwicklung:

Die bisherigen Überlegungen galten für den Erhebungsbereich „Persönlichkeit des Zeugen“ im Sinne der biographischen Analyse. Identische Probleme stellen sich aber auch für den Erhebungsbereich „Rekonstruktion von Aussageentstehung und weiterer Aussageentwicklung“. Hier kommt die Besonderheit hinzu, daß häufiger als bei der biographischen Analyse erstens mehrere Personen als Informationsträger in Frage kommen können und zweitens auch solche Personen bedeutsam sein können, die nicht in Verwandtschaft oder engen Beziehungsverhältnissen (z. B. Heimerzieher) zu den zu begutachtenden Zeugen stehen. Die mögliche Grenzüberschreitung von sachverständigen Befunderhebungen zu nicht sachverständigen Ermittlungshandlungen ergibt sich auch aus den zu erfragenden Sachverhalten: Zwar ist die Rekonstruktion der Aussageentwicklung Bestandteil aussagepsychologischer Methodik, doch sind die zu erfragenden Sachverhalte nicht ausschließlich psychologischer Natur. Zur adäquaten Befragung nach möglichen suggestiven Bedingungen bei Aussageentwicklungen wird allerdings in der Regel aussagepsychologische Kompetenz notwendig sein (vgl. Abschnitt 3.5).

In der eigenen aussagepsychologischen Begutachtungspraxis für Gerichte ist es bisher unbeanstandet geblieben, bei der Begutachtung kindlicher Zeugen außer einer fremdanamnestischen Befragung der vorstellenden Bezugsperson auch eine Befragung dieser Person zu Umständen der Erstbekundung und möglichen weiteren sachverhaltsbezogenen Befragungen des Kindes vorzunehmen – und diese im Gutachten zu berichten und für die eigenen Schlußfolgerungen zu verwenden.

Greuel et al. (1998, S. 269) halten auch die Befragung *mehrerer* Personen, die zur Rekonstruktion der Aussagegenese beitragen könnten, durch den Sachverständigen für angemessen. Sie betonen allerdings die Abhängigkeit solcher Entscheidungen von den Bedingungen des Einzelfalles und empfehlen die jeweilige Rückversicherung beim Auftraggeber im Einzelfall.

Nach hiesiger Auffassung ist allerdings der Regelung der Vorzug zu geben, daß der aussagepsychologische Sachverständige keine (auch keine informativ-rischen) Befragungen *aller* Drittpersonen vornimmt, die in der Aussageentwicklung eine Rolle gespielt haben können. Die spektakulären Massenmißbrauchsverfahren mit ihrem enormen Umfang seien exemplarisch als Beleg dafür angeführt, daß schnell Kapazitätsgrenzen des Sachverständigen überschritten sein können. Darüber hinaus ergeben sich rechtliche Probleme wie die fehlende Wahrheitspflicht der Befragten gegenüber dem Sachverständigen und die Notwendigkeit der Reflexion des Wahrheitsgehalts einzelner Befragungsergebnisse bei möglicherweise auftretenden Widersprüchen zwischen mehreren Befragten.

Die angesprochenen Probleme gelten im Grundsatz natürlich auch für die Befragung *einer* Bezugsperson durch den Sachverständigen. Dennoch sollte diese Möglichkeit in der Praxis schon aus Gründen der Verfahrensökonomie genutzt werden. Dem Sachverständigen sollten allerdings die rechtlichen Probleme bewußt sein, d. h., aussagepsychologische Schlußfolgerungen können nur unter dem Vorbehalt erfolgen, daß sich die erfragten Sachverhalte in einer Hauptverhandlung bestätigen, gegebenenfalls muß der Sachverständige alternative Schlußfolgerungen unter Zugrundelegung unterschiedlicher Anknüpfungstatsachen vortragen.

Wenn man sich die allgemeinen Regeln der Sachverständigentätigkeit vor Augen führt, so wird klar, daß es eine Verpflichtung eines aussagepsychologischen Sachverständigen zur Befragung von Drittpersonen im Umfeld der Aussageentstehung und ihrer weiteren Entwicklung weder aus fachlicher noch aus rechtlicher Sicht gibt. Der Sachverständige hat eben nicht „einen Fall zu lösen“, sondern seine spezifische Sachkunde zur Problemlösung einzubringen. Dies kann er ausreichend durch Hinweise an das Gericht auf weitere relevante Auskunftspersonen oder durch Alternativüberlegungen im Hinblick auf die Bedeutung denkbarer unterschiedlicher Anknüpfungstatsachen für die aussagepsychologische Bewertung der Aussagegenese.

Die Trennung zwischen sachverständiger Befunderhebung und nicht sachverständiger Ermittlungstätigkeit funktioniert in der Praxis problemlos, wenn Sachverständige strenge Maßstäbe bei der Definition ihrer fachspezifischen Erhebungsbereiche und -methoden anlegen. Eine höchstrichterliche Klärung erscheint dennoch für komplexe Fallkonstellationen hilfreich.

2.3 Das schriftliche (vorbereitende) Gutachten

Die Trivialität der Regeln zur Erstellung von schriftlichen Sachverständigengutachten (Gutachtentechnik) steht in auffallendem Kontrast dazu, wie häufig in der Praxis gegen sie verstoßen wird. Hier soll dennoch keine vollständige Anleitung zur Erstellung von forensisch-psychologischen Gutachten erfolgen, sondern es wird lediglich auf einige Aspekte eingegangen,

die in jüngster Zeit kontrovers diskutiert werden und die für die Beantwortung der Fragen des BGH bedeutsam sind.

Formalia:

Natürlich muß ein Gutachten einleitende Formalia enthalten: Auftrag, Ort und Zeitpunkt der Begutachtung, verwendete Methoden und weitere durchführungsrelevante Angaben. Aus der Mitteilung dieser Formalia sollte sich auch der zeitliche Umfang der Begutachtung bzw. einzelner Teilschritte ergeben, sofern diese nicht durch den Befundbericht deutlich werden. Aus fachlicher Sicht geboten ist ein exaktes Verlaufsprotokoll der Begutachtungen, das allerdings nicht regelmäßig in schriftlichen Gutachten mitgeteilt werden muß. Dies ermöglicht die Beurteilung sequentieller Effekte wie möglicher Belastungs- bzw. Ermüdungserscheinungen, gelegentlich kann es auch für die inhaltliche Interpretation von Befunden Bedeutung haben. Der Umfang des einleitenden Teils (der hier so genannten Formalia) richtet sich nach den Gegebenheiten des Einzelfalls, starre Vorschriften erscheinen weder möglich noch notwendig.

Befundbericht und Schlußfolgerungen:

Bereits aufgrund logischer Erwägungen ergibt sich für Inhalt und Aufbau von Gutachten, daß die Anknüpfungs-¹⁵ und Befundtatsachen mitgeteilt werden, von denen der Sachverständige ausgeht (Berichtsteil), und daß die auf die Gutachtenfrage(n) angestellten Schlußfolgerungen sowie ihr Ergebnis mitgeteilt werden (schlußfolgernder Teil). Dieses Grundkonzept von Gutachten wird auch in den einführenden Lehrbüchern über psychologische Diagnostik dargestellt. Offenbar wegen häufig zu beobachtender Mängel hat der Berufsverband Deutscher Psychologen (1988) diese Grundregeln auch in einer Broschüre zusammengestellt.

Der Aufbau von Gutachten läßt sich auch aus der Gehilfenrolle des Sachverständigen bei der gerichtlichen Problemlösung (Beantwortung einer Beweisfrage) ableiten. Eine ausschließliche Beantwortung der Beweisfrage(n) reicht nicht aus, vielmehr muß das Gutachten nachvollziehbar (und damit überprüfbar) aufgebaut sein. Das Transparenzgebot beinhaltet die Benennung und Beschreibung aller Anknüpfungs- und Befundtatsachen sowie die Darlegung der schlußfolgernden Überlegungen in einer Weise, die den Abwägungsprozeß (Denkprozeß) des Gutachters erkennbar werden läßt. Bei umfangreichen Ausarbeitungen ergibt sich damit die Notwendigkeit für klare Untergliederungen und gegebenenfalls eine Zusammenfassung.

¹⁵ Die Wiedergabe der Anknüpfungstatsachen impliziert keinen umfangreichen Aktenbericht, es genügt die Nennung der berücksichtigten Akten gegebenenfalls mit Hervorhebung relevanter Vernehmungen. Da Gutachten in sich verständlich sein sollten, empfiehlt sich allerdings auch eine kurze inhaltliche Beschreibung des (Ausgangs-)Sachverhaltes, dessen Umfang in der Regel nur wenige Seiten umfaßt.

Aus dem Transparenzgebot ergibt sich also in gewisser Weise ein regelhafter Aufbau von schriftlichen Gutachten, auch ohne daß etwa gesetzliche Vorschriften dazu existieren (Jessnitzer & Frieling, 1992, Rdnr. 430) oder fachliche Regeln ein starres Schema vorschreiben würden. Das Transparenzgebot regelt aber nicht den Umfang der schriftlichen Ausarbeitungen. Es gilt prinzipiell für Kurzgutachten ebenso wie für ausführliche Ausarbeitungen und kann unabhängig von der Gutachtenlänge erfüllt werden.

In der forensischen Praxis kommt es vor, daß einige Auftraggeber – unter anderem mit dem Hinweis auf die Vorläufigkeit eines schriftlichen Gutachtens im Strafprozeß – ausdrücklich kompakte (kurze) Gutachten, andere dagegen – häufig unter Hinweis auf die Nützlichkeit eines ausführlichen Explorationsberichts für die weiteren forensischen Entscheidungen – eher ausführliche schriftliche Gutachten anfordern. In der Mehrzahl der Fälle erfolgt der Auftrag aber ohne jeglichen Hinweis zum erwarteten Umfang des schriftlichen Gutachtens. Hier könnten Auftraggeber ihre Leitungsfunktion durchaus prägnant wahrnehmen, d.h. ihre Anforderungen deutlich kommunizieren.

Aus den dargestellten Überlegungen ergibt sich unseres Erachtens, daß eine regelhafte Forderung nach einem vollständigen Transkript der Exploration zur Sache in aussagepsychologischen Gutachten bereits aus pragmatischen Gründen nicht sinnvoll erscheint. Nicht zuletzt sind im Strafprozeß Aufträge denkbar, die eine Untersuchung (d.h. auch Exploration) eines Zeugen beinhalten, aber auf ein schriftliches (vorläufiges) Gutachten verzichten. Natürlich ist die transparente Darstellung der Explorationsinhalte ein wesentlicher Teil von (schriftlichen) aussagepsychologischen Gutachten, aber auch inhaltliche Überlegungen sprechen nicht notwendigerweise für ein vollständiges Wortprotokoll der Exploration zur Sache. Die erforderliche Transparenz kann durchaus durch einen mehr oder weniger zusammenfassenden Explorationsbericht gegebenenfalls mit zahlreichen Zitaten gewährleistet werden. Bedenkt man die spezifische Explorationsmethode bei aussagepsychologischen Begutachtungen („Trichtertechnik“) mit dem Schwerpunkt auf offenen Erzählanstößen, so ist – insbesondere bei Infragestellen multipler anklagerelevanter Vorfälle – bei vielen Zeugen mit relativ unstrukturierten Darstellungen zu rechnen. Der (sachkundige) Gutachter wird auch bei assoziativen Übergängen eines Zeugen zwischen verschiedenen Sachverhalten weder unterbrechen noch steuernd eingreifen, da gerade solche Darstellungsweisen aussagepsychologisch von Interesse sind. In solchen Konstellationen kann es zu Wortprotokollen von erheblichem Umfang (auch über 100 Seiten) kommen, die ohne Strukturierungshilfen nur schwer lesbar sind.

Ein anderer Gesichtspunkt ist der einer Kontrolle des Sachverständigen bei Zweifeln an seiner Kompetenz oder Integrität. Da es aus fachlicher Sicht unerläßlich erscheint, daß ein aussagepsychologischer Sachverständiger für die

eigene Analysetätigkeit eine Tonaufnahme zur Verfügung hat (vgl. Abschnitt 2.2), ist das Material zur Kontrolle des Sachverständigen prinzipiell vorhanden, denn Sachverständige sollten ihre Befundunterlagen bis zur Rechtskraft eines Urteils aufbewahren. Ob ein Sachverständiger diese Unterlagen aus der Phase der Befunderhebung auf Anforderung eines Gerichts vorzulegen hat, ist ausschließlich eine rechtliche Frage, die sich analog auch im Hinblick auf Testunterlagen und andere Originalaufzeichnungen während der Befunderhebung stellt.

Die unseres Erachtens nicht praxisgerechte Forderung nach regelmäßigen vollständigen Wortprotokollen der Exploration als verbindlichem Teil von forensisch-aussagepsychologischen Gutachten entstand wahrscheinlich als Reaktion auf einen unzulänglichen Umgang mit dem Verdacht auf sexuellen Kindesmißbrauch (Stichwort: suggestive Aufdeckungsarbeit) und auf damit verbundene fehlerhafte schriftliche Stellungnahmen, durchaus auch in Form von gerichtlichen Glaubwürdigkeitsgutachten (vgl. u.a. Steller, 1998). Angesichts der Komplexität suggestiver Prozesse kann ein Wortprotokoll einer forensisch-aussagepsychologischen Exploration allerdings in vielen Fällen gar nicht zur Aufklärung der suggestiven Faktoren beitragen. Ist der Befragung ein längerer suggestiver Prozeß vorausgegangen, werden suggerierte Inhalte unter Umständen bereits im spontanen Bericht produziert. Die Suggestionodynamik wird zu diesem Zeitpunkt in der Regel durch spezifische Verstärkungen und einseitige Informationserhebung aufrechterhalten (vgl. Abschnitt 3.5), suggestiv formulierte Fragen spielen allenfalls eine untergeordnete Rolle. Der wesentliche Analyseschritt für die Problemstellung „Suggestion oder Realitätsgehalt“ besteht von daher in der Rekonstruktion der gesamten Aussagegenese. Auf der anderen Seite führt nicht jede geschlossene oder gar fehlerhaft suggestive Frage zu Falschaussagen – auch nicht bei jungen Kindern. Die Befürchtung, daß aussagepsychologische Sachverständige durch ihre Exploration erst das Produkt schaffen würden, das sie begutachten (sollen), erscheint nicht realistisch.

Geht man also von dem sachkundigen Gutachter aus, der auftragsgemäß nach bestem Wissen und Gewissen seine aussagepsychologische Begutachtung durchführt und darüber ein Gutachten erstellt, so erscheint der nach fachlichen Gesichtspunkten gestaltete Explorationsbericht gegenüber einem Wortprotokoll in bezug auf das Gebot der Nachvollziehbarkeit in vielen Fällen eventuell sogar überlegen. Denn Transparenz kann dadurch erhöht werden, daß der Sachverständige in seinem Explorationsbericht deutlich kennzeichnet, an welchen Stellen direkte Fragen, Vorhalte oder andere Vorgaben bzw. an welchen Stellen spontane Schilderungen erfolgten.

Die sachgerechte Gestaltung des Explorationsberichts zur Erfüllung des Transparenzgebotes ist bereits eine sachverständige Tätigkeit. Die Gestaltung des Explorationsberichts mit allen denkbaren Varianten zwischen vollständi-

gem Wortprotokoll und zusammenfassendem Bericht mit mehr oder weniger ausführlichen einschlägigen Zitaten von Frage und Antwort ist abhängig von der Gestaltung des gesamten schriftlichen Gutachtens und obliegt daher der Entscheidung des Sachverständigen; es sei denn, der Auftraggeber hat durch explizite Aufträge Vorgaben gemacht.

Das praxeologische Problem der Erstellung von Wortprotokollen bzw. der Herausgabe von Tonaufnahmen durch Sachverständige sollte auch in Verbindung mit Überlegungen zum Zeugenschutz gesehen werden. Die vorstehenden Erörterungen bezogen sich auf die gängige Praxis, daß aussagepsychologische Explorationen in der Regel nach einer polizeilichen Vernehmung und gegebenenfalls nach einer weiteren staatsanwaltschaftlichen oder richterlichen Vernehmung erfolgen und daß der Zeuge trotz Audioaufnahme beim Sachverständigen in der Hauptverhandlung aussagen muß. Im Zeugenschutzgesetz vom 30. April 1998 wird u.a. die Aufzeichnung von Zeugenvernehmungen auf Bild-Ton-Träger bei Opfer-Zeugen unter 16 Jahren vorgesehen (§ 58 a StPO). Ein Ersatz der Vernehmung von Zeugen unter 16 Jahren in der Hauptverhandlung ist durch Vorführung der Bild-Ton-Aufzeichnung einer *richterlichen* Vernehmung möglich (§ 255 a StPO). Für die zukünftige Praxis sollte überlegt werden, ob und wie aussagepsychologische Explorationen durch Sachverständige in diese Bild-Ton-Dokumentation einbezogen werden können. Dadurch könnte sich das Problem von Wortprotokollen aussagepsychologischer Explorationen erledigen.

Testdarstellungen und Testergebnisse:

Eine ausführliche Beschreibung der Indikation und Methodik von standardisierten diagnostischen Verfahren (Leistungstests, Fragebögen) in schriftlichen Gutachten erscheint grundsätzlich entbehrlich. Die Nennung der Verfahren reicht aus, da damit Zielsetzung, Reichweite, Methodik und Gütekriterien der Verfahren grundsätzlich bekannt sind: Sie sind in den Testmanualen und umfangreicher Sekundärliteratur dargestellt. Regelmäßige Testbeschreibungen in aussagepsychologischen Gutachten würden diese im Umfang unnötig „aufblähen“ und damit eher gegen als für das Transparenzgebot wirken. Eine fallbezogene Verdeutlichung des Stellenwertes von verwendeten Testverfahren kann ohne ausführliche Testbeschreibungen erfolgen (z. B. durch einfache Hinweise, wie: „Zur Einschätzung des allgemeinen kognitiven Leistungsniveaus des Zeugen wurde der sprachfreie Intelligenztest XY durchgeführt. ...“). Eine (im Umfang dem Gesamtgutachten entsprechende) Darstellung, falls und in welcher Weise von einer standardisierten Testdurchführung abgewichen wurde, erscheint allerdings geboten. Ebenso sind angewandte Proben in angemessenem Umfang zu erläutern, so daß ihr Stellenwert im Rahmen der hypothesengeleiteten Befunderhebung deutlich wird.

Für psychometrische Verfahren (Leistungstests, Fragebögen) ist noch die Frage zu diskutieren, ob Testergebnisse in den testspezifischen Skalenwerten

(z.B. T-Werte, Z-Werte oder IQ-Werte) oder gar in sogenannten Rohwerten mitgeteilt werden sollten oder ob eher eine Verbalisierung der Ergebnisse angezeigt erscheint. Letzteres wird hier favorisiert, da es wenig sinnvoll erscheint, testspezifische Skalenwerte mitzuteilen. Zwar verbinden die meisten psychologischen Laien (und das sind die Adressaten von Gutachten) eine gewisse Vorstellung mit der IQ-Skalierung, doch fragt sich, ob diese wirklich zutrifft. Einem psychologischen Laien wird kaum bekannt sein, daß eine Leistungseinordnung auf der IQ-Skala zum Beispiel in Höhe von 115 völlig identisch ist mit dem Standardwert (Z-Wert) von 110 (z. B. bei Verwendung des Intelligenz-Struktur-Tests [IST 2000] von Amthauer, Brocke & Liepmann, 1998) bzw. mit dem Standardwert von 13 (sogenannte Wertpunkte bei den Wechsler-Subtests, vgl. Tewes, 1994) und dem Wert von 60 auf der T-Skala.

Hier soll nicht auf Einzelheiten der Quantifizierung bei psychometrischen Verfahren eingegangen werden. Der Grundgedanke testpsychologischen Quantifizierens besteht in der Angabe der individuellen Position im Vergleich zu einer Bezugsgruppe (meistens bestehend aus altersgleichen, es sind aber andere Bezugsgruppen, z. B. Absolventen desselben Schultyps u.a. denkbar), d.h. in der Position des individuellen Werts im Vergleich zum mittleren Wert und der Werteverteilung (Streuung) in der gewählten Bezugsgruppe. Dieser Grundgedanke erscheint unter dem Transparenzgebot von Gutachten eher vermittelbar durch Verbalisierungen als durch Angaben von Zahlenwerten auf prinzipiell austauschbaren Skalen. Verbalisierungen sollten die Relativität der Einordnungen deutlich machen, zum Beispiel: Die Leistungen des Zeugen bei allgemeinen Denkaufgaben lagen im Vergleich zu Gleichaltrigen über dem Durchschnitt, oder: Die Selbstbeschreibung des Zeugen im Hinblick auf psychosomatische Beschwerden lag im Vergleich zu klinisch auffälligen Personen im Durchschnitt, im Vergleich zu klinisch unauffälligen Personen dagegen über dem Durchschnitt. Dem sachverständigen Kollegen erschließt sich durch solche Formulierungen der Skalenbereich, in dem das individuelle Testergebnis liegt. Der Informationsgehalt für den psychologischen Laien dürfte eher in den Verbalisierungen als in Zahlenwerten gegeben sein.

Es wird nicht übersehen, daß Transparenz im Sinne von Kontrolle bzw. Überprüfbarkeit der Richtigkeit von Testergebnissen besser durch die Mitteilung von Zahlenwerten gewährleistet werden kann. Letztlich ist ja auch eine Kombination von Zahlenwerten und Verbalisierungen im Gutachten denkbar. Dies müßte sich nach dem Gesamtumfang des Gutachtens richten. Im folgenden soll gezeigt werden, daß eine Überprüfbarkeit durch Angabe von Zahlenwerten aber nur in sehr engen Grenzen gegeben ist. So ist zum Beispiel bei Mitteilung von individuellen Testergebnissen in testspezifischen Skalenwerten nur die Richtigkeit des (ausgesprochen trivialen, da in den Testmanualen vorgegebenen) Schrittes von dieser Zahlenangabe zur Verbalisi-

sierung überprüfbar. Vorgeschaltet sind die Umwandlung von Rohwerten in Standardwerte anhand von Normentabellen bzw. zunächst die Erhebung der Rohwerte. Die Mitteilung von individuellen Roh- und Standardwerten würde die Prüfung ermöglichen, ob der Sachverständige möglicherweise Fehler beim Ablesen von Tabellen gemacht hat – eine im Kontext aussagepsychologischer Begutachtungen wohl eher marginale Fehlerquelle. Wollte man überprüfen, ob die Rohwerte adäquat erhoben wurden, wären die Testprotokolle als solche zu prüfen – bei vielen standardisierten Verfahren entzieht sich die Erhebungsphase allerdings der Kontrolle, da sie sich auch anhand der Testprotokolle nicht vollständig rekonstruieren läßt.

Mit dem obenstehenden Quantifizierungsbeispiel sollte darauf hingewiesen werden, daß die Verwendung einer spezifischen Fachsprache wahrscheinlich gerade nicht zur Erhöhung gutachterlicher Transparenz beiträgt. Forderungen in der Literatur gehen eher in die Richtung, daß Sachverständige ihre Befunde allgemeinverständlich darstellen (z. B. Jessnitzer & Frieling, 1992, Rdnr. 448). Für die prinzipielle Nachprüfbarkeit der Tätigkeit von Sachverständigen erscheint es eher sinnvoll, daß diese ihre Befunderhebungen in ihren Unterlagen möglichst genau dokumentieren, als daß schriftliche Gutachten durch Scheingenaugkeiten überfrachtet werden.

Schriftliche Gutachten erfolgen in Strafverfahren zur Vorbereitung von Hauptverhandlungen.¹⁶ Da immer damit zu rechnen ist, daß in einer Hauptverhandlung neue Tatsachen bekannt werden, wird noch einmal die Vorläufigkeit der schriftlichen (vorbereitenden) aussagepsychologischen Gutachten deutlich. Jessnitzer und Frieling (1992, Rdnr. 499) empfehlen, daß sich der Sachverständige im schriftlichen Gutachten „noch nicht allzusehr festlegt, wenn erst in der Hauptverhandlung die endgültige Klärung des zu beurteilenden Sachverhalts zu erwarten ist“. In der Praxis kommt es immer wieder vor, daß Auftraggeber derartige Gutachten als „uneindeutig“ kritisieren und auf mangelnde Sachkunde oder mangelnde Entscheidungsfreudigkeit des Sachverständigen attribuieren. Gutachten, die rechtlichen und fachlichen Standards genügen, beinhalten aber gerade eine Darstellung des Abwägens des Sachverständigen zwischen verschiedenen Denkmöglichkeiten. Natürlich sind Gewichtungen nötig und möglich. Sie haben aber immer mit dem Vorbehalt zu erfolgen, daß in der Hauptverhandlung keine Tatsachen bekannt werden, die eine andere Gewichtung angezeigt sein lassen. Insofern stellen entsprechende Hinweise am Schluß von schriftlichen Gutachten keine Leerformeln dar.

¹⁶ In Strafverfahren wird ein Sachverständiger, der zuvor von der Staatsanwaltschaft oder dem Gericht mit einem Gutachten beauftragt worden war, grundsätzlich vernommen – sofern ein Hauptverfahren eröffnet wurde –, von wenigen Ausnahmen abgesehen, die in der Regel für aussagepsychologische Gutachten nicht zutreffen (Jessnitzer & Frieling, 1992, Rdnr. 417).

Aus der Funktion des mündlichen Gutachtenvortrages von Sachverständigen in Hauptverhandlungen ergibt sich, daß einengende Vorschriften in bezug auf vorbereitende schriftliche Gutachten den verschiedenen Anforderungen der forensischen Gutachtenpraxis nicht gerecht werden. Der Gutachtenvortrag in der Hauptverhandlung ermöglicht weitere Darlegungen durch den Sachverständigen und gibt in Zweifelsfällen die Möglichkeit zur Kontrolle seiner Tätigkeit.

Abschließend soll noch eine Überlegung angestellt werden, die eher prozeßrechtliche Fragen berührt. Analysen von schriftlichen aussagepsychologischen Gutachten, die nach einer Hauptverhandlung bzw. nach einem Urteil angefertigt und vorgelegt werden (z. B. im Rahmen von Berufungs- oder Revisionsverfahren – nach eigenen Beobachtungen auch in Wiederaufnahmeverfahren), berücksichtigen unseres Erachtens nicht ausreichend die Tatsache, daß das mündliche („eigentliche“) Gutachten, das in der Hauptverhandlung erstattet wurde, möglicherweise trotz identischer Gesamtbewertung eine vom schriftlichen Vorgutachten abweichende Argumentationsfigur gehabt haben kann – eben unter Berücksichtigung der Ergebnisse der Hauptverhandlung. Da diesbezügliche aussagekräftige Protokolle¹⁷ fehlen, müßte der Gutachtenkritiker sich eigentlich an den Feststellungen über das Gutachten im Urteil orientieren. Daß dieses wiederum die Perzeption des Gutachtens durch das Gericht wiedergibt, wäre dabei zu berücksichtigen. Der von Deckers (1996, S. 3109) aus der Sicht der Verteidigung betonte Weg der Auseinandersetzung mit schriftlichen aussagepsychologischen Gutachten durch Inanspruchnahme eines Gutachtenkritikers betrifft also im wesentlichen die Vorbereitung erstinstanzlicher Hauptverhandlungen, während spätere Kritik an vorbereitenden Gutachten häufig ins Leere gehen dürfte.

2.4 Standards forensisch-aussagepsychologischer Begutachtungen

Allgemeine Standards (Gütemaßstäbe) der forensisch-psychologischen Diagnostik lassen sich unter vier Aspekten betrachten (Steller, 1988a): Neben rechtlichen und ethischen Standards geht es um theoretische Gesichtspunkte des diagnostischen Entscheidungsprozesses, um die inhaltlichen Grundlagen für die Bearbeitung spezieller Gutachtenfragen und um die Regeln für die Erstellung von Gutachten (Gutachtentechnik).

Auf ethisch-moralische oder rechtliche Standards der forensisch-psychologischen Begutachtung wird hier nicht näher eingegangen. Zu diskutieren wären Zielkonflikte zwischen der rein diagnostischen Aufgabenstellung und dem beruflichen Selbstverständnis auch vieler Psychologen (nicht nur von Ärzten) als Behandelnde, als Therapeuten. Es ginge auch um die Grenzen forensisch-psychologischer Begutachtung unter den mög-

¹⁷ Ausnahmen sind Hauptverhandlungen bei Amtsgerichten, bei denen aufgrund der Komplexität des mündlichen Gutachtenvortrags aber nicht selten entstellende Protokollierungen vorkommen.

licherweise konkurrierenden Gesichtspunkten der individuellen Rechte des zu Begutachtenden und den Ansprüchen der Rechtsgemeinschaft. Nicht alles, was methodisch möglich ist, muß gesetzlich erlaubt sein. Andererseits können rechtliche und ethisch-moralische Bewertungen sich ändern, wie u.a. die Beurteilung der sogenannten Polygraphie durch den Bundesgerichtshof in seinen unterschiedlichen Einschätzungen von 1998 gegenüber 1954 zeigt.

Der theoretische Aspekt wurde in Abschnitt 1.3 diskutiert. Hier geht es um das allgemeine Verständnis von psychologischer Diagnostik als hypothesengeleitete Problemlösemethode in Abgrenzung zu einem überholten Verständnis von psychologischer Diagnostik als Methode der deutenden (aufdeckenden) Menschenbeurteilung im Sinne einer charakterologischen Röntgendiagnostik (vgl. auch Steller, in Vorbereitung). Dieses allgemeine Verständnis von psychologischer Diagnostik korrespondiert mit der fachlichen Absage an eine charakterologische Glaubwürdigkeitsbegutachtung von Zeugen und mit der ausschließlichen Orientierung an aussagebezogenen Glaubhaftigkeitsbeurteilungen (vgl. Kapitel 1.).

Natürlich gehören inhaltliche Kenntnisse eines Gegenstandsbereichs zu Aspekten der Qualitätssicherung. Das notwendige inhaltliche Wissen zur Glaubhaftigkeitsbeurteilung ist Gegenstand dieses Gutachtens. Ergänzend wird betont: Während vor einigen Jahren für die aussagepsychologische Glaubhaftigkeitsbeurteilung bei Anlegung strenger Maßstäbe noch festzustellen war, daß wissenschaftliche Gütekriterien nicht erfüllt waren, da eine umfassende Theorie und folglich eine aus der Theorie abgeleitete Erhebungstechnologie fehlte (Steller, 1988b), läßt sich zehn Jahre später konstatieren, daß der damals prognostizierte Forschungsschub gerade auf diesem Gebiet der forensischen Psychologie tatsächlich eingetreten ist. Auch wenn nach wie vor ein Theoriedefizit besteht, so hat es doch erheblichen Wissenszuwachs gegeben. Wie in Kapitel 3. dargestellt werden wird, wurde die in der Praxis bereits angewandte, empirisch aber kaum geprüfte Methodik zur Differenzierung zwischen wahren und erfundenen Aussagen im letzten Jahrzehnt vielfachen empirischen Prüfungen unterzogen, bei denen die Brauchbarkeit ihrer Grundidee insgesamt bestätigt wurde. Zeitlich parallel wurde in internationaler Forschung eine bis dahin in der Praxis marginale Fragestellung umfangreich untersucht, nämlich die Bedingungen für das Entstehen suggerierter Aussagen. Hier hat es einen regelrechten Forschungsboom gegeben (vgl. Abschnitt 3.5). Es ist umfangreiches Wissen über die Entstehung und die Randbedingungen von Suggestionseffekten gesammelt worden, das unmittelbare Bedeutung für praktische aussagepsychologische Begutachtungen besitzt.

Auch die Ebene der konkreten Gutachtenerstattung (Gutachtentechnik) ist Gegenstand der hier vorliegenden Ausarbeitung (vgl. Abschnitt 2.3).

Zusammenfassend gelten die folgenden wesentlichen Standards (nach Volbert, im Druck):

- **Bezugnahme auf die spezifische Aussage:**
Die gutachterliche Stellungnahme hat sich auf die spezifische Aussage zu beziehen. Es ist zu prüfen, ob aussagepsychologische Indikatoren dafür vorliegen, daß es sich um eine erlebnisbasierte Schilderung handelt. Dabei ist keine globale, sondern eine auf die anklagerelevanten Vorfälle bezogene Einschätzung vorzunehmen. Charakterologische oder motivationsbezogene Überlegungen reichen nicht aus.
- **Erkennbare Spezifizierung der globalen gerichtlichen Fragestellung für den Einzelfall, d.h. Formulierung von relevanten Fragestellungen und Hypothesen:**
Diese Forderung impliziert, daß überhaupt Alternativhypothesen zur Wahrheitsannahme, d.h. zur Hypothese, daß es sich um eine erlebnisbasierte Darstellung handelt, aufgestellt werden müssen.
- **Datensammlung auf der Basis der ausgewählten Fragestellungen:**
Die Datenerhebung (also sowohl Explorationsinhalte wie Testverfahren und andere diagnostische Erhebungsmethoden) muß sich orientieren an den Fragestellungen und kann nicht routinemäßig festgelegt werden (hypothesengeleitete Diagnostik).
- **Erkennbare Überprüfung relevanter Alternativhypothesen:**
Es muß aus dem Gutachten ersichtlich sein, ob und wie relevante Alternativhypothesen zur Wahrheitsannahme geprüft wurden. Dabei reicht eine einfache Benennung nicht aus, entscheidend ist, daß der Abwägungsprozeß des Gutachters, sein diagnostisches Schlußfolgern, deutlich wird.
- **Verwendung einer wissenschaftlich begründeten Methodik:**
Der wesentliche methodische Schritt besteht in einer Analyse der Aussagequalität. Die Qualitätsanalyse umfaßt im Schwerpunkt eine merkmalsorientierte Inhaltsanalyse, besteht aber ebenfalls aus der Konstanzanalyse und der Analyse der Aussageweise. Die Tatsache, daß eine merkmalsorientierte Inhaltsanalyse vorgenommen wurde, bedeutet noch nicht, daß ein Gutachten dem wissenschaftlichen Stand entspricht. Wenn die Inhaltsanalyse ohne ausreichende Berücksichtigung der individuellen Fähigkeiten, vor allem aber ohne ausreichende Beachtung der Aussagegenese erfolgt, kann das Ergebnis ebenso fehlerhaft sein, als wenn gar keine inhaltsanalytische Methodik angewandt wurde.

3 Wissenschaftliche Grundlagen der Aussageanalyse

Trotz ständiger Hinzuziehung von Sachverständigengutachten mit inhaltsanalytischer Methodik durch Gerichte existierten über Jahrzehnte kaum em-

pirische Untersuchungen hierzu. Als die Methodik der Aussageanalyse in größerem Rahmen im internationalen Bereich vorgestellt wurde (Statement Reality Analysis; Undeutsch, 1982), kritisierten insbesondere amerikanische Kollegen die Anwendung eines empirisch nicht überprüften Verfahrens in der forensischen Praxis (Wells & Loftus, 1991). Von den Begründern der Methodik wurde lange Zeit bezweifelt, daß eine experimentelle Überprüfung der Glaubwürdigkeitsmerkmale überhaupt möglich sei (z.B. Arntzen, 1983, 1993) – ein Standpunkt, der in dieser absoluten Form nicht haltbar ist (Wegener, 1997). In den letzten Jahren wurde eine Reihe von Studien durchgeführt, um zu überprüfen, ob sich der angenommene qualitative Unterschied zwischen erlebnisbegründeten und erfundenen Schilderungen empirisch nachweisen läßt. Zu unterscheiden sind dabei Feld- und Simulationsstudien; Feldstudien beziehen sich auf tatsächliche forensische Fälle, in Simulationsstudien werden Versuchspersonen aufgefordert, Berichte über wahre und erfundene Schilderungen mit einer anderen inhaltlichen Thematik abzugeben. Beide Herangehensweisen haben Nachteile: Bei Feldstudien fehlt in der Regel ein sicheres Außenkriterien zur Bestimmung des Wahrheitsstatus einer Aussage; Simulationsstudien weisen als Schwäche eine geringe Lebensnähe auf, wodurch die Übertragbarkeit ihrer Ergebnisse auf die Realsituation von forensischen Begutachtungen eingeschränkt sein kann. Da diese Probleme grundsätzlich nicht auflösbar sind, ist es notwendig, beide Forschungszugänge ergänzend zu nutzen (Bekerian & Dennett, 1995).

3.1 Feldstudien

Erste Feldstudien mit forensischem Fallmaterial wurden in den USA durchgeführt (Boychuk, 1991; Raskin & Esplin, 1991a, 1991b). Als Außenkriterium für den Wahrheitsstatus der Aussage wurde eine Kombination von Merkmalen benutzt: Vorhandensein bzw. Fehlen medizinischer Befunde mit Hinweischarakter auf sexuellen Mißbrauch, Ergebnisse einer Polygraph-Untersuchung des Beschuldigten, Geständnisse und Aussagen anderer Zeugen; in der Arbeit von Boychuk wurde außerdem die strafrechtliche Sanktionierung bzw. das Fehlen einer solchen berücksichtigt. Analysiert wurden die Aussagen von 40 Kindern im Alter von 3 bis 15 Jahren (Raskin & Esplin, 1991a, 1991b) bzw. 75 Aussagen von Kindern im Alter zwischen 4 und 16 Jahren (Boychuk, 1991). Beide Untersuchungen erbrachten, daß in „bestätigten“ Kinderaussagen über sexuelle Mißbrauchserfahrungen deutlich mehr Qualitätsmerkmale enthalten waren als in „zweifelhaften“ Aussagen. In der Arbeit von Raskin und Esplin (1991a) war jedes einzelne Merkmal häufiger in „bestätigten“ als in „zweifelhaften“ Aussagen zu finden. In der Feldstudie von Boychuk (1991) wurden 12 Merkmale deutlich häufiger in „bestätigten“ als in „äußerst zweifelhaften Fällen“ kodiert, bei der Mehrzahl der motivationsbezogenen Merkmale sowie bei einigen anderen Merkmalen („Phänomengemäße Schilderung unverstandener Handlungselemente“ und „Schilderung psychischer Vorgänge des Angeschuldigten“) fanden sich keine Mittelwertunterschiede zwischen den beiden Gruppen. Hierbei ist aber zu

berücksichtigen, daß diese Merkmale ohnehin nur in wenigen Fällen nachweisbar waren; bei zwei anderen Merkmalen („Schilderung nebensächlicher Einzelheiten“ und „Entlastung des Angeschuldigten“) fanden sich zwar hypothesenkonforme Mittelwertunterschiede, diese verfehlten jedoch knapp das statistische Signifikanzniveau von 5 %.

Um Fehlklassifizierungen zu vermeiden, wurden in diese beiden Untersuchungen nur Fälle einbezogen, die mit sehr hoher Wahrscheinlichkeit erlebnisbegründet bzw. mit sehr hoher Wahrscheinlichkeit nicht erlebnisbasiert waren. Von daher ist von einer Einschränkung der Repräsentativität der Fälle auszugehen, und es stellt sich die Frage, ob sich qualitative Unterschiede auch in den Aussagen des „Mittelfelds“ finden lassen. Bei einer kürzlich in Israel durchgeführten Feldstudie (Lamb, Sternberg, Esplin, Hershkowitz, Orbach & Hovav, 1997) wurden 98 Transkripte über sexuelle Mißbrauchserfahrungen von Kindern zwischen 4 und 13 Jahren analysiert. Die Fälle wurden ohne Berücksichtigung der Aussagen der Kinder anhand verschiedener Dimensionen (medizinische Befunde, andere Zeugenaussagen, Sachbeweise, Aussage des Beschuldigten, Verschiedenes) auf einer fünfstufigen Skala hinsichtlich der Wahrscheinlichkeit des Zutreffens der Vorwürfe von Ratern eingeschätzt, denen die Aussagen der Kinder nicht bekannt waren.¹⁸ Die Aussagen in den als „sehr wahrscheinlich“ eingeschätzten Fällen wiesen einen deutlich höheren Gesamt-Score von Qualitätsmerkmalen auf als die übrigen Gruppen. Die „unwahrscheinlichen“ und „sehr unwahrscheinlichen“ Fälle wurden zu einer Gruppe und die „wahrscheinlichen“ und „sehr wahrscheinlichen“ Fälle wurden für weitere Analysen ebenfalls zu einer Gruppe zusammengefaßt. Es zeigten sich signifikante Unterschiede bezüglich der Gesamtmenge der Qualitätsmerkmale in den Aussagen zwischen den beiden Gruppen, allerdings waren die Mittelwertdifferenzen zwischen den beiden Gruppen nicht sehr hoch. In dieser Studie wurden aus Überlegungen zur Beurteilerübereinstimmung (hierzu siehe unten) lediglich 14 Qualitätsmerkmale einbezogen. Hiervon fanden sich fünf Merkmale signifikant häufiger in „wahrscheinlichen“ als in „unwahrscheinlichen“ Aussagen („Unstrukturierte Darstellung“, „Quantitativer Detailreichtum“, „Raum-zeitliche Verknüpfungen“, „Interaktionsschilderungen“, „Wiedergabe von Gesprächen“), bei dem Merkmal „Schilderung ausgefallener Einzelheiten“ wurde das statistische Signifikanzniveau knapp verfehlt. Einige Merkmale traten insgesamt so selten auf, daß sie in den Berechnungen nicht berücksichtigt werden konnten („Schilderung nebensächlicher Einzelheiten“, „Phänomengemäße Schilderung unverstandener Handlungselemente“, „Indirekt handlungsbezogene

¹⁸ Die 98 Aussagen blieben aus einer Gesamtmenge von insgesamt 1.187 Interviews, nachdem folgende Fälle ausgeschlossen waren: a) Fälle mit unbekanntem Tatverdächtigen, b) außer der Aussage des Kindes lagen kaum andere Beweise vor, c) das geschädigte Kind war jünger als 4 oder älter als 13 Jahre, d) bei dem fraglichen Mißbrauch war es nicht zu einem körperlichen Kontakt zwischen Beschuldigtem und Geschädigtem gekommen.

Schilderungen“), bei anderen Merkmalen fanden sich hypothesenkonforme Mittelwertunterschiede („Schilderung von Komplikationen im Handlungsverlauf“, „Schilderung eigener psychischer Vorgänge“, „Spontane Verbesserungen der eigenen Aussage“), die Unterschiede waren jedoch nicht statistisch signifikant, wobei zu berücksichtigen ist, daß die Gruppe der „unwahrscheinlichen“ Fälle lediglich 13 Aussagen umfaßte. Hypothesenkonträre Ergebnisse lagen bezüglich zweier Merkmale vor: „Logische Konsistenz“ wurde in allen Aussagen signiert, das Merkmal „Schilderung psychischer Vorgänge des Beschuldigten“ wurde (nicht signifikant) häufiger in „unwahrscheinlichen“ als in „wahrscheinlichen“ Fällen kodiert.

3.2 Simulationsstudien

Bei Simulationsstudien werden Versuchspersonen typischerweise aufgefordert, über ein wahres und ein erfundenes Erlebnis zu berichten. Um wichtige Grundvariablen des forensisch relevanten Sachverhalts „sexueller Mißbrauch“ abzubilden, ohne die Grenzen ethischer Zumutbarkeit der kindlichen Teilnehmer der Studie zu überschreiten, forderten Steller, Wellershaus und Wolf (1992) beispielsweise Kinder auf, über einen Sachverhalt mit folgenden Eigenschaften zu berichten: Eigenbeteiligung, weitgehender Kontrollverlust des Betroffenen, vorwiegend negative emotionale Tönung; diese Charakteristika treffen insbesondere auf viele medizinische Eingriffe oder körperliche Angriffe zu. Der in der Undeutsch-Hypothese postulierte qualitative Unterschied konnte an acht Merkmalen besonders deutlich aufgezeigt werden („Logische Konsistenz“, „Quantitativer Detailreichtum“, „Raum-zeitliche Verknüpfungen“, „Schilderung von Komplikationen im Handlungsverlauf“, „Schilderung ausgefallener Einzelheiten“, „Schilderung nebensächlicher Einzelheiten“, „Phänomengemäße Schilderung unverstandener Handlungselemente“, „Indirekt handlungsbezogene Schilderungen“). Insgesamt wurde deutlich, daß die Qualitätsunterschiede zwischen wahren und unwahren Aussagen im wesentlichen für die Merkmalsgruppen „Allgemeine Merkmale“, „Spezielle Inhalte“ und „Inhaltliche Besonderheiten“ bestätigt wurden. Keine signifikanten Differenzen ergaben sich für die Merkmale der Gruppe „Motivationsbezogene Merkmale“, was damit zu tun haben mag, daß die Motivation kindlicher Teilnehmer einer als „Erzählwettbewerb“ konzipierten Studie mit der Motivation kindlicher Zeugen in Realsituationen nicht hinreichend vergleichbar ist.

Mittlerweile liegen eine Reihe von Simulationsstudien vor, die hier nicht im einzelnen referiert werden (Überblicke bei Greuel et al., 1998; Steller, Volbert & Wellershaus, 1993). Es ist festzuhalten, daß die ökologische Validität der experimentellen Untersuchungen zu den Glaubhaftigkeitsmerkmalen deutlich zugenommen hat. Während in den ersten Untersuchungen beispielsweise Filme als Stimulusmaterial verwendet wurden, nehmen neuere Untersuchungen auf tatsächliche Erlebnisse Bezug. Teilweise handelt es sich dabei um Erlebnisse, die in der Vergangenheit lagen und nicht der experi-

mentellen Kontrolle zugänglich waren, teilweise wurde in den Untersuchungen auch das Ereignis, über das die Aussage erfolgen sollte, in das Untersuchungsdesign einbezogen, so daß exakt kontrolliert werden konnte, welche Aussageteile tatsächlich erlebnisbasiert waren und welche nicht. Zunächst wurden vor allem Aussagen von Kindern analysiert, in späteren Untersuchungen lieferten auch erwachsene Versuchspersonen erlebnisbasierte und erfundene Schilderungen (z.B. Wolf & Steller, 1997). Trotz aller methodischen Unterschiede der einzelnen Untersuchungen stützen die empirischen Befunde durchgehend die Hypothese eines qualitativen Unterschieds zwischen wahren und erfundenen Aussagen (Greuel et al., 1998).¹⁹

Auch die Tatsache, daß nicht in allen Untersuchungen bei allen Qualitätsmerkmalen signifikante Unterschiede zwischen wahren und erfundenen Berichten auftraten, relativiert die generelle Bestätigung der Undeutsch-Hypothese nicht: Die Systematisierung von Qualitätsmerkmalen, die von Steller und Köhnken (1989) vorgenommen wurde (vgl. Abschnitt 1.2.2), basiert auf Publikationen, in denen praxisorientiertes Erfahrungswissen zusammengetragen wurde, welches vor allem an Fällen gewonnen wurde, in denen der Verdacht des sexuellen Mißbrauchs von Kindern bestand. Die Systematisierung umfaßt infolgedessen sowohl allgemeine Merkmale wie auch Merkmale, die einen Bezug zu dieser Thematik haben (z.B. „Phänomengemäße Schilderung unverstandener Handlungselemente“, „Schilderung psychischer Vorgänge des Angeschuldigten“, „Entlastungen des Angeschuldigten“ etc.). Das hat zur Folge, daß in den meisten Simulationsstudien, in denen ja regelmäßig eine andere inhaltliche Thematik gegeben ist, das Auftreten von Merkmalen, die einen inhaltlichen Bezug zur Mißbrauchsthematik ha-

¹⁹ Zu einem etwas anderen Ergebnis kamen Krahé und Kundrotas (1992). Sie fanden in einer Feldstudie, daß nur neun der 19 Qualitätsmerkmale häufiger in erlebnisbasierten als in erfundenen Aussagen gefunden wurden, während die übrigen zehn häufiger in unwahren Aussagen anzutreffen waren. Das Analysematerial bestand aus 30 authentischen Vernehmungprotokollen von Frauen, die eine Vergewaltigung angezeigt hatten, von denen je die Hälfte eingestandene Falschaussagen bzw. anhand von Tätergeständnissen oder Indizien als wahr klassifizierte Aussagen waren. Abgesehen von dem Problem eines validen Außenkriteriums für die Zuordnung von wahren und unwahren Aussagen, hat diese Untersuchung einen gravierenden Mangel, der die Gültigkeit der Ergebnisse erheblich in Frage stellt: Das verwendete Aussagematerial (polizeiliche Protokolle) stellt kein adäquates Analysematerial für die merkmalsorientierte Inhaltsanalyse dar, da es sich nicht um Wortprotokolle, sondern um zusammenfassende und gewichtete Nachschriften von Vernehmungen handelt. Wesentliche Inhalte, die Qualitätsmerkmale darstellen können, sind daher möglicherweise entfallen, da sie aus Sicht des protokollierenden Polizeibeamten als nebensächlich bzw. nicht zur Sache gehörig eingeschätzt wurden. Darüber hinaus dienten als Beurteiler zur Kodierung der Qualitätsmerkmale Polizeibeamte, die lediglich schriftlich mit der Systematisierung der Realkennzeichen von Steller und Köhnken (1989) vertraut gemacht worden waren (Wolf & Steller, 1997).

ben, auch in erlebnisbasierten Schilderungen wenig wahrscheinlich ist.²⁰ Darüber hinaus ist davon auszugehen, daß die Auftretenswahrscheinlichkeit der einzelnen Qualitätsmerkmale per se unterschiedlich verteilt ist. So können die Kriterien der Gruppe „Allgemeine Merkmale“ in allen erlebnisbasierten Aussagen auftreten, während beispielsweise das Merkmal „Komplikationen im Handlungsverlauf“ sich in einer erlebnisbasierten Schilderung nur finden läßt, wenn eine solche im wirklichen Geschehen auch tatsächlich stattgefunden hat. Schließlich hängt das Vorhandensein einzelner Qualitätsmerkmale mit hoher Wahrscheinlichkeit auch von Merkmalen der aussagenden Person ab, z.B. der Ausdrucksfähigkeit oder spezieller Ausdruckspräferenzen (Dahle & Wolf, 1997). Zu berücksichtigen ist ferner der Zusammenhang von Qualitätsmerkmalen und angewandten Interviewstrategien. Hershkowitz, Lamb, Sternberg und Esplin (1997) fanden, daß allgemeine Erzählaufforderungen zu signifikant längeren Antworten von Kindern führten als Befragerinterventionen, die die Aufmerksamkeit der Kinder auf spezifische Aspekte oder Details fokussierten, und die Ausführungen der Kinder auf Erzählaufforderungen enthielten mehr Qualitätsmerkmale als die Antworten auf andere Befragerinterventionen. Diese Ergebnisse unterstreichen die besondere Bedeutung einer „trichterförmigen“ Befragung (vgl. Abschnitt 2.2). Befragungen ohne offene Erzählaufforderungen vermindern die diagnostische Kraft der Aussageanalyse bzw. können sie vollständig invalidieren. Lamb, Sternberg, Esplin, Hershkowitz und Orbach (1997) haben darauf hingewiesen, daß vorliegende wissenschaftliche Untersuchungen zur merkmalsorientierten Inhaltsanalyse häufig auf Interviews von sehr unterschiedlicher Qualität basieren.

Unter Berücksichtigung dieser Überlegungen ist die Erwartung, Mittelwertunterschiede müßten sich in allen Untersuchungen, insbesondere auch in allen Simulationsstudien, bei allen Qualitätsmerkmalen finden, wenig plausibel. Von daher erfassen Bewertungen, daß diejenigen Merkmale am wichtigsten seien, die in den meisten Untersuchungen zwischen erlebnisbasierten und erfundenen Schilderungen diskriminieren (z. B. Ruby & Brigham, 1997), die Bedeutung der Qualitätsmerkmale nur unzureichend; sie berücksichtigen nicht ausreichend die unterschiedliche inhaltliche Validität der Merkmale.

3.3 Bewertung der empirischen Studien

Zusammenfassend ist festzuhalten, daß empirische Studien eine generelle Bestätigung der Undeutsch-Hypothese ergeben haben. Bei vielen Qualitätsmerkmalen finden sich jedoch in Gruppenvergleichen nur geringe Mittelwertunterschiede zwischen wahren und erfundenen Aussagen, und die Trefferquoten (also die richtigen Zuordnungen von wahren und unwahren Aussagen durch die Transkripte bearbeitenden Rater, die keine Informationen

²⁰ Ein Zusammenhang zwischen dem Vorliegen einzelner Merkmale mit dem Inhalt des Erlebnisses wurde beispielsweise von Hommers (1997) belegt.

über die aussagenden Personen oder das relevante Ereignis haben) liegen in der Regel zwar über dem Zufallsniveau, insbesondere bei der Klassifizierung der unwahren Aussagen gibt es jedoch nicht unerhebliche Fehlerspannen. Dies hat manche Autoren veranlaßt, die forensische Anwendung der Methodik in Frage zu stellen (z.B. Ruby & Brigham, 1997; Zaparniuk, Yuille & Taylor, 1995). Diese Schlußfolgerung basiert jedoch auf einem fundamentalen Mißverständnis: Wie weiter oben ausführlich dargestellt wurde (vgl. Abschnitt 2.2), nahmen Steller und Köhnken (1989) eine Systematisierung der *inhaltlichen* Qualitätsmerkmale vor, da in der bis dato vorliegenden Literatur einzelne Analyseschritte (Inhalts-, Motivations-, Konstanzanalyse) nicht immer präzise voneinander getrennt wurden. Nachdem diese Systematisierung vorlag, wurde insbesondere in der amerikanischen Rezeption der komplexe diagnostische Ansatz im wesentlichen auf die Liste der Qualitätsmerkmale reduziert, die darüber hinaus auch noch als ein Testverfahren, gewissermaßen als „Wahrheitstest“, mißverstanden wurde.²¹ So finden sich in vielen englischsprachigen Publikationen im einleitenden Teil durchaus Hinweise, daß die dort „Criteria-Based Content Analysis“ (CBCA) genannte merkmalsorientierte Inhaltsanalyse lediglich ein Teil einer umfassenden und die spezifischen Voraussetzungen der Person berücksichtigenden Methode zur Beurteilung der Glaubhaftigkeit einer Aussage („Statement Validity Assessment“ (SVA)) sei, dennoch wird in der Regel lediglich geprüft, inwieweit die merkmalsorientierte Inhaltsanalyse Testgütekriterien erfüllt, um auf der Basis dieser Ergebnisse Bewertungen abzugeben, inwieweit das gesamte Verfahren wissenschaftlich ausreichend abgesichert ist, um vor Gericht angewandt zu werden.²² Tatsächlich läßt sich aus diesen Untersuchungen ableiten, daß die Ergebnisse einer merkmalsorientierten Inhaltsanalyse auf der Basis eines Aussagetranskripts ohne zusätzliche Informationen über die aussagende Person oder das relevante Ereignis für das abschließende Glaubhaftigkeitsurteil nicht ausreichen. Ein solches Vorgehen ist aber von aussagepsychologischer Seite auch niemals vorgeschlagen worden, statt dessen wurde von vornherein die *intraindividuelle* Perspektive bei der Schlußfolgerung von der Qualität zur Glaubhaftigkeit einer Aussage betont (vgl. auch Abschnitt 1.3). *Lege artis* würde in der Praxis ohne Kenntnis des Entwicklungsstandes, der indivi-

²¹ Vergleichbare Mißverständnisse finden sich auch im deutschsprachigen Raum, z. B. Plaum, 1998.

²² In der Übersicht von Ruby und Brigham (1997) beispielsweise wird dieser Prozeß besonders deutlich: „Criteria-based content analysis“ wird in der Einleitung bereits fälschlicherweise als eine Komponente einer umfassenderen *Interviewtechnik* und nicht als Bestandteil eines komplexen Auswertungsverfahrens bezeichnet. Es findet sich dann zwar noch ein Hinweis, daß CBCA nicht ohne die anderen Komponenten Anwendung finden sollte. Im weiteren wird mit dem Hinweis, daß CBCA das Kernstück der Methode sei, nur noch geprüft, inwieweit die merkmalsgestützte Analyse eines Aussagetranskripts zu einer richtigen Klassifizierung von wahren und erfundene Schilderungen führt, und die Einbindung der merkmalsgestützten Analyse in das gesamte Auswertungsmodell wird im weiteren nicht mehr angesprochen.

duellen kognitiven Fähigkeiten, der relevanten Kenntnisse des Zeugen und der Aussagegeschichte nur auf der Basis eines Aussagetranskripts keine gutachterliche Stellungnahme zur Glaubhaftigkeit der Aussage erfolgen.²³

In den referierten Studien wurde also bisher nur die Hypothese über einen qualitativen Unterschied zwischen erlebnisbegründeten und erfundenen Schilderungen geprüft. Dieser wurde generell bestätigt. Ausführungen über die Güte forensisch-aussagepsychologischer Glaubhaftigkeitsbegutachtungen lassen sich aus diesen Untersuchungen nicht ableiten, da die merkmalsorientierte Inhaltsanalyse lediglich ein Bestandteil des Vorgehens ist. Will man Aussagen über die Trefferquoten gutachterlicher Urteile abgeben, müßten Untersuchungen durchgeführt werden, bei denen den Experten die gesamte relevante Information und nicht nur ein Interviewtranskript zur Verfügung steht.

Untersucht wurde in einigen Arbeiten auch die Interrater-Reliabilität, also die Übereinstimmung verschiedener Beurteiler bei der Einschätzung der Qualitätsmerkmale sowie die Test-Retest-Reliabilität, also die Übereinstimmung eines Beurteilers zu verschiedenen Zeitpunkten (Horowitz, Lamb, Esplin, Boychuk, Krispin & Reiter-Lavery, 1997). Die Test-Retest-Reliabilität erwies sich als hoch; die Interrater-Reliabilität war bei den meisten Merkmalen ebenfalls zufriedenstellend, einige Merkmale ließen sich jedoch nicht mit zufriedenstellender Reliabilität erfassen. Tully (1998) hat allerdings darauf hingewiesen, daß bei der Überprüfung von Beurteilerübereinstimmungen auch berücksichtigt werden muß, inwieweit in dem zu analysierenden Material prägnante Merkmale tatsächlich vorhanden sind. Wenn das Material uneindeutig ist und die Beurteiler im Rahmen der Untersuchung genötigt werden, zu entscheiden, ob ein Merkmal vorhanden ist oder nicht, ist mit geringeren

²³ Tatsächlich resümieren beispielsweise Ruby und Brigham (1997) auch lediglich: „Although the CBCA technique shows some promise in enabling raters to differentiate true from false statements, the authors conclude that the presentation of expert testimony derived from CBCA analyses of an individual child would be premature and unwarranted“ (S. 705). Sie beziehen sich also selbst nur auf eine Auswertungskomponente, nehmen jedoch diesen Teil für das Ganze und erwecken durch die Art ihrer Darstellung den Eindruck, ihr Resümee gelte für die gesamte Methodik. Auch Lamb, Sternberg et al. (1997) resümieren: „The findings generally affirm the validity of Undeutsch's (1982) observations about the characteristics of credible accounts and thus suggest that the underlying principles may be of value to forensic investigators and fact finders. They also underscore that CBCA scores should not yet – and perhaps should never – be used in forensic contexts to evaluate individual statements, however. The statistically significant relationship between plausibility and CBCA scores is not sufficiently strong to support opinions about the veracity of individual accounts solely on the basis of the CBCA scores. Even if high scores were indicative of plausibility, furthermore, low scores may be attributable to a variety of causes (including poor interview quality and reticence on the part of the child) independent of veracity“ (S. 191).

Übereinstimmungen zu rechnen als bei eindeutigem Material. Er argumentiert, die Raterübereinstimmung bei möglicherweise wenig eindeutigem Material und der Vorgabe, eine Einschätzung darüber abgeben zu *müssen*, ob ein Merkmal vorhanden ist oder nicht, sage nichts darüber aus, wie hoch die Beurteilerübereinstimmung in tatsächlichen Fällen sei, wenn die Beurteiler der Auffassung seien, es handele sich um ein prägnantes Merkmal. Darüber hinaus ist erneut darauf hinzuweisen, daß es sich bei der Systematik der Merkmale nicht um ein Testverfahren handelt und die Auflistung der einzelnen Merkmale im wesentlichen erfolgte, um eine Inhaltsanalyse zu erleichtern. Eine Textstelle kann in vielen Fällen aber durchaus mehreren Merkmalen zugeordnet werden, von daher mag es bei Zuordnung zu einem spezifischen Merkmal Divergenzen zwischen Beurteilern geben, auch wenn Konsens darüber besteht, daß eine bestimmte Textstelle im Sinne der Qualitätsanalyse prägnant ist.

Tully (1998) führt aus, daß es sich bei der merkmalsorientierten Inhaltsanalyse nicht um ein eigenes Instrument oder eine eigene Technik handelt, sondern um eine Systematik, die auf dem wissenschaftlichen Verständnis der zugrundeliegenden Prozesse basiert. Die referierten Forschungen der letzten Jahre haben eine wesentliche zugrundeliegende Behauptung bestätigt, nämlich die der Undeutsch-Hypothese letztlich zugrundeliegende Annahme, daß es die menschliche Informationsverarbeitungskapazität übersteigt, eine qualitativ einer erlebnisgestützten Aussage gleichwertige Sachverhaltsschilderung zu erfinden (Greuel et al., 1998, S. 159). Tully führt als Analogie gängige Routinen medizinischer Untersuchungen an, welche angewendet werden, wenn ein Patient beispielsweise über bestimmte Schmerzen klagt. Diese Untersuchungsprotokolle würden ebenfalls auf wissenschaftlichen Erkenntnissen basieren, ohne ein Testverfahren zu sein. Welche spezifischen Anzeichen oder Symptome in den diagnostischen Entscheidungsprozeß eingehen, sei abhängig davon, wie prägnant die Symptome im individuellen Fall seien. Unklare und widersprüchliche Symptome würden mit Vorsicht behandelt, und es gebe diesbezüglich Beurteilerdifferenzen zwischen Ärzten. Wenn die Symptome ausgeprägt seien, sei aber davon auszugehen, daß kompetente Ärzte diese Symptome und die verschiedenen hiermit verbundenen Möglichkeiten unter Berücksichtigung anderer klinischer Befunde, der Vorgeschichte etc. erkennen würden. Tully argumentiert, daß es abwegig sei, 100 Patienten von 20 Ärzten untersuchen zu lassen und bezüglich jedes potentiell relevanten Symptoms eine Einschätzung in „vorhanden“ oder „nicht vorhanden“ abgeben zu lassen und am Ende der Studie zu resümieren, die Symptome, bei denen für die Übereinstimmung der Ärzte keine idealen statistischen Kennwerte erreicht wurden, sollten in Zukunft bei der Erstellung einer Diagnose nicht mehr berücksichtigt werden. In Fortführung dieses Gedankens könnte man ausführen, daß aus Untersuchungen, die zum Ergebnis haben, daß die einzelnen Symptome nicht ideal diskriminieren zwischen Patienten, die eine be-

stimmte Erkrankung haben und solchen, die sie nicht haben, nicht abzuleiten ist, daß diese Erkrankung nicht richtig diagnostiziert werden kann.

Daß Prüfungen von Trefferquoten ausschließlich auf der Basis von merkmalsorientierten Analysen von Aussagetranskripten für die Bewertung der gutachterlichen Praxis wenig aussagekräftig sind, wurde bereits erwähnt. Für den inhaltsorientierten Ansatz der Glaubhaftigkeitsbegutachtung spricht allerdings, daß Trefferquoten trotz des gegenüber der Praxis erheblich reduzierten Analysematerials zum Teil meistens signifikant besser sind als die von Beurteilern, die mit diesem Auswertungssystem nicht vertraut sind (Lehr, 1995; Steller, 1989; in einer Untersuchung von Landry und Brigham [1992] wurden allerdings nur die wahren Schilderungen von den kundigen Beurteilern besser klassifiziert als von den unkundigen). Von daher belegen die vorhandenen Studien nicht nur die Undeutsch-Hypothese, sondern sie zeigen auch, daß durch die Systematisierung der Qualitätsmerkmale nicht nur Wissen zur Beurteilung der Glaubhaftigkeit einer Aussage expliziert wird, das ohnehin intuitiv benutzt wird, sondern daß die Fokussierung auf bestimmte inhaltliche Aspekte in der Regel bereits zu einer signifikanten Verbesserung der Klassifikation führt, selbst wenn einem nichts anderes als ein Aussagetranskript zur Verfügung steht und man keine weiteren Informationen über die aussagende Person hat. Wahre Aussagen werden durchgängig besser erkannt als unwahre Aussagen (Landry & Brigham, 1992; Lehr, 1995; Wellershaus & Wolf, 1989). Dies entspricht auch den Ergebnissen aus Studien zur verhaltensorientierten Glaubwürdigkeitsattribution und steht in Zusammenhang mit einer generellen Tendenz, Aussagen eher als „wahr“ denn als „unwahr“ zu beurteilen (vgl. Köhnken, 1990).

Auch Forderungen nach standardisierten diagnostischen Entscheidungsregeln oder normativen Mengenangaben von erfüllten Qualitätsmerkmalen beruhen auf einem falschen Verständnis der Qualitätsmerkmale als Testverfahren. Die in manchen Publikationen angegebenen Entscheidungsregeln (z.B. Zaparniuk, Yuille & Taylor, 1995), nach denen angeblich eine bestimmte Anzahl bzw. Kombination von Kriterien darauf verweist, daß die Aussage erlebnisbegründet ist, sind theoretisch abwegig und empirisch nicht untermauert.²⁴ Greuel et al. (1998) führen aus, daß es sich bei den Qualitätsmerkmalen nicht um „Items eines wie auch immer gearteten „Wahrheitstests“ [handelt], die über Aufsummierungen oder andere mathematische Verknüpfungen zu einem Gesamtergebnis in bezug auf den wahrscheinlichen Erlebnisgehalt einer Aussage führen könnten. Eine derartige *elementaristische Sichtweise* ist der Logik einer einzelfallorientierten und hypothesengeleiteten Diagnostik fremd“ (S. 159). Sie fügen hinzu (ebd., S.161), auch wenn der Wunsch nach leicht

²⁴ Die dabei vorgenommenen Verweise auf angebliche Regeln in Publikationen von Steller sind schlicht falsch. In Steller (1989, S. 136) findet sich ebenfalls eine diesbezüglich falsche Feststellung, die vom Herausgeber des Buches ohne Rückfrage beim Autor eingefügt wurde.

anwendbaren und transparenten Entscheidungsregeln verständlich sei, würden sich Entscheidungsregeln für alle denkbaren Fall- und Personenkonstellationen eben nicht festlegen lassen, und resümieren (ebd., S. 162): „Letztlich ist und bleibt die Datenintegration zur Bestimmung des Erlebnisbezugs einer Aussage immer ein Prozeß der einzelfallbezogenen psychodiagnostischen Urteilsbildung und damit – wie das Gerichtsverfahren selbst – im Ergebnis ein interpretativer Akt im Sinne einer im forensischen Kontext konstruierten Wirklichkeit. Die Qualität der zugrunde liegenden Sachverständigenleistung wird sich in letzter Konsequenz also daran messen lassen müssen, inwieweit die im Einzelfall zur Anwendung gekommenen Bewertungs- und Verknüpfungsregeln transparent und damit überprüfbar gemacht worden sind“.

3.4 Neue Forschungsrichtungen

Wie weiter oben ausgeführt wurde, sind ein Teil der Kriterien eng mit der Mißbrauchsthematik verknüpft und auf andere Inhalte wahrscheinlich nicht ohne weiteres übertragbar, obwohl die zugrundeliegende Annahme allgemeine Gültigkeit haben dürfte, daß die Informationsverarbeitungskapazität mit der Erfindung einer qualitativ einer erlebnisbasierten Aussage gleichwertigen Sachverhaltsschilderung überfordert ist. Von Sporer (1997a) sind Bemühungen unternommen worden, inhaltlich unspezifische Qualitätsmerkmale zusammenzutragen, indem er die Merkmale aus der Systematisierung von Steller und Köhnken (1989) mit den sogenannten „Reality-monitoring criteria“ zu integrieren versucht hat.

Völlig unabhängig von der Rechtspsychologie wurden in der Kognitionspsychologie in Arbeiten zur Realitäts- bzw. Quellenüberwachung (Johnson, Hashtroudi & Lindsay, 1993; Johnson & Raye, 1981) qualitative Charakteristika beschrieben, mit denen Personen im Alltag (meist implizit) Unterscheidungen zwischen Erinnerungen an tatsächlich erlebte Ereignisse und anders generierten Erinnerungen treffen. Diese Kriterien werden beispielsweise benutzt zur Differenzierung zwischen Erinnerungen an tatsächliche Erlebnisse und Träume oder zwischen Erinnerungen an tatsächliche Handlungen und Erinnerungen an bloße Überlegungen, diese Handlungen durchzuführen. Dem Modell nach beinhalten extern generierte Erinnerungen mehr sensorische Informationen, wie zum Beispiel Wissen über Farbe und Form eines Objektes oder über Geruch und Geschmack, mehr kontextuelle Informationen, also Zeitangaben und Informationen über räumliche Anordnungen von Personen und Objekten, mehr Informationen über Emotionen und mehr sinnvolle, semantische Details wie die inhaltliche Einbettung in einen Kontext oder das Wissen über die Funktion eines Objektes. Bei internal erzeugten Erinnerungen sollen sich dagegen mehr kognitive Operationen finden, die während der Elaborierung mitverarbeitet worden sind wie Reflexionen, Assoziationen oder Entscheidungsprozesse (Sporer & Küpper, 1995). Sporer (1997 a, b) hat in einer ersten Untersuchung auf der Basis von Faktorenanalysen beiden Ansätzen gemeinsame zugrundeliegende Dimensionen („Logi-

sche Konsistenz und Realitätsnähe“, „Gefühle und Gedanken“, „Klarheit“, „Quantitativer Detailreichtum“, „Interaktionen“, „Kontextuelle Einbettung und zeitliche Informationen“) extrahiert. Hieraus ergeben sich Ansätze für eine theoretische Integration, aufgrund derer schließlich auch Anwendungsbereiche und -grenzen exakter beschrieben werden können. Auf dieser Basis könnten dann auch „domänenspezifische“ Kriterien (Sporer, 1997a, S. 84), also für bestimmte Altersgruppen oder bestimmte Inhalte relevante Merkmale, konzipiert werden.

Offene Forschungsfragen bestehen bezüglich der Bedeutung einzelner Qualitätsmerkmale in verschiedenen Altersgruppen. Entwicklungspsychologische Fragestellungen sind bislang wenig systematisch untersucht worden. Saacke (1995) argumentierte, daß bestimmte Qualitätsmerkmale („Logische Konsistenz“, „Indirekt handlungsbezogene Schilderungen“, „Spontane Verbesserungen der eigenen Aussage“, „Einwände gegen die Richtigkeit der eigenen Aussage“) in den Aussagen voroperational denkender Kinder (zwischen 2 und 7 Jahren) aufgrund kognitiver Faktoren eine geringe Auftretenswahrscheinlichkeit haben. Sie zeigte in einer Simulationsstudie, daß einige Merkmale in den wahren Schilderungen von Erstkläßlern a) insgesamt sehr selten und b) signifikant seltener produziert wurden als in den Aussagen von Viertkläßlern. Allerdings fanden sich auch einige andere Merkmale seltener in den Aussagen der Erstkläßler als in den Schilderungen der Viertkläßler („Quantitativer Detailreichtum“, „Schilderungen ausgefallener Details“, „Eingeständnis von Erinnerungslücken“).

Auch in anderen Studien zeigten sich in der Regel Zusammenhänge zwischen dem Alter und der Menge der Qualitätsmerkmale. Lamb, Sternberg et al. (1997) fanden, daß die Aussagen älterer Kindern insgesamt mehr Qualitätsmerkmale enthielten als die Schilderungen jüngerer Kinder (Alter der Gesamtstichprobe lag zwischen 4 und 13 Jahre; vgl. auch Horowitz et al., 1997). Anson, Golding und Gully (1993) zeigten, daß sechs der 19 Qualitätsmerkmale signifikant mit Alter korreliert waren: „Logische Konsistenz“, „Raum-zeitliche Verknüpfungen“, „Interaktionsschilderungen“, „Wiedergabe von Gesprächen“, „Entlastung des Angeschuldigten“, „Deliktspezifische Aussageelemente“ (Alter der Gesamtstichprobe ebenfalls zwischen 4 und 13 Jahren). In der Feldstudie von Boychuk (1991) traten lediglich die Merkmale „Interaktionsschilderungen“ und „Eingeständnis von Erinnerungslücken“ bei älteren Kindern häufiger auf (Alter der Gesamtstichprobe zwischen 4 und 16 Jahren). Lamers-Winkelmann und Buffing (1996) untersuchten den Zusammenhang von Alter und Qualitätsmerkmalen im Rahmen einer niederländischen Feldstudie anhand der Aussagen von zwei- bis 12jährigen Kindern und fanden einen Zusammenhang für sechs Merkmale („Raum-zeitliche Verknüpfungen“, „Interaktionsschilderungen“, „Wiedergabe von Gesprächen“, „Darstellung nebensächlicher Details“, „Eingeständnis von Erinnerungslücken“, „Deliktspezifische Aussageelemente“).

Ein gegenüber den bisher referierten gruppenstatistischen – d.h. vor allem auf die Untersuchung universeller Gesetzmäßigkeiten (z. B. Gültigkeit der Undeutsch-Hypothese) und allgemeiner Zusammenhänge (z. B. Alterseinflüsse) zielenden – Forschungsstrategien ganz anders angelegter Ansatz zur Untersuchung der Validität inhaltlicher Aussageanalysen wurde kürzlich von Dahle (1997a) entwickelt und vorgeschlagen. Das Konzept bietet vor allem eine Methodik zur Überprüfung der Leistungsfähigkeit von Aussageanalysen in konkreten Einzelfällen, da es auf der systematischen Analyse der individuellen Aussagebesonderheiten von Einzelpersonen basiert. Das Prinzip besteht darin, anhand von Aussagen einer Person über mehrere Erlebnisse und frühere Ereignisse eine Auswahl individuell geeigneter (d. h. differenzierender) Inhaltsmerkmale vorzunehmen und aus diesen nach den Regeln psychometrischer Testentwicklung eine spezifische Quantifizierungsmethode von Aussagequalitäten dieser Person zu entwickeln. In einem zweiten Schritt werden anhand der Skala und der Beispielaussagen der Person intraindividuelle Vergleichsnormen der Qualitäten wahrer und unwahrer Berichte abgeleitet. Diese ermöglichen es, für neue Berichte dieser Person individuelle Wahrscheinlichkeiten ihres Wahrheitsstatus zu bestimmen und unter Zugrundelegung geeigneter Sicherheitskriterien entsprechende Zuordnungen vorzunehmen. Die Auswahl der Entscheidungskriterien ist dabei grundsätzlich variabel, jedoch geht jede Erhöhung der Urteilssicherheit mit einer Erhöhung von Fällen einher, die mit der geforderten Sicherheit nicht mehr zuzuordnen sind. In den von Dahle (1997a) bzw. Dahle und Wolf (1997) vorgelegten Einzelfallanalysen wurden z. B. bei Zugrundelegung einer 95 %igen Urteilssicherheit bei insgesamt 155 Zuordnungen nur 6 Fehler (3,87 %) begangen, eine solchermaßen sichere Zuordnung war jedoch nur in knapp 40 % der Fälle möglich, die restlichen Fälle fielen in die Kategorie „unentscheidbar“.

Wegen des erheblichen Aufwands²⁵ eignet sich die Methodik kaum für eine regelhafte Übertragung in die Praxis als eigenständige diagnostische Strategie, sie hat ihre Stärken vor allem bei der systematischen Untersuchung der einzelfallbezogenen Validität von Aussageanalysen. Die Möglichkeit der Quantifizierung von Zuordnungs- und Irrtumswahrscheinlichkeiten bietet jedoch auch Ansatzpunkte zur gezielten Beforschung der Bedeutung von Einzelaspekten und der Voraussetzungen für qualitativ hochwertige Aussageanalysen. So fand Dahle beispielsweise, daß einige der von Steller und Köhnken (1989) zusammengestellten Qualitätsmerkmale mit hoher Regelmäßigkeit zur Trennung wahrer und unwahrer Berichte beitragen (d. h. *interindividuell* bedeutsam sind), andere hingegen nur in bestimmten Einzel-

²⁵ Neben der zu beurteilenden Aussage müßte eine Person zahlreiche wahre und fiktive Schilderungen als Vergleichsmaßstab liefern (in den hier erwähnten wissenschaftlichen Untersuchungen waren es je 15 wahre und fiktive Aussagen).

fällen (d. h. potentiell) bedeutsam sind²⁶. Weitere Befunde unterstreichen die Notwendigkeit einer sorgfältigen Schulung der Auswerter von Aussageanalysen: So war es durch geeignete Strategien (Gruppenleistung mehrerer Auswerter oder Herbeiziehung besonders praxiserfahrener Auswerter) möglich, die in der o. g. Untersuchung (auf Auswertungen methodisch unterwiesener Psychologiestudentinnen beruhende) vorgefundene Fehlerquote von 3,87 % auf 0 zu reduzieren und gleichzeitig die Quote nicht entscheidbarer Fälle von 60 % auf rund 30 % zu senken (Dahle, 1997b).

Auf Grundlage der einzelfallexperimentellen Untersuchungsmethodik nach Dahle stellten sich in einer Studie von Hustedt (1998) die von Steller und Köhnken (1989) zusammengetragenen inhaltlichen Qualitätsmerkmale gegenüber den Realitätsüberwachungskriterien (Sporer & Küpper, 1995) sowie einer Zusammenstellung von Merkmalen des Ausdrucks- und Sprechverhaltens (vgl. zusammenfassend Köhnken, 1982; 1990) als die effizienteste Grundlage für Aussageanalysen zur Beurteilung des Wahrheitsgehalts verbaler Berichte dar. Im direkten Vergleich wies die auf den Realkennzeichen beruhende Analyse die besten Trefferquoten auf, die sich auch durch Hinzunahme von Einzelkriterien aus den anderen genannten Bereichen nicht wesentlich verbessern ließen.

3.5 Inhaltliche Qualität in suggerierten Aussagen

Angesichts der Entwicklung der letzten Jahre ist von besonderer Bedeutung, inwieweit die inhaltlichen Merkmale geeignet sind, nicht nur zwischen erlebnisbasierten und erfundenen Aussagen zu trennen, sondern auch zwischen erlebnisbasierten und *suggestierten*. Ausgangspunkt der Verdachtsbildung ist in diesen Fällen oft nicht eine Bekundung von Kindern über sexuelle Mißbrauchserfahrung, sondern der Verdacht entsteht durch die Ausdeutung von sogenannten Signalen. Besonders häufig zu finden ist eine einseitige Interpretation unspezifischer Verhaltensweisen (wie Schlafstörungen, Einnässen, Angst etc.), obwohl belegt ist, daß kein spezifisches sexuelles Mißbrauchssyndrom existiert (Kendall-Tackett et al., 1993). Wenn sich ein Anfangsverdacht aufgrund solcher Überinterpretationen verdichtet hat, wird Kindern häufig mit Techniken begegnet, die stark suggestive Wirkung haben. So kommt es zur Durchführung langfristiger „Aufdeckungsarbeit“ mit oft indirekten wiederholten Befragungen, zum Teil aber auch mit direkten Vorgaben, bedingungslosem Akzeptieren und Verstärken von Beschreibungen sexueller Mißbrauchshandlungen, auch wenn diese vage, widersprüchlich oder sogar unrealistisch sind, und der Deutung von Schweigen und Verneinung als „Noch-nicht-bereit-Sein“ zur Verbalisierung sexueller Mißbrauchserfahrungen.

²⁶ Vgl. Dahle, 1997a, S.19; siehe hierzu auch die differentialpsychologische Unterscheidung „guter“ und „schlechter“ Lügner von Hommers (1997).

Mittlerweile liegt eine Fülle von Untersuchungen vor, die belegen, daß die Verwendung von suggestiven Methoden zu nicht erlebniskongruenten Schilderungen auch über persönlich bedeutsame und belastende Ereignisse von Kindern und möglicherweise zu länger bestehenden Pseudoerinnerungen führen kann. Der aktuelle Forschungsstand zu Suggestionseffekten in Kinderaussagen ist an anderer Stelle ausführlich referiert worden (Ceci & Bruck, 1993, 1995; Volbert, 1997; Volbert & Pieters, 1996). Hier können nur Beispiele genannt werden, um die inhaltliche Nähe zwischen empirischer Forschung und Anwendungspraxis zu veranschaulichen. So berichteten Kinder auf entsprechende Fragen fälschlicherweise, die Helferin habe ihnen beim Arztbesuch ins Ohr gepustet oder am Knie geleckt, und blieben über einen dreimonatigen Zeitraum wiederholter Befragungen bei diesen Behauptungen (Ornstein, Gordon & Larus, 1992). Kinder gaben an, ein Mann habe ihnen etwas Ekeliges in ihren Mund getan (Poole & Lindsay, 1995), behaupteten, der Kinderarzt habe einen Finger oder einen Stock in ihre Genitalien eingeführt (Bruck, Ceci, Francoeur & Barr, 1995), oder erklärten, ein Mann habe ihre Freunde angefaßt und auf die Lippen geküßt und einige Kleidungsstücke ausgezogen (Lepore & SESCO, 1994).

Daß suggestive Einflußnahmen Effekte haben können, läßt sich aufgrund der Forschungslage nicht mehr bezweifeln. Allerdings führen suggestive Einflußnahmen keineswegs immer zu einer tatsächlichen Beeinflussung einer Aussage. Von daher stellt sich die Frage, in welchen Situationen und unter welchen Bedingungen mit Suggestionseffekten zu rechnen ist. Hier sind vor allem drei Aspekte zu betonen:

Befragervoreinstellung. Suggestive Einflußnahmen erfolgen nicht ausschließlich und wahrscheinlich auch nicht hauptsächlich durch suggestiv formulierte Fragen (z.B. „Und dann hat er dich ausgezogen, nicht wahr?“); das Hauptcharakteristikum suggestiver Interviews besteht vielmehr in einer Voreinstellung des Interviewers, einem sog. „Interviewer bias“ (Ceci & Bruck, 1995). Diese Voreinstellung ist gekennzeichnet durch A-priori-Annahmen darüber, daß bestimmte Ereignisse tatsächlich passiert sind, sowie durch eine Befragung, die auf die Bestätigung dieser Annahme orientiert ist: Der Interviewer sammelt Informationen, die geeignet sind, die Vorabhypothese zu unterstützen, der Interviewerhypothese widersprechenden Auskünften des Kindes wird nicht weiter nachgegangen, Informationen zur Abklärung von Alternativhypothesen werden nicht gesammelt, inkonsistente oder objektiv unmögliche Angaben des Kindes ignoriert oder im Rahmen der Ausgangshypothese interpretiert. Konkret bedeutet das, daß keine offenen, sondern sehr direkte, teilweise suggestive Fragen gestellt und innerhalb einer oder mehrerer Befragungen wiederholt werden. Produziert das Kind nicht die erwartete Information, erfolgen erneute Befragungen, zum Teil über einen langen Zeitraum. Bei den Bemühungen, eine unterstützende Atmosphäre für das Kind zu schaffen, kommt es zur selektiven Verstärkung von Äußerungen,

die konsistent mit der Interviewerhypothese sind (z.B. durch Kopfnicken oder durch Lob dafür, daß das Kind so tapfer sei, über die Vorfälle zu berichten), während andere Äußerungen nicht weiter beachtet werden.

Stärke der Gedächtnisspur. Viele Befunde sprechen dafür, daß suggestive Einflußnahmen dann besonders wirksam sind, wenn die Erinnerung (die Gedächtnisspur) für das relevante Ereignis besonders schwach oder gar nicht vorhanden ist und die Gedächtnisspur für die suggerierte Information besonders stark ist (z.B. Pezdek & Roe, 1995). Damit ist eher zu rechnen, wenn zwischen fraglichem Ereignis und suggerierter Information ein längeres Intervall liegt, wenn das Kind zum angeblichen Tatzeitpunkt sehr jung war, wenn die Suggestion wiederholt erfolgt bzw. wenn sie von mehreren unterschiedlichen Personen, eventuell auch noch in unterschiedlicher Weise an das Kind herangetragen wird.

Sozialpsychologische Faktoren. Es ist davon auszugehen, daß Fehlinformationen in bestimmten Fällen auch dann übernommen werden, wenn eine richtige Erinnerung an das ursprüngliche Ereignis besteht (McCloskey & Zaragoza, 1985). Dies geschieht entweder, weil die beeinflusste Person ihre eigene Erinnerung als weniger zuverlässig einschätzt als die von einem kompetenten Dritten vermittelte Information oder weil sich die beeinflusste Person entsprechend den angenommenen Erwartungen verhält. Bei einem Kind kann das Bemühen, die Erwartungen der erwachsenen Autoritätsperson zufriedenzustellen, besonders relevant sein (vgl. Zaragoza, Dahlgreen & Muench, 1992). Für Kinder ist es üblich, Informationen von Erwachsenen in kommunikativen Prozessen zu erwerben. Vermitteln Erwachsene dem Kind eine spezifische Auffassung eines Ereignisses, kann das dazu führen, daß das Kind seine eigene Erinnerung revidiert, die sich im Widerspruch zur Vermutung des Erwachsenen befindet, weil das Kind den Erwachsenen für kompetenter hält.

Bislang existiert kein theoretisches Modell, aufgrund dessen die Determinanten von Suggestionseffekten vollständig geklärt werden könnten. Aufgrund der vorhandenen Forschung ist davon auszugehen, daß Suggestionprozessen ein multidimensionales Bedingungsgefüge zugrunde liegt, bei dem ein Reihe von kognitiven und sozialpsychologischen Faktoren, individuelle Voraussetzungen, Aufgabenmerkmale und Abfragebedingungen wirksam sind (Bruck, Ceci & Hembrooke, 1998).

Das Vorliegen von suggestiven Bedingungen impliziert nicht zwangsläufig, daß eine Aussage auch tatsächlich hierdurch beeinflusst wurde. Von besonderer Relevanz ist von daher in der praktischen Arbeit die Frage, ob eine individuelle Aussage, auch wenn sie unter suggestiven Bedingungen zustande gekommen ist, tatsächlich als suggeriert oder doch als erlebnisbegründet zu klassifizieren ist. Einige Untersuchungen, bei denen kein spezifisches Analy-

semodell vorgegeben wurde, lassen annehmen, daß eine richtige Kategorisierung kaum möglich ist. So wurden beispielsweise in Untersuchungen von Ceci, Huffman et al. (1994), Ceci, Loftus et al. (1994) sowie von Leichtman und Ceci (1995) Praktikern und Wissenschaftlern verschiedener Berufsgruppen, die sich mit der Befragung von Kindern in Fällen des Verdachts auf sexuellen Mißbrauch befassen, Videos von suggerierten Kinderaussagen gezeigt. (Da es sich nicht um Felduntersuchungen handelte, bei denen nie ein sicheres Außenkriterium vorliegt, bezogen sich die Aussagen in diesen Interviews nicht auf sexuellen Mißbrauch.) Die Trefferquoten bei der Zuordnung, ob es sich um erlebnisbegründete oder fiktive Ereignisse handelte, lagen im Durchschnitt nicht oberhalb des Zufallsniveaus. Dies scheint einerseits darauf zurückzuführen zu sein, daß die suggerierten Aussagen viele Details enthielten. Zudem erschienen die Kinder von ihren Aussagen subjektiv überzeugt, so daß sich auch in ihrem Aussageverhalten kein Hinweis für eine nicht erlebnisbegründete Aussage ergab.

Von besonderem Interesse für die aussagepsychologische Begutachtungspraxis ist daher die bereits einleitend formulierte Frage, ob sich auch die Qualität einer suggerierten Aussage von der einer erlebnisbegründeten unterscheidet. Bereits unter theoretischer Perspektive ist eine Hypothese hierzu nicht ohne weiteres zu formulieren. Knüpft man an die weiter oben dargestellte Prämisse für die Unterscheidung zwischen erlebten und erfundenen Aussagen an – nämlich die Konzeptualisierung einer Aussage als Leistung (vgl. Abschnitt 1.2.1) –, so läßt sich argumentieren, daß diese Prämisse bei der Unterscheidung zwischen erlebnisbasierten und suggerierten Aussagen nicht gegeben ist. Der suggestiv beeinflusste Zeuge muß keine kognitive Energie auf kreative und Kontrollprozesse verwenden, da er keine Aussage erfindet, sondern in vielen Fällen auf vermeintliche Erinnerungen rekurriert. Die motivationsbezogenen Glaubhaftigkeitsmerkmale können daher kaum geeignet sein, zwischen erlebnisbasierten und suggerierten Schilderungen zu differenzieren, da diese sich ausschließlich auf die Frage beziehen, ob ein falsch aussagender Zeuge sich in dieser Weise darstellen würde, während ein suggestiv beeinflusster Zeuge eine subjektiv wahre Schilderung abgibt. Bei einer nicht bewußten Veränderung des Gedächtnisinhalts fehlt die Intentionalität der Täuschung, es liegt eben kein motiviertes und zielgerichtetes Verhalten vor, eine Selbstpräsentation als glaubwürdiger Kommunikator als Teil einer Täuschungsstrategie ist daher nicht notwendig (vgl. Köhnken, 1990; 1997).

Wie weiter oben ausgeführt wurde, werden andererseits in Arbeiten zur Quellenüberwachung (Johnson et al., 1993; Johnson & Raye, 1981) qualitative Charakteristika genannt, mit denen Unterscheidungen zwischen Erinnerungen an tatsächlich erlebte Ereignisse und anders generierte Erinnerungen getroffen werden, die teilweise den forensischen Glaubhaftigkeitsmerkmalen ähneln. Dieser Ansatz soll auch auf Irrtümer und unabsichtliche Verfä-

schungen übertragbar sein (Schooler, Clark & Loftus, 1988; Schooler, Gerhard & Loftus, 1986), so daß man argumentieren könnte, solche qualitativen Merkmale sollten auch geeignet sein, zwischen Aussagen über Erinnerungen an tatsächliche Erlebnisse und Schilderungen induzierter Erinnerungen zu unterscheiden.

Empirische Untersuchungen hierzu existieren kaum. In einer am Institut für Forensische Psychiatrie der FU Berlin durchgeführten unveröffentlichten Untersuchung (Lehr, 1995; vgl. auch Volbert & Pieters, 1997) fanden sich kaum signifikante Unterschiede, allerdings gab es aufgrund der sehr kurzen relevanten Handlungsepisoden und des jungen Alters der Kinder (4-6jährige Kinder) bzw. der damit verbundenen sehr kurzen narrativen Aussageabschnitte Bodeneffekte bei den Glaubhaftigkeitsmerkmalen, bei einer relativ kleinen Stichprobe waren von daher signifikante Gruppenunterschiede kaum zu erreichen. In derzeit noch laufenden Doktor- und Diplomarbeiten, die am selben Institut durchgeführt werden, zeichnet sich ebenfalls ab, daß die qualitativen Unterschiede zwischen erlebnisbasierten und suggerierten Schilderungen nach mehrmaligen suggestiven Befragungen bei Erstkläßlern ausgesprochen gering sind und sich Unterschiede im wesentlichen bezüglich des allgemeinen Merkmals „Detailreichtum“, nicht aber bei einzelnen spezifischen Qualitätsmerkmalen finden lassen. In einer Untersuchung von Huffman und Ceci (1997; zit. nach Ruby & Brigham, 1997) unterzogen mit der Methodik vertraute Beurteiler 10 wahre und 10 suggerierte Aussagen von Vorschulkindern merkmalsorientierten Inhaltsanalysen. In den erlebnisbasierten Schilderungen fanden sich mehr Qualitätsmerkmale als in den suggerierten Aussagen, signifikante Mittelwertunterschiede fanden sich bei vier Merkmalen „Logische Konsistenz“, „Raum-zeitliche Verknüpfungen“, „Wiedergabe von Gesprächen“ und „Schilderung ausgefallener Einzelheiten“. Hier waren die Trefferquoten von unkundigen Beurteilern insbesondere bei der Identifizierung von suggerierten Aussagen besser als Zuordnungen auf der Basis der Summe der Qualitätsmerkmale (Huffman & Ceci, 1997; zitiert nach Poole & Lindsay, 1998). Kürzlich untersuchten Bruck et al. (1997), ob sich erlebnisbasierte und suggerierte Schilderungen von Vorschulkindern differenzieren lassen anhand der Menge der berichteten Details, der Anzahl der spontanen Nennungen der Kinder, der Aussagekohäsion (Gebrauch von zeitlichen Markern, Wiedergabe von Gesprächen) und der Aussageelaboration (Verwendung von emotionsbezogenen Ausdrücken, Gebrauch von Adjektiven und Adverbien). Die Autoren fanden, daß sich erlebnisbasierte und induzierte Schilderungen im Laufe von wiederholten Interviews zunehmend angleichen und induzierte Aussagen schließlich sogar mehr deskriptive Elemente enthielten als erlebnisbasierte Schilderungen.

Zusammenfassend ist festzuhalten, daß es – anders als bei der Unterscheidung zwischen erlebnisbasierten und erfundenen Schilderungen – keine konsistenten empirischen Belege dafür gibt, daß sich erlebnisbasierte und suggerierte Aussagen in ihrer Qualität unterscheiden. Die empirischen

Erkenntnisse, die in diesem Bereich vorliegen, lassen sogar eher annehmen, daß sich zumindest in der im Hinblick auf mögliche Suggestionseinflüsse besonders problematischen Gruppe der jungen Kinder solche qualitativen Unterschiede zwischen erlebnisbasierten und suggerierten nicht oder allenfalls in geringem Umfang finden lassen. Eine aussagepsychologische Beurteilung, die auch bei dieser Fragestellung vor allem auf die Aussagequalität abstellt, kann von daher zu falschen Ergebnissen kommen.

Für die Beurteilung von Fällen mit mäßigem oder geringem Suggestionspotential hat Greuel (1997b) kürzlich auf die Relevanz von Eigenständigkeitsmerkmalen (Arntzen, 1993) hingewiesen. Entscheidende Bedeutung kommt dabei den Spontanpräzisierungen und -ergänzungen zu, insbesondere wenn diese die logische Konsistenz und Anschaulichkeit der Aussage erhöhen, wenn sie widerspruchsfrei in das bisherige Aussagematerial integriert werden können, wenn sie in beiläufiger Form vorgebracht werden, wenn sie bei unsystematischer Befragung in individualtypischem Erinnerungstempo vorgebracht und/oder weitergeführt werden können oder wenn sie ihrerseits qualifizierte Merkmale einer erlebnisfundierte Aussage aufweisen (Greuel, 1997b, S. 218, vgl. Arntzen, 1993).

Greuel (1997 a, b) argumentiert, daß in Fällen sehr intensiver Suggestion die Anwendung der inhaltsorientierten Methodik gar nicht mehr erfolgen kann. Dieser Auffassung ist insoweit zuzustimmen, daß ein positiver Beleg eines tatsächlichen Erlebnisbezugs nach sehr intensiven suggestionsfördernden Bedingungen mit inhaltsanalytischer Methodik kaum noch gefunden werden kann (vgl. Steller & Volbert, 1997). Umgekehrt lassen sich aber häufig Elemente in einer Aussage oder in der Aussagegeschichte finden, die nicht nur auf potentiell suggestive Wirkung verweisen, sondern aufgrund derer Suggestionseffekte konkret anzunehmen sind. Nicht zuletzt haben die spektakulären Mißbrauchsfälle der letzten Jahre, bei denen die Frage einer möglicherweise sehr intensiven suggestiven Beeinflussung zu prüfen war, gezeigt, daß aussagepsychologischer Sachverstand erheblich zur Klärung beitragen konnte (Köhnken, 1997; Steller, 1998).

4 Zur Kritik forensisch-aussagepsychologischer Begutachtungen in Strafverfahren

Psychologische Glaubwürdigkeitsbegutachtungen sind in der letzten Zeit von verschiedenen Seiten kritisiert worden. Fischer (1994) geht von einem Überhandnehmen von Glaubwürdigkeitsbegutachtungen in Strafverfahren wegen Verstößen gegen die sexuelle Selbstbestimmung aus und fordert ein Zurückdrängen von Glaubwürdigkeitsgutachtern aus der Beweiswürdigung und eine Stärkung des Verantwortungsbewußtseins gegenüber der „ureigenen Aufgabe“ von Strafrichtern.

Es erscheint bereits fraglich, ob von einem Überhandnehmen von Glaubwürdigkeitsbegutachtungen ausgegangen werden kann: In einer Totalerhebung der bei der Staatsanwaltschaft Berlin im Jahre 1991 registrierten Fälle von sexuellem Kindesmißbrauch (etwa 1.000) wurden 34 Glaubwürdigkeitsbegutachtungen gezählt. Es kam zu 178 Hauptverhandlungen, in denen 28 Glaubwürdigkeitsgutachten erstattet wurden (16 %; Busse & Volbert, 1997). Bei einer anderen Berliner Stichprobe im Erhebungszeitraum Mai 1995 bis Mai 1996 wurde über 18 von insgesamt 68 Opferzeugen eines sexuellen Mißbrauchsdelikts, die zu einer Hauptverhandlung geladen wurden, ein Glaubwürdigkeitsgutachten erstattet (26 %; Busse, Volbert & Steller, 1996). Von einem Überhandnehmen forensisch-aussagepsychologischer Begutachtungen kann also mindestens für diesen Gerichtsbezirk nicht die Rede sein. Auch in einer bundesweiten Befragung von Richtern und Staatsanwälten (Volbert & Erdmann, 1996) gab die Mehrheit der Befragten an, nur in besonderen Fällen Glaubwürdigkeitsgutachten in Auftrag zu geben.

Ebenso ist zu hinterfragen, ob die Forderung nach einem Zurückdrängen von Glaubwürdigkeitsgutachten aus der Beweiswürdigung die aussagepsychologische Praxis angemessen aufgreift. Es ist eben nicht so, daß von aussagepsychologischen Gutachtern „ohne weiteres die Gesamtheit sonstiger Beweismittel herangezogen wird“ (a.a.O., S. 3). Auf die Grenzen der sachverständigen Befunderhebungen wurde in Abschnitt 2.2 hingewiesen, die Schlußfolgerungen von aussagepsychologischen Sachverständigen bewegen sich ausschließlich im Rahmen psychologischer Analysen. Insofern ist der aussagepsychologische Sachverständige Beweismittel, er stellt keine Beweiswürdigung an. Vereinzelt oder gegebenenfalls auch häufige Fehler in der Praxis sprechen nicht gegen die prinzipielle Aufgabenteilung. Wie bei anderen Gutachten so ist auch bei Glaubwürdigkeitsbegutachtungen eine strikte Trennung der sachverständigen Befunderhebungen und -interpretationen von der richterlichen Beweiswürdigung möglich, wenn beide Seiten ihre Sachgebiete beherrschen.

Fischer selbst (a.a.O., S. 5) betont ausdrücklich den Wert aussagepsychologischer Forschungserkenntnisse für die forensische Praxis. Er fordert eine „stärkere Beschäftigung der Strafrichter mit den Forschungsergebnissen dieser Wissenschaft“ und beklagt, daß Gerichte Glaubwürdigkeit „nach wie vor weithin *intuitiv*“ beurteilen. Natürlich ist es Sache einzelner Gerichte, zu entscheiden, ob sie sich aussagepsychologische Erkenntnisse per Fortbildung oder per Sachverständigenbeweis zunutze machen. Es erscheint auch keineswegs negativ, wenn Gerichte durch (wiederholte) Glaubwürdigkeitsgutachten „lernen“ und Glaubwürdigkeitsgutachter sich somit für diese Gerichte zunehmend überflüssig machen. Wenn Fischers Klage über die weitgehende Unkenntnis von Richtern über aussagepsychologische Grundsätze aber zutrifft, so erscheint ein Verzicht von Glaubwürdigkeitsbegutach-

tungen im Strafprozeß jedenfalls nicht zur Optimierung der Wahrheitsfindung geeignet.

Die Strafrechtskommission des Deutschen Juristinnenbundes hat einen Gesetzentwurf zur „Reform der Nebenklage und anderer Verletztenrechte“ erarbeitet, in dem – auf dem Hintergrund kindgerechter Rahmenbedingungen für die Vernehmungen von Kindern – „u. a. die Glaubwürdigkeitsbegutachtung geregelt werden soll“. Als „Vorschlag für einen neu einzufügenden § 68 b StPO (Glaubwürdigkeitsgutachten)“ wird (auf S. 49) formuliert: „Gutachten, die der Beurteilung der Glaubwürdigkeit von Zeuginnen und Zeugen dienen, sind nur zulässig, wenn die zu begutachtende Person nach Belehrung einwilligt. Glaubwürdigkeitsgutachten bei Personen unter 16 Jahren sind unzulässig.“ Weiter heißt es (S. 22): „Konsequenz ist, daß auch Gutachten, einschließlich Parteigutachten, ohne Exploration der Aussageperson unzulässig sind. Diese wären zudem zur Aufklärung generell ungeeignete Beweismittel.“

Die Begründung des Entwurfes zielt vor allem auf den Kinderschutz ab. Einen Verzicht auf aussagepsychologische Begutachtungen legen kinderschutzorientierte Überlegungen jedoch nicht nahe. Nicht zuletzt ist das Ausmaß der Belastung von Kindern durch Strafverfahren auch abhängig von der Kompetenz der Beteiligten in bezug auf adäquate Interview- und Beurteilungsverfahren. So führt die Kenntnis solcher Methoden beispielsweise zu einer Vermeidung von überflüssigen – weil für die Klärung des Sachverhalts und für die Beurteilung der Glaubhaftigkeit der Aussage nicht relevanten – Fragen an das Kind. Um die Belastung der Kinder möglichst gering zu halten, könnten aussagepsychologische Begutachtungen im Zusammenhang mit der Konservierung von Erstvernehmungen erfolgen. Hier ist darauf hinzuweisen, daß eine dem Kinderschutz dienende Verwendung von Videodokumentationen nur dann zu erwarten ist, wenn die Vernehmungen bzw. Explorationen bei den dokumentierten Befragungen sachgerecht erfolgt sind. Auch dazu kann ein Beitrag durch aussagepsychologische Sachverständige geleistet werden.

Burgsmüller (1997) bezeichnet den Entwurf des Deutschen Juristinnenbundes als „Gegenreaktion“ auf die Tätigkeiten aussagepsychologischer Gutachter in den sogenannten Wormser Verfahren vor dem Landgericht Mainz. Insbesondere kritisiert sie die Begutachtungen ohne Exploration der kindlichen Zeugen zu Person und Sache. Hierzu ist anzumerken, daß die Tätigkeit der vom Gericht bestellten aussagepsychologischen Zweitgutachter in den Verfahren Worms II und III vor dem Landgericht Mainz nach der Prozeßlage vordringlich in der Analyse der Entstehungsbedingungen und der weiteren Entwicklung der kindlichen Zeugenaussagen bestand.²⁷ Diese

²⁷ Es handelte sich eben um einen Sachverständigenauftrag der Art „Beurteilung von Tatsachen“ (ohne eigene Tatsachenfeststellungen) – vgl. auch Abschnitt 2.2.

konnte ohne (eigene) Exploration der Kinder erfolgen, da das notwendige Tatsachenmaterial in den Akten vorlag und in den Hauptverhandlungen reproduziert wurde. Eine (nochmalige) Befragung der Kinder nach deren zum Teil mehrmonatiger suggestiver Beeinflussung war nicht nur überflüssig, sondern methodisch und ethisch kontraindiziert.

Sowohl die „konservative“ als auch die „radikale“ juristische Kritik an Glaubwürdigkeitsbegutachtungen gehen zum Teil von falschen Voraussetzungen aus (Glaubwürdigkeitsbegutachtung als „Routinevorgang“). Sie unterscheiden nicht zwischen sogenannten Glaubwürdigkeitsgutachten mit eklektisch-intuitivem Vorgehen (vgl. Steller, 1998) und forensisch-aussagepsychologischen Begutachtungen mit Verwendung der hier beschriebenen Logik und Methodik und berücksichtigen nicht ausreichend den derzeitigen Entwicklungsstand der Methodik. Bei einem Verzicht auf einzelfallbezogene forensisch-aussagepsychologische Begutachtungen ist zu befürchten, daß neue Entwicklungen oder Problemstellungen der Aussagepsychologie nicht rechtzeitig von der forensischen Praxis wahrgenommen werden. Zum Beispiel kam es Anfang der neunziger Jahre im gerichtlichen Umgang mit dem Verdacht auf sexuellen Kindesmißbrauch sowohl im Familien- als auch im Strafrecht zu erheblichen Verunsicherungen. Diese waren auch dadurch begründet, daß Gerichte und einige Gutachter nicht ausreichend auf die Beurteilung von sexuellem Mißbrauchsverdacht nach suggestiver Aufdeckungsarbeit mit Kindern vorbereitet waren. Massiven Fehlentwicklungen konnte allerdings dadurch vorgebeugt werden, daß der vorhandene Sachverstand auch per Sachverständigenbeweis in Gerichtsverfahren eingeführt wurde. Folgt man den Überlegungen des Deutschen Juristinnenbundes, verbietet aussagepsychologische Begutachtungen, verzichtet weitgehend auf eine Mitwirkung von Kindern im Strafverfahren und greift vor allem auf mittelbare Beweismittel zurück (Nelles, 1997), könnte dies erneut dazu führen, daß vermehrt Äußerungen und Verhaltensweisen von Kindern im Hinblick auf deren Bedeutung für die Feststellung von sexuellem Kindesmißbrauch vor Gericht vorgetragen werden, die Resultat suggestiver Aufdeckungsarbeit sind. Ein Verbot aussagepsychologischer Begutachtungen würde die Gefahr erhöhen, daß die zugrundeliegenden Suggestionenprozesse nicht ausreichend beleuchtet und falsche Mißbrauchsdiagnosen durch Strafgerichte übernommen würden.

Als Beispiel für eine andere mögliche zukünftige Problematik sei auf das Phänomen der sogenannten wiedererwachten Erinnerungen (recovered memories) von erwachsenen Frauen an sexuelle Mißbrauchserfahrungen als Kind hingewiesen. Es handelt sich um als Erinnerungen erlebte Vorstellungen über bisher verdrängte Erfahrungen (repressed memories), bei denen sich die Frage nach möglicher Induktion von Pseudoerinnerungen durch Elemente der Therapie oder durch Autosuggestion stellt. Aus der US-amerikanischen forensischen Praxis wurden nicht nur spektakuläre Fälle berichtet, die Häu-

figkeit der entsprechenden Fälle erscheint ebenfalls enorm (vgl. u. a. Loftus, 1997). Vereinzelt Gutachtaufträge signalisieren, daß offenbar auch deutsche Strafgerichte mit der Problematik konfrontiert werden. Der Umfang der (kontroversen) wissenschaftlichen Literatur zum Thema ist erheblich. Auch hier handelt es sich vorwiegend um gedächtnispsychologische (d. h. aussagepsychologische) Problemstellungen, bei deren sachverständiger Bearbeitung neben den allgemeinen aussagepsychologischen Kenntnissen das Spezialwissen zum Thema „repressed memories“ nötig ist.

Ein Zurückdrängen aussagepsychologischer Sachkunde aus Gerichten erscheint daher auch im Hinblick auf mögliche neue Problemkonstellationen nicht indiziert, da eine Interaktion zwischen forensischer Inanspruchnahme und wissenschaftlicher Entwicklung der Aussagepsychologie besteht. Die von einem Pionier der angewandten Psychologie, dem von Deutschland nach Harvard berufenen Professor Hugo Münsterberg (1914, S. 515), ausgesprochene Empfehlung, daß „Psychologie und Jurisprudenz ohne wechselseitiges Mißtrauen wirklich in Fühlung miteinander vorwärts arbeiten“, gilt weiterhin.

Literatur

- Amthauer, R., Brocke, B. & Liepmann, D. (1998). *Intelligenz-Struktur-Test 2000 (I-S-T 2000)*. Göttingen: Hogrefe.
- Anson, D. A., Golding, S. L. & Gully, K. J. (1993). Child sexual abuse allegations: Reliability of criteria-based content analysis. *Law and Human Behavior, 17*, 331-342.
- Arntzen, F. (1970/1983/1993). *Psychologie der Zeugenaussage. System der Glaubwürdigkeitsmerkmale*. München: Beck.
- Bekerian, D. A. & Dennett, J. L. (1995). Assessing the truth in children's statements. In T. Ney (Ed.), *True and false allegations of child sexual abuse. Assessment and case management* (pp. 263-175). New York: Brunner/Mazel.
- Berufsverband Deutscher Psychologen e. V. (1988). *Richtlinien für die Erstellung Psychologischer Gutachten*. Bonn: Deutscher Psychologen Verlag.
- Boyчук, T. D. (1991). *Criteria-based content analysis of children's statements about sexual abuse: A field-based validation study*. Unpublished doctoral dissertation, Arizona State University.
- Bruck, M., Ceci, S., Francoeur, E. & Barr, R. (1995). „I hardly cried when I got my shot!“ Influencing children's reports about a visit to their pediatrician. *Child Development, 66*, 193-208.
- Bruck, M., Ceci, S. J. & Hembrooke, H. (1998). Reliability and credibility of young children's reports. From research to policy and practice. *American Psychologist, 53*, 136-151.
- Bruck, M., Hembrooke, H. & Ceci, S. J. (1997). Children's reports of pleasant and unpleasant events. In D. Read & D. S. Lindsay (Eds.), *Re-*

collections of trauma. Scientific evidence and clinical practice (NATO ASI Series. Series A: Life Sciences Vol. 291) (pp. 199-219). New York: Plenum.

Burgsmüller, C. (1997). *Plädoyer für eine konsequente anwaltliche Vertretung sexuell mißbrauchter Kinder - Schutz vor Grenzüberschreitungen durch BeraterInnen, GutachterInnen und Vernehmungspersonen*. Abstract eines Vortrags auf der 7. Tagung der Fachgruppe Rechtspsychologie in der Deutschen Gesellschaft für Psychologie in Kiel vom 09. bis 11. Oktober 1997.

Busse, D. & Volbert, R. (1997). Glaubwürdigkeitsgutachten in Strafverfahren wegen sexuellen Mißbrauchs - Ergebnisse einer Gutachtenanalyse. In L. Greuel, T. Fabian & M. Stadler (Hrsg.), *Psychologie der Zeugenaussage* (S. 131-142). Weinheim: Beltz/Psychologie Verlags Union.

Busse, D., Volbert, R. & Steller, M. (1996). *Belastungserleben von Kindern in Hauptverhandlungen*. (Reihe Recht, herausgegeben vom Bundesministerium der Justiz). Bonn.

Ceci, S. J. & Bruck, M. (1993). The suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, 113, 403-439.

Ceci, S. J. & Bruck, M. (1995). *Jeopardy in the courtroom*. Washington, DC: American Psychological Association.

Ceci, S. J., Huffman, M. L. C., Smith, E. & Loftus, E. W. (1994). Repeatedly thinking about a non-event: Source misattributions among preschoolers. *Consciousness and Cognition*, 3, 388-407.

Ceci, S. J., Loftus, E. W., Leichtman, M. D. & Bruck, M. (1994). The possible role of source misattributions in the creation of false beliefs among preschoolers. *International Journal of Clinical and Experimental Hypnosis*, 42, 304-320.

Dahle, K.-P. (1997a). Wege zu einem linguistischen Wahrheitstest? Perspektiven einer einzelfallexperimentellen Weiterentwicklung der Kriterienorientierten Aussageanalyse für die forensisch-psychologische Glaubwürdigkeitsdiagnostik. *Diagnostica*, 43, 3-26.

Dahle, K.-P. (1997b). Zur potentiellen Validität der Kriterienorientierten Aussageanalyse. Unveröffentl. Manuskript zum Vortrag auf der 7. Arbeitstagung der Fachgruppe Rechtspsychologie der Deutschen Gesellschaft für Psychologie in Kiel vom 9. bis 11. Oktober 1997.

Dahle, K.-P. & Wolf, P. (1997). Realkennzeichen im Einzelfall. In L. Greuel, T. Fabian & M. Stadler (Hrsg.), *Psychologie der Zeugenaussage*. Weinheim: Psychologie Verlags Union.

Deckers, R. (1996). Verteidigung in Verfahren wegen sexuellen Mißbrauchs von Kindern. *Neue Juristische Wochenschrift*, 47, 3105-3111.

Dettenborn, H., Fröhlich, H.-H., & Szewczyk, H. (1984). *Forensische Psychologie. Ein Lehrbuch der gerichtlichen Psychologie für Juristen, Kriminalisten, Psychologen, Pädagogen und Mediziner*. Berlin: VEB Deutscher Verlag der Wissenschaften.

- Deutscher Juristinnenbund (Strafrechtskommission)(ohne Jahr). *Reform der Nebenklage und anderer Verletztenrechte*. Unveröffentlichtes Papier.
- Fischer, T. (1994.) Glaubwürdigkeitsbeurteilung und Beweiswürdigung – Von der Last der „ureigenen Aufgabe“ –. *Neue Zeitschrift für Strafrecht*, 1, 1-56.
- Fürniss, T. (1991). *The multi-professional handbook of child sexual abuse. Integrated management, therapy, and legal intervention*. London: Routledge.
- Greuel, L. (1997a). Glaubwürdigkeit - Zur psychologischen Differenzierung eines umgangssprachlichen Konstrukts. *Praxis der Rechtspsychologie*, 7, 154-169.
- Greuel, L. (1997b). Suggestibilität und Aussagezuverlässigkeit: ein (neues) Problem in der forensisch-psychologischen Praxis? In L. Greuel, T. Fabian & M. Stadler (Hrsg.), *Psychologie der Zeugenaussage* (S. 211-220). Weinheim: Psychologie Verlags Union.
- Greuel, L., Offe, S., Fabian, A., Wetzels, P., Fabian, T., Offe, H. & Stadler, M. (1998). *Glaubhaftigkeit der Zeugenaussage*. Weinheim: Psychologie Verlags Union.
- Hershkowitz, I., Lamb, M. E., Sternberg, K. J. & Esplin, P. W. (1997). The relationships among interviewer utterance type, CBCA cores, and the richness of children's responses. *Legal and Criminological Psychology*, 2, 169-176.
- Hommers, W. (1997). Die aussagepsychologische Kriteriologie unter kovarianzstatistischer und psychometrischer Perspektive. In L. Greuel, T. Fabian & M. Stadler (Hrsg.), *Psychologie der Zeugenaussage* (S. 87-100). Weinheim: Psychologie Verlags Union.
- Horowitz, S. W. (1991). Empirical support for statement validity analysis. *Behavioral Assessment*, 13, 293-313.
- Horowitz, S. W., Lamb, M. E., Esplin, P. W., Boychuk, T. D., Krispin, O. & Reiter-Lavery, L. (1997). Reliability of criteria-based content analysis of child witness statements. *Legal and Criminological Psychology*, 2, 11-21.
- Huffman, M. L. & Ceci, S. J. (1997). *Can criteria-based content analysis distinguish true and false beliefs of preschoolers? An exploratory analysis*. Unpublished manuscript, Cornell University, Ithaca, NY.
- Hustedt, V. (1998). Erweiterung der Einzelfallexperimentellen Kriterienorientierten Aussageanalyse für forensisch-psychologische Glaubwürdigkeitsdiagnostik durch sprachformale und andere potentiell trennfähige Kriterien. Unveröffentl. Diplomarbeit, Technische Universität Braunschweig.
- Jäger, R. S. (1983). *Der diagnostische Prozeß. Eine Diskussion psychologischer und methodischer Randbedingungen*. Göttingen. Hogrefe.
- Jessnitzer, K. & Frieling, G. (1992, 10. Aufl.). *Der gerichtliche Sachverständige*. Köln: Carl Heymanns Verlag.
- Johnson, M. K., Hashtroudi, S. & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3-28.

- Johnson, M. K. & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, 88, 67-85.
- Kendall-Tackett, K. A., Williams, L. M. & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin*, 113, 164-180.
- Köhnken, G. (1982). *Sprechverhalten und Glaubwürdigkeit. Eine experimentelle Studie zur extralinguistischen und textstilistischen Aussageanalyse*. Unveröffentlichte Dissertation, Universität Kiel.
- Köhnken, G. (1990). *Glaubwürdigkeit. Untersuchungen zu einem psychologischen Konstrukt*. München: Psychologie Verlags Union.
- Köhnken, G. (1997). Suggestive Prozesse in Zeugenbefragungen: Formen und theoretische Erklärungsansätze. *Monatsschrift für Kriminologie und Strafrechtsreform*, 80, 290-299.
- Krahé, B. & Kundrotas, S. (1992). Glaubwürdigkeitsbeurteilung bei Vergewaltigungsanzeigen: Ein aussageanalytisches Feldexperiment. *Zeitschrift für experimentelle und angewandte Psychologie*, 39 (4), 598-620.
- Lamb, M. E., Sternberg, K. J., Esplin, P. W., Hershkowitz, I. & Orbach, Y. (1997). Assessing the credibility of children's allegations of sexual abuse: A survey of recent research. *Learning and Individual Differences*, 9, 175-194.
- Lamb, M. E., Sternberg, K., Esplin, P. W., Hershkowitz, I., Orbach, Y. & Hovav, M. (1997). Criterion-based content analysis: a field validation study. *Child Abuse & Neglect*, 21, 255-264.
- Lamers-Winkelmann, F. & Buffing, F. (1996). Children's testimony in the Netherlands. A study of statement validity analysis. In B. L. Bottoms & G. S. Goodman (Eds.), *International perspectives on child abuse and children's testimony. Psychological research and law* (pp. 45-61). Thousand Oaks: Sage.
- Landry, K. L. & Brigham, J. C. (1992). The effect of training in criteria-based content analysis on the ability to detect deception in adults. *Law and Human Behavior*, 16, 663-676.
- Lefferenz, H. (1972). Die Beurteilung der Glaubwürdigkeit. In H. Göppinger & H. Witter (Hrsg.), *Handbuch der Forensischen Psychiatrie. Band II* (S. 1314-1345). Berlin: Springer.
- Lehr, S. (1995). *Qualitativer Vergleich von instruierten, suggerierten und erlebnisbegründeten Kinderaussagen*. Unveröffentlichte Diplomarbeit, FU Berlin.
- Leichtman, M. D. & Ceci, S. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology*, 31, 568-578.
- Lempp, R. (1989). Begutachtung. In C. Eggers, R. Lempp, G. Nissen & P. Strunk (Hrsg.), *Kinder- und Jugendpsychiatrie* (5. Aufl.) (S. 594-631). Berlin: Springer.
- Lepore, S. J. & SESCO, B. (1994). Distorting children's reports and interpretations of events through suggestions. *Applied Psychology*, 79, 108-120.
- Littmann, E. & Szewczyk, H. (1983). Zu einigen Kriterien und Ergebnissen forensisch-psychologischer Glaubwürdigkeitsbegutachtung von sexuell

- mißbrauchten Kindern und Jugendlichen. *Forensia*, 4, 55-72.
- Loftus, E. F. (1997). Dispatch from the (un)civil memory wars. In J. D. Read & D. S. Lindsay (Eds.), *Recollections of trauma. Scientific evidence and clinical practice* (NATO ASI Series. Series A: Life Sciences Vol. 291) (pp. 171-194). New York: Plenum Press.
- McCloskey, M. & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against the memory impairment hypothesis. *Journal of Experimental Psychology: General*, 114, 1-16.
- Münsterberg, H. (1914). *Grundzüge der Psychotechnik*. Leipzig: Barth.
- Nelles, U. (1997). Persönlichkeitsrechte und Pflichten kindlicher Zeugen im Strafprozeß. *STREIT*, 15, 99-110.
- Ornstein, P., Gordon, B. N. & Larus, D. (1992). Children's memory for a personally experienced event: Implications for testimony. *Applied Cognitive Psychology*, 6, 49-60.
- Pezdek, K. & Roe, C. (1995). The effect of memory trace strength on suggestibility. *Journal of Experimental Child Psychology*, 60, 116-128.
- Plaum, E. (1998). Ein „Wahrheitstest“ für Glaubhaftigkeitsgutachten. *Kriminalistik*, 8-9 (52), 549-557.
- Poole, D. A. & Lindsay, D. S. (1995). Interviewing preschoolers: Effects of nonsuggestive techniques, parental coaching, and leading questions on reports of nonexperienced events. *Journal of Experimental Child Psychology*, 60, 129-154.
- Poole, D. A. & Lindsay, D. S. (1998). Assessing the accuracy of young children's reports: Lessons from the investigation of child sexual abuse. *Applied & Preventive Psychology*, 7, 1-26.
- Raskin, D. C. & Esplin, P. W. (1991a). Statement validity assessments: Interview procedures and content analyses of children's statements of sexual abuse. *Behavioral Assessment*, 13, 265-291.
- Raskin, D. C. & Esplin, P. W. (1991b). Assessment of children's statements of sexual abuse. In J. Doris (Ed.), *The suggestibility of children's recollections* (pp. 153-164). Washington, DC: American Psychological Association.
- Ruby, C. L. & Brigham, J. C. (1997). The usefulness of the criteria-based content analysis technique in distinguishing between truthful and fabricated allegations: A critical review. *Psychology, Public Policy, and Law*, 3, 705-737.
- Saacke, I. (1995). Aussagequalität und kognitiver Entwicklungsstand. Unveröffentlichte Diplomarbeit, FU Berlin.
- Scholz, O. B. & Endres, J. (1995). Aufgaben des psychologischen Sachverständigen beim Verdacht des sexuellen Kindesmißbrauchs - Befunde, Diagnostik, Begutachtung -. *Neue Zeitschrift für Strafrecht*, 15 (1), 6-12.
- Schooler, J. W., Clark, C. & Loftus, E. F. (1988). Knowing when memory is real. In M. M. Gruneberg, P. E. Morris & R. N. Sykes (Eds.), *Practical*

- aspects of memory: current research and issues* (vol. 1, pp. 83-88). Chichester: Wiley.
- Schooler, J. W., Gerhard, D. & Loftus, E. F. (1986). Qualities of the unreal. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12 (2), 171-181.
- Sporer, S. L. (1997a). Realitätsüberwachungskriterien und forensische Glaubwürdigkeitskriterien im Vergleich: Validitätsüberprüfung anhand selbsterlebter und erfundener Geschichten. In L. Greuel, T. Fabian & M. Stadler (Hrsg.), *Psychologie der Zeugenaussage* (S. 71-85). Weinheim: Psychologie Verlags Union.
- Sporer, S. L. (1997b). The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology*, 11, 373-397.
- Sporer, S. L. & Küpper, B. (1995). Realitätsüberwachung und die Beurteilung des Wahrheitsgehaltes von Erzählungen: Eine experimentelle Studie. *Zeitschrift für Sozialpsychologie*, 173-193.
- Steller, M. (1988a). Standards der forensisch-psychologischen Begutachtung. *Monatsschrift für Kriminologie und Strafrechtsreform*, 71, 16-27.
- Steller, M. (1988b). Die vierte Phase der Aussagepsychologie. *Forensia*, 9, 23-28.
- Steller, M. (1989). Recent development in statement analysis. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 135-154). Dordrecht: Kluwer.
- Steller, M. (1994). Diagnostischer Prozeß. In R.-D. Stieglitz & U. Baumann (Hrsg.), *Psychodiagnostik psychischer Störungen* (S. 37-46). Stuttgart: Ferdinand Enke Verlag.
- Steller, M. (1998). Aussagepsychologie vor Gericht - Methodik und Probleme von Glaubwürdigkeitsgutachten mit Hinweisen auf die Wormser Mißbrauchsprozesse. *Recht & Psychiatrie*, 16, 11-18.
- Steller, M. (in Vorb.). Psychologische Diagnostik - Menschenkenntnis oder angewandte Wissenschaft? In H.-L. Kröber (Hrsg.), *Psychologische Begutachtung im Strafverfahren - Indikationen und Qualitätsstandards*. Lübeck: Römhild-Verlag.
- Steller, M. & Boychuk, T. (1992). Children as witnesses in sexual abuse cases: Investigative interview and assessment techniques. In H. Dent & R. Flin (Eds.), *Children as witnesses* (pp. 47-71). Chichester: John Wiley & Sons.
- Steller, M. & Köhnken, G. (1989). Criteria-based statement analysis. Credibility assessment of children's statements in sexual abuse cases. In D. C. Raskin (Ed.), *Psychological methods for investigation and evidence* (pp. 217-245). New York: Springer.
- Steller, M. & Volbert, R. (1997). Glaubwürdigkeitsbegutachtung. In M. Steller & R. Volbert (Hrsg.), *Psychologie im Strafverfahren* (S. 12-39). Bern: Huber.
- Steller, M., Volbert, R. & Wellershaus, P. (1993). Zur Beurteilung von Zeugenaussagen: Aussagepsychologische Konstrukte und methodische Strate-

- gien. In L. Montada (Hrsg.), *Bericht über den 38. Kongreß der Deutschen Gesellschaft für Psychologie in Trier 1992* (S. 367-376). Göttingen: Hogrefe.
- Steller, M., Wellershaus, P. & Wolf, T. (1992). Realkennzeichen in Kinderaussagen: Empirische Grundlagen der Kriterienorientierten Aussageanalyse. *Zeitschrift für experimentelle und angewandte Psychologie*, 39, 151-170.
- Stern, W. (1904). Die Aussage als geistige Leistung und als Verhörprodukt: Experimentelle Schüleruntersuchungen. In W. Stern (Hrsg.), *Beiträge zur Psychologie der Aussage* (Heft 3). Leipzig: Barth.
- Stieglitz, R.-D. & Baumann, U. (Hrsg.) (1994). *Psychodiagnostik psychischer Störungen*. Stuttgart: Ferdinand Enke Verlag.
- Szewczyk, H. & Littmann, E. (1982). Untersuchungen zur Glaubwürdigkeitsbeurteilung kindlicher Zeugen. In A. Trankell (Ed.), *Reconstructing the past. The role of psychologists in criminal trials* (pp. 73-103). Stockholm: Norstedt & Söners.
- Tewes, U. (Hrsg.) (1994, 2. korr. Auflage). *Hamburg-Wechsler-Intelligenztest für Erwachsene - Revision 1991*. Bern: Verlag Hans Huber.
- Trankell, A. (1971). *Der Realitätsgehalt von Zeugenaussagen*. Göttingen: Hogrefe.
- Tully, B. (1998). Reliability of criteria-based content analysis of child witness statements: Cohen's kappa doesn't matter. *Legal and Criminological Psychology*, 3, 183-188.
- Undeutsch, U. (1967). Beurteilung der Glaubhaftigkeit von Zeugenaussagen. In U. Undeutsch (Hrsg.), *Handbuch der Psychologie, Bd. 11, Forensische Psychologie* (S. 26-181). Göttingen: Hogrefe.
- Undeutsch, U. (1982). Statement reality analysis. In A. Trankell (Ed.), *Reconstructing the past. The role of psychologists in criminal trials* (pp. 27-56). Stockholm: Norstedt & Söners.
- Volbert, R. (1997). Suggestibilität kindlicher Zeugen. In M. Steller & R. Volbert (Hrsg.), *Psychologie im Strafverfahren* (S. 40-62). Bern: Huber.
- Volbert, R. (im Druck). Standards der psychologischen Glaubhaftigkeitsdiagnostik. In H.-L. Kröber (Hrsg.), *Psychologische Begutachtung im Strafverfahren - Indikationen und Qualitätsstandards*. Lübeck: Römhild-Verlag.
- Volbert, R. & Erdmann, K. (1996). Kinder als Zeugen in Strafverfahren wegen sexuellen Mißbrauchs: Einstellungen und Erfahrungen von Richtern und Staatsanwälten. *Monatsschrift für Kriminologie und Strafrechtsreform*, 79, 238-252.
- Volbert, R. & Pieters, V. (1996). Suggestive Beeinflussungen von Kinderaussagen. *Psychologische Rundschau*, 47, 183-198.
- Volbert, R. & Pieters, V. (1997). Instructions and suggestions: Effects on the amount of details in children's statements. In S. Redondo, V. Garrido, J. Pérez, J. Bajet & R. M. Martínez (Eds.), *Advances in psychology and law. International contributions* (pp. 136-146). Berlin: de Gruyter.

- Wegener, H. (1997). Die Entwicklung der experimentellen Rechtspsychologie. In L. Greuel, T. Fabian & M. Stadler (Hrsg.), *Psychologie der Zeugenaussage* (S. 13-21). Weinheim: Psychologie Verlags Union.
- Wellershaus, P. & Wolf, T. (1989). *Kriterienorientierte Aussageanalyse. Empirische Überprüfung der diagnostischen Leistungsfähigkeit inhaltlicher Realkennzeichen zur Trennung wahrer von unwahren Kinderaussagen*. Unveröffentlichte Diplomarbeit, Universität Kiel.
- Wells, G. L. & Loftus, E. F. (1991). Commentary: Is this child fabricating? Reactions to a new assessment technique. In J. Doris (ed.), *The suggestibility of children's recollections* (pp. 168-171). Washington, DC: American Psychological Association.
- Wolf, P. & Steller, M. (1997). Realkennzeichen in Aussagen von Frauen. Zur Validierung der Kriterienorientierten Aussageanalyse für Zeugenaussagen von Vergewaltigungsoffern. In L. Greuel, T. Fabian & M. Stadler (Hrsg.), *Psychologie der Zeugenaussage* (S. 122-130). Weinheim: Psychologie Verlags Union.
- Zaparniuk, J., Yuille, J. C. & Taylor, S. (1995). Assessing the credibility of true and false statements. *International Journal of Law and Psychiatry*, 18, 343-352.
- Zaragoza, M. S., Dahlgreen, D. & Muench, J. (1992). The role of memory impairment in children's suggestibility. In M. L. Howe, C. J. Brainerd & V. F. Reyna (Eds.), *Development of long-term retention* (pp. 184-216). New York: Springer.

Zusammenfassende Beantwortung des Fragenkatalogs des BGH

Im folgenden werden die explizit ausformulierten Fragen des 1. Strafsenats des BGH, die zusammen mit dem Gutachtauftrag übersandt wurden, zusammenfassend beantwortet (= kursiver Text). Ausführliche Begründungen und wissenschaftliche Belege für diese Antworten ergeben sich aus den entsprechenden Ausführungen im Gutachten.

1. Fragen zur Methodik bei Erstellung eines psychologischen Glaubwürdigkeitsgutachtens

1.1 Ist es erforderlich, angewandte Testverfahren hinsichtlich ihrer Indikation und Methodik zu erklären?

Nein. Diese Angaben sind in Testmanualen bzw. Sekundärliteratur enthalten. Ihre regelmäßige Mitteilung im schriftlichen Gutachten würde einen erheblichen Umfang benötigen. Kurze Erwähnungen über den Stellenwert der jeweils verwendeten Verfahren im Rahmen der hypothesengeleiteten Datenerhebung reichen nicht nur aus, sondern erfüllen das Transparenzgebot besser als Testdarstellungen, die nicht explizit auf den Begutachtungsfall bezogen sind.

1.2 Ist es erforderlich, die Ergebnisse der Testverfahren mitzuteilen (wenn ja, in welchem Umfang), oder genügt es, die Befunde interpretativ zu beschreiben?

Unter dem Transparenzgebot im Sinne der Nachvollziehbarkeit von Gutachten erscheint eine (interpretierende) Verbalisierung von Testergebnissen sinnvoller als die Angabe von Skalenwerten, die von psychologischen Laien nicht ausreichend verstanden werden können. Unter dem weiteren Aspekt von Transparenz, nämlich der Überprüfbarkeit der Tätigkeit von Sachverständigen, hat die Mitteilung von Zahlenwerten nur einen sehr eingeschränkten Erkenntniswert: Sie ermöglicht das Auffinden eventueller basaler Fehler mit wahrscheinlich allenfalls marginalem Stellenwert im Kontext aussagepsychologischer Gutachten.

2.1. Ist es ein methodischer Mangel, wenn die Sachverständige kein Wortprotokoll des Aussageberichts und der Befragung der Zeugin zur Sache erstellt?

2.2 Müßte ein solches Protokoll im Gutachten mitgeteilt werden?

Die Fragen 2.1 und 2.2 werden im Zusammenhang beantwortet. Sie unterscheiden zwischen dem Material der Befunderhebung und -bearbeitung und deren Mitteilung im schriftlichen Gutachten. Während Tonaufnahmen von Explorationen für die Analysetätigkeit von Sachverständigen unerlässlich erscheinen, kann diese aber ohne Wortprotokolle, nämlich durch Abhören der Tonaufnahmen, erfolgen. Die Mitteilung der Explorationsinhalte kann je nach Einzelfall in sehr unterschiedlicher

Form erfolgen, nämlich in allen denkbaren Varianten zwischen vollständigem Wortprotokoll und zusammenfassendem Explorationsbericht, gegebenenfalls mit einer mehr oder weniger umfangreichen Auswahl relevanter Zitate. Unter dem Aspekt der Nachvollziehbarkeit erscheint der sachverständig aufbereitete Explorationsbericht sinnvoller als reine Wortprotokolle. Hat der Sachverständige allerdings ein Wortprotokoll für die eigene Analysetätigkeit erstellt, so steht nichts dagegen, dieses Protokoll ins Gutachten ein- oder als Anlage beizufügen. Denkbar ist in Einzelfällen auch eine Kombination von zusammenfassendem Explorationsbericht mit Herausarbeitung der wesentlichen Inhalte der Exploration als Teil des Gutachtens mit der Beifügung eines Wortprotokolls als Anlage.

Wenn es im wesentlichen um die Nachprüfbarkeit im Sinne der Kontrolle von Sachverständigen geht, so könnte die Überprüfung direkt anhand der Tonaufnahme erfolgen bzw. die Anfertigung eines Wortprotokolls auch zu diesem Zwecke vom Gericht nachträglich in Auftrag gegeben werden. Zusätzlich wird darauf hingewiesen, daß eine regelmäßige Anhörung von Tonaufnahmen des Sachverständigen in Hauptverhandlungen nicht nur unter dem Gesichtspunkt der Kontrolle von Sachverständigen diskutiert werden sollte, sondern mit Aspekten des Zeugenschutzes verbunden werden sollte. Die Anhörung von Ton- bzw. Videoaufnahmen anstelle bzw. vor der Entscheidung über die Vernehmungen von Zeugen könnte zu deren Entlastung beitragen und Mehrfachvernehmungen vermeiden helfen. Die isolierte Betonung der Kontrolle von Sachverständigen ist auf dem Hintergrund denkbarer fehlerhafter aussagepsychologischer Gutachten verständlich. Sie erscheint aber auch befördert durch einen nicht gerechtfertigten Transfer der Bedenken gegen suggestive Aufdeckungsmethoden bei Verdacht auf sexuellen Mißbrauch auf die Explorationsmethodik von aussagepsychologischen Sachverständigen.

2.3. Wie wirken sich entsprechende Mängel auf Ergebnisse und Nachprüfbarkeit des Gutachtens aus?

Mängel in der Transparenz von Gutachten schränken die (gedankliche) Nachvollziehbarkeit und Überprüfbarkeit von Gutachten ein. Die umfangreiche Beschreibung von Tests, die Mitteilung von quantifizierten Testergebnissen sowie reine Wortprotokolle der Exploration sind aber nicht die wesentlichen Elemente zur Erfüllung des Transparenzgebotes. Die Fokussierung der Diskussion auf diese Bestandteile könnte dadurch erfolgt sein, daß hier relativ konkrete, um nicht zu sagen „einfache“ Gesichtspunkte angesprochen sind. Die Qualität aussagepsychologischer Gutachten im Sinne des Transparenzgebotes ergibt sich aber nicht daraus, daß Sachverständige ihre basalen Diagnostikschritte umfangreich dokumentieren, sondern vielmehr daraus, daß der komplexe hypothesengeleitete Erhebungs- und Urteilsprozeß der Sachverständigen in schriftlichen (oder mündlichen Gutachten) erkennbar gemacht wird. Eine ele-

mentaristische Sichtweise wird den komplexen Anforderungen von forensisch-aussagepsychologischen Gutachten nicht gerecht. Übergeordnete Gütemaßstäbe solcher Gutachten sind:

- Bezugnahme auf die spezifische Aussage anstelle rein charakterologischer oder motivationsbezogener Überlegungen,
- erkennbare Spezifizierung der globalen Fragestellung „Glaubwürdigkeit“ im Hinblick auf fallbezogene (Unwahr-)Hypothesen,
- hypothesengeleitete Datensammlung anstelle routinemäßiger, gegebenenfalls umfangreicher Testapplikationen,
- Darstellung der gedanklichen Auseinandersetzungen mit den Alternativhypothesen zur Wahrheitsannahme,
- Anwendung einer wissenschaftlich begründeten Methodik: Qualitätsanalyse mit Rückbezug auf aussagerelevante Personenparameter und denkbare Fehlerquellen in der Aussagegenese.

3. Ist es erforderlich, die Dauer der psychologischen Untersuchung festzuhalten und mitzuteilen?

Der ungefähre zeitliche Umfang der gesamten Begutachtung und relevanter Teilschritte (z. B. Exploration) sollte im Gutachten erkennbar sein. Auf jeden Fall sollte mitgeteilt werden, ob die Begutachtung auf mehrere Tage verteilt war; in solchen Fällen auch, wann jeweils Exploration oder andere methodische Schritte erfolgten. Es sollte auch deutlich werden, falls in einer Exploration von sehr langer Dauer nur sehr wenig Inhalte produziert wurden.

Starre Regeln sollten nicht vorgegeben werden. Es empfiehlt sich, daß ein Sachverständiger ein detailliertes Ablaufprotokoll über seine Begutachtung führt. Welche Details in welcher Weise im Gutachten mitteilenswert erscheinen, ist bereits eine sachverständige Entscheidung.

- 4.1 Ist es Aufgabe der Sachverständigen, insbesondere zur Analyse und Aufhellung der Entstehung und der Geschichte der Aussage alle Personen, mit denen die Zeugin über die mutmaßlichen Vorfälle gesprochen hat, informatorisch zu befragen?

Die Rekonstruktion der Aussagegenese ist wesentlicher Bestandteil bei allen denkbaren aussagepsychologischen Fragestellungen. Bei der Fragestellung „Suggestion vs. Erlebnisgehalt“ ist sie sogar der zentrale Analyseschritt. Dennoch kann es nicht Aufgabe von aussagepsychologischen Sachverständigen sein, regelmäßig alle Personen, mit denen Zeugen über die mutmaßlichen Fälle gesprochen haben, informatorisch zu befragen. Ein solches Vorgehen wirft sowohl Fragen hinsichtlich der Kapazitätsgrenzen von Sachverständigen als auch – was wichtiger erscheint – Fragen möglicher Grenzüberschreitungen von der Sachverständigentätigkeit zur Ermittlungstätigkeit auf – selbst dann, wenn die Befragungen lediglich als informatorisch deklariert werden. Daneben stellen sich auch Fragen der Praktikabilität, zum Beispiel die Frage, ob

relevante Personen der Einladung von Sachverständigen folgen würden. Darüber hinaus sind rechtliche Probleme berührt (u.a. Belegungspflichten).

Die von Sachverständigen anzustellenden erforderlichen Überlegungen über die Aussageentstehung und weitere Aussageentwicklung können in den meisten Fällen auf der Grundlage der in den Akten enthaltenen Anknüpfungstatsachen und in Verbindung mit den von den zu begutachtenden Zeugen erhaltenen Befundtatsachen erfolgen. Eine vom Sachverständigen für notwendig gehaltene (zusätzliche) Vernehmung weiterer Zeugen kann von diesem beim Gericht angeregt werden.

Bei sehr jungen Kindern (Vor- und Grundschulalter) kann die informativische Befragung einer relevanten Auskunftsperson auch zu Fragen der Aussagegenese (neben biographischen Fragen) notwendig sein, zumal wenn der Gutachtauftrag zu Beginn von Ermittlungen (z. B. durch die Staatsanwaltschaft) erfolgt und noch wenige Anknüpfungstatsachen in der Ermittlungsakte vorliegen.

Wegen Unsicherheiten in der Praxis der Glaubwürdigkeitsbegutachtung könnte hier durch verbindliche Interpretation der rechtlichen Vorgaben Klärung geschaffen werden.

4.2 Wären die Ergebnisse solcher Befragungen festzuhalten und im Gutachten mitzuteilen?

Ja. Der Sachverständige hat nicht nur die Tatsache mitzuteilen, daß er solche Befragungen durchgeführt hat, sondern auch den Inhalt solcher Befragungen.

5.1 Welchen Wert haben Phantasieproben?

Im Rahmen der Beurteilung der personalen Kompetenz von Zeugen als Bezugsgröße für den Wert der Qualitätsanalyse zur Glaubhaftigkeitseinschätzung können auch „Phantasieproben“ ihren Stellenwert haben. Kreative Kompetenzen des Zeugen können alternativ oder zusätzlich mit Hilfe biographischer (auch fremdanamnestic) Details eingeschätzt werden oder aus der Beobachtung des Zeugen bei der allgemeinen Exploration erschlossen werden. Wie für Suggestionsproben gilt auch für Phantasieproben, daß sie desto weniger Erkenntniswert für den Einzelfall besitzen, je mehr sie sich in Art und Inhalt von dem in Frage stehenden Sachverhalt unterscheiden. Direkte Schlüsse von entsprechenden Proben auf den Realitätsgehalt von Zeugenaussagen sind nicht möglich.

5.2 Ist es erforderlich, Sexualwissen und vorhandene Sexualerfahrungen zu explorieren?

Im Rahmen aussagepsychologischer Begutachtungen über in Frage stehende sexuelle Handlungen beinhaltet die Etablierung des Bezugswissens für die Interpretation der Qualitätsanalyse im Hinblick auf den Realitäts-

gehalt der Zeugenaussage, daß sich der Sachverständige Kenntnis über das Sexualwissen und mögliche sexuelle Erfahrungen des zu begutachtenden Zeugen verschafft. Es geht um die Abschätzung, ob sexuelle Parallelerlebnisse bzw. Sexualkenntnisse die Qualität der Zeugenaussage ausreichend erklären können, ohne daß auf eine Erlebnisgrundlage des zu beurteilenden Sachverhalts geschlossen werden kann. Der Umfang der sexualbezogenen Exploration oder fremdanamnestischer Erhebungen (z. B. bei jungen Kindern) richtet sich dabei nach den Gegebenheiten des Einzelfalles. So ist bei jungen Kindern und retardierten Zeugen auch die Exploration von Sexualwissen angezeigt, während dies bei älteren Zeugen vorausgesetzt werden kann (eine routinemäßige Abfrage von Begriffen für Körperteile anhand anatomisch korrekter Puppen ist bei älteren Kindern oder Jugendlichen überflüssig und dilettantisch), bei ihnen steht dagegen die Exploration nach sexuellen Erfahrungen im Vordergrund.

- 5.3 Ist es statthaft, für die Beurteilung der Glaubwürdigkeit eines Zeugen sog. Außenkriterien heranzuziehen?

Gemäß herrschendem Verständnis hat ein Sachverständiger spezifische Fragen mit Hilfe seiner spezifischen fachlichen Kenntnisse und Methoden zu bearbeiten. Der spezifische fachpsychologische methodische Zugang bei der Glaubwürdigkeitsbegutachtung besteht nach derzeitigem Verständnis in der Qualitätsanalyse mit Rückbezug auf Personen- und Fehlerquellenanalysen. Die Verwertung von Außenkriterien wie medizinische Befunde oder andere Sachbeweise bzw. Zeugenaussagen anderer Personen ist Beweiswürdigung und würde eine Kompetenzüberschreitung des aussagepsychologischen Sachverständigen darstellen. Eine solche ist allerdings nicht gegeben, wenn ein Sachverständiger vorfindbare Anknüpfungstatsachen in seine aussagepsychologischen Schlußfolgerungen integriert.

6. Sind weitere Anmerkungen zur Erstattung psychologischer Glaubwürdigkeitsgutachten zu machen?

Auf den Abschnitt 2.3 über das schriftliche Gutachten wird hingewiesen. Erneut sei betont, daß fachliche Standards im Hinblick auf die vielfältigen und jeweils unterschiedlichen Anforderungen bei unterschiedlichen praktischen Problemstellungen nicht in einfachen Regelwerken formuliert werden können und daß zu starre Vorschriften auf praxeologischer Ebene der Komplexität des Gegenstandsbereichs forensisch-aussagepsychologische Begutachtung nicht gerecht werden.

Zu 1.1, 1.2, 2.1, 2.2 ist anzumerken, daß die Gutachten mündlich zu erstatten sind und die schriftlichen Gutachten nur der Vorbereitung dienen.

Dieser Hinweis im Fragenkatalog des BGH verweist darauf, daß Umfang und Detaillierungsgrad bei schriftlichen (vorläufigen) Gutachten auch durch Vereinbarungen zwischen Auftraggebern und Sachverständigen

nach den Gegebenheiten des Einzelfalls geregelt werden können. Starre Vorschriften würden den unterschiedlichen Notwendigkeiten nicht gerecht werden.

- II. Welches sind die wesentlichen sachlichen Kriterien für die Beurteilung der Glaubwürdigkeit insbesondere kindlicher und jugendlicher Zeugen? *Diese Frage betrifft die Methodik forensisch-aussagepsychologischer Begutachtungen insgesamt sowie ihre wissenschaftlichen Grundlagen. Dazu wurden umfangreiche Ausführungen gemacht, deren Zusammenfassung in „übergeordneten Gütemaßstäben“ bereits in der Antwort zu Frage I., 2.3 wiedergegeben wurde*.*

Anschrift der Verfasser:

Prof. Dr. Max Steller und Dr. Renate Volbert
Institut für Forensische Psychiatrie
der Freien Universität Berlin
Limonenstraße 27
12203 Berlin

* Ein III. und IV. Komplex von (insgesamt 9) Fragen über das primäre Glaubwürdigkeitsgutachten und die darauf bezogene Methodenkritik durch Prof. Sch. wird hier nicht veröffentlicht, da diesbezügliche Vorbehalte der Erstgutachterin nicht ausgeräumt werden konnten.