

# ***Berufsfelder der Rechtspsychologie***

*Dokumentation der 9. Arbeitstagung der Fachgruppe Rechtspsychologie  
in der Deutschen Gesellschaft für Psychologie (DGPs)  
an der Westfälischen Wilhelms-Universität Münster  
vom 13.-15. September 2001*

***Wolfgang Bilsky und Cordula Kähler (Hrsg.)***

*unter Mitarbeit von  
Julia Müller, Jutta Siepert, Arne Weigold  
und Nina Bergmann*

Programmbeirat der 9. Arbeitstagung:

Prof. Dr. Thomas Bliesener (Kiel)  
Prof. Dr. Rudolf Egg (Wiesbaden)  
PD Dr. Werner Greve (Hannover)  
Prof. Dr. Günter Köhnken (Kiel)  
Prof. Dr. Michael Stadler (Bremen)  
Dr. Renate Volbert (Berlin)

Westfälische Wilhelms-Universität  
Münster 2001

ISBN 3-00-008097-X



## Vorwort

Die 9. Arbeitstagung der Fachgruppe Rechtspsychologie in der Deutschen Gesellschaft für Psychologie (DGPs) steht unter dem Leitthema "Berufsfelder der Rechtspsychologie". Den Anstoß zu dieser Themenwahl haben zum einen Gespräche gegeben, die im Juni 1998 zwischen dem damaligen Vorstand der DGPs und Vertretern der Fachgruppen zum Thema "Neue Berufsfelder der Psychologie" geführt worden sind. Zum anderen war im Frühjahr 2000 von seiten der DGPs die Frage aufgeworfen worden, ob die derzeit gültige Rahmenprüfungsordnung, unabhängig von der um BA/MA-Studiengänge geführten Diskussion, zur Sicherung der Kontinuität in der akademischen Ausbildung von Psychologen überarbeitet werden solle. In diesem Zusammenhang wurde angedeutet, die Zulassung neuer Anwendungsfächer wie z.B. der Rechtspsychologie zu ermöglichen. Angesichts dieser Initiativen und Überlegungen erschien es sinnvoll, innerhalb der Fachgruppe Rechtspsychologie die Berufsfelddiskussion aufzugreifen und Forschungsarbeiten ebenso wie Forschungsdesiderate aus unterschiedlichen Arbeitsbereichen zur Diskussion zu stellen. Außer durch die Tagungsbeiträge der Teilnehmer sollte dies durch ausgewählte Plenumsvorträge geschehen.

Unsere Einladung, Beiträge für die diesjährige Arbeitstagung einzureichen, ist, trotz mehrerer sich in diesem Jahr überschneidender Veranstaltungen, auf eine erfreuliche Resonanz gestoßen. Neben den klassischen Themen aus dem Bereich der forensischen Psychologie sind dabei, ganz im Sinne unseres Leitthemas, auch zahlreiche Beiträge angemeldet worden, die sich auf andere und weniger prominente Arbeitsfelder der Rechtspsychologie beziehen. Die

**Programmübersicht** vermittelt einen ersten Eindruck hiervon und vom Ablauf der Veranstaltung.

Neben einer inhaltlich möglichst breit gefächerten Diskussion war uns zudem an einer Intensivierung des innerfachlichen Informationsaustausches gelegen. Aus diesem Grund bestand bei dieser Arbeitstagung erstmals die Möglichkeit, **Langfassungen** der angemeldeten Beiträge vorab einzureichen, so daß die vorliegende Tagungsdokumentation bereits von Beginn der Tagung an zur Verfügung steht. Die Tatsache, daß mehr als die Hälfte der Referenten von diesem Angebot Gebrauch gemacht und Langfassungen eingereicht hat, spricht aus unserer Sicht für ein entsprechendes Interesse und bestätigt die positiven Erfahrungen, die mit vergleichbaren Tagungsdokumentationen gemacht worden sind.

Angesichts der Kürze der für Anfertigung und Herausgabe der Langfassungen zur Verfügung stehenden Zeit war eine über grundlegende formale Aspekte hinausgehende Überarbeitung der Texte seitens der Herausgeber ebensowenig möglich wie eine Revision durch die Autoren. Das Spektrum der mit dieser CD vorgelegten Beiträge ist dementsprechend heterogen und reicht von eher knappen Arbeitsskizzen bis zu ausgearbeiteten Manuskripten. Dies scheint uns im Hinblick auf eine möglichst zügige fachinterne Information im Rahmen der Arbeitstagung vertretbar, zumal alle Rechte der weiteren Nutzung der eingereichten Manuskripte bei den Autoren verbleiben. Über die Beiträge derjenigen Autoren, bei denen Datenerhebung, Datenanalysen oder aber die Dokumentation der auf der Arbeitstagung vorzustellenden Ergebnisse zum Zeitpunkt der Beitragsannahme noch nicht abgeschlossen waren, informieren schließlich die auf dieser CD ebenfalls wiedergegebenen **Abstracts**. Durch die **Programmübersicht**, die **Abstracts** und die **Langfassungen** der Beiträge kann sich der Leser so einen umfassenden Überblick über die gesamte Tagung verschaffen.

Die mit dieser CD vorgelegte Dokumentation der 9. Arbeitstagung der Fachgruppe Rechtspsychologie hätte nicht ohne die tatkräftige Unterstützung durch eine Vielzahl von Freunden und Kollegen realisiert werden können. Ihnen möchten wir an dieser Stelle ausdrücklich danken.

Münster, September 2001

Wolfgang Bilsky und Cordula Kähler

# 9. Arbeitstagung der Fachgruppe Rechtspsychologie in der Deutschen Gesellschaft für Psychologie e. V.

## Programmübersicht

Zeit	Mittwoch, 12.09.01	Donnerstag, 13.09.01	Freitag, 14.09.01	Samstag, 15.09.01
08:00	Fortbildungsveranstaltung für Juristen  (Vortagung)	<i>Öffnung des Tagungsbüros</i>	<i>Öffnung des Tagungsbüros</i>	<i>Öffnung des Tagungsbüros</i>
09:00-09:45		Begrüßung	Plenumsvortrag: Mediation	7.3 Strafzumessung
09:50-10:50		1.1 Aussagepsychologie	4.1 Zeugen	8.1 Diagnostik
10:50-11:10		2.1 Täter und Opfer	5.1 Delinquentes Verhalten	"
11:10-12:10		<i>Kaffeepause</i>	<i>Kaffeepause</i>	<i>Kaffeepause</i>
12:15-13:00		1.2 Aussagepsychologie	4.2 Zeugen	7.4 Strafzumessung (AG)
13:00-14:30		2.2 Täter und Opfer	5.2 Delinquentes Verhalten	8.2 Diagnostik
14:30-15:30		Plenumsvortrag: Polizei	Plenumsvortrag: Sachverständige	Schlußplenum
15:30-15:50		<i>Mittagspause</i>	<i>Mittagspause</i>	-
15:50-16:50		1.3 Aussagepsychologie	6.1 Polizei	
17:00-17:45		3.1 Straf- und Maßregelvollzug	7.1 Strafzumessung	
17:45-19:00		<i>Kaffeepause</i>	<i>Kaffeepause</i>	
19:00		1.4 Aussagepsychologie	6.2 Polizei	
20:00	3.2 Straf- und Maßregelvollzug	7.1 Strafzumessung		
	Plenumsvortrag: Strafvollzug	Mitgliederversammlung		
	-			
19:00	<i>Anreise zur FG-Tagung</i>	<i>Geselliger Abend im</i>	<i>Abendbummel durch Münster</i>	
20:00	<i>- ab 20 Uhr -</i>	<i>Mühlenhof-Museum</i>		
	<i>informeller Begrüßungs-</i>			
	<i>abend im "Töddenhoek"</i>			



# Übersicht

## *Abstracts*

*Heike Argstatter & Sandra Loohs*

Was wissen Jugendliche über Sexualdelikte? Eine Untersuchung zu Skripten von sexuellem Mißbrauch an Kindern.....35

*Katrin Behruzi & Udo Undeutsch*

Gründe für das Verschweigen sexuellen Mißbrauchs – Ergebnisse einer Analyse der Aussagen von Opferzeugen .....49

*Doris Bender, Friedrich Lösel & Thomas Bliesener*

Zum Einfluß personaler Ressourcen auf den Zusammenhang zwischen Substanzgebrauch und Delinquenz (Abstract) .....6

*Thomas Bliesener & Friedrich Lösel*

Konfliktverhalten und Konfliktwahrnehmung von Tätern und Opfern der Gewalt unter Jugendlichen (Abstract) .....6

*Michaela Brocke, Judith Bus, Heinz Holling & Wolfgang Bilsky*

Deliktmerkmale und Strafzumessung bei einem Vergewaltigungsdelikt: Vergleich einer Fragebogenstudie mit einer Conjointanalyse.....59

*Christine Brugger*

Psychologische Sachverständigengutachten im Entlassungsverfahren von geistig abnormen zurechnungsfähigen Rechtsbrechern in Österreich .....71

*Thomas P. Busch & O. Berndt Scholz*

Die Generierung empirischer Täterprofile: Eine post-hoc Klassifikation am Beispiel der Tötung des Intimpartners .....81

*Andrea Bütikofer & Margit Oswald*

Das Kind als Zeuge: Einfluss des Erzählverbots auf die Wiedergabeleistung von Vorschulkindern (Abstract) .....9

*Klaus-Peter Dahle*

Zur prognostischen Validität des HCR-20 und des LSI-R bei deutschen Strafgefangenen (Abstract) .....9

*Rudolf Egg*

Sozialtherapeutische Einrichtungen im Justizvollzug – aktueller Stand und Entwicklungen (Abstract) .....10

*Johann Endres & Wilfried Hommers*

Bindungsstile bei Straftätern – Konstruktvalidität eines neuen Verfahrens (Abstract) ..10

*Birte English*

Psychologische Rhetorik für Juristen – Evaluation eines Trainingsprogramms und die Wirkung der rhetorischen Gestaltung von Verteidigungsplädoyers auf richterliche Strafurteile (Abstract) ..... 11

*Matthias Fischer*

Vergleich von kriminellen Handlungen mit und ohne Gewaltanwendung hinsichtlich Tatplanung, Tathergang und Nachtatverhalten.....91

*Immo Fritsche*

Experimentelle Testung der Neutralisationstheorie abweichenden Verhaltens .....99

*Uwe Füllgrabe*

Psychologie der Eigensicherung..... 127

*Ute Gabriel & Colette Marti*

Kriminalitätsfurcht und Strafhärte: Es kommt darauf an, ob man an „Prävention durch Strafe“ glaubt.....137

*Lutz Gretenkord*

Empirisch fundierte Prognosestellung im Maßregelvollzug nach § 63 StGB – EFP-63 ..... 147

*Luise Greuel*

Gedächtnispsychologische Grundlagen der Glaubhaftigkeitsbegutachtung – Ansätze zu einer theoretischen Fundierung der “Undeutsch-Hypothese” (Abstract) .....14

*Werner Greve & Dirk Enzmann*

Etikettierungen im Jugendstrafvollzug? Wider die Gewißheiten des Labeling-Ansatzes (Abstract) ..... 15

*Everhard von Grootte*

Prognose von Täterverhalten bei Geiselnahmen ..... 157

*Wilfried Hommers*

Maßgeschneiderte psychometrische Verfahren für die familienrechtliche Begutachtung..... 169

*Wilfried Hommers & Martin Lewand*

Zum Einfluss des Tatmotivs auf eine Voraussetzung der strafrechtlichen Verantwortungsreife..... 181

*Daniela Hosser*

Soziale Unterstützung im Strafvollzug: Der Einfluss sozialer Beziehungen auf das Befinden und die Normorientierung junger Männer ..... 191



<i>Jörg Hupfeld</i>	
Dimensionen retributiver Gerechtigkeit .....	199
<i>Prisca Jager &amp; O. Berndt Scholz</i>	
Die Verwendung ereigniskorrelierter Potentiale im Rahmen der Psychophysiologischen Aussagebegutachtung .....	213
<i>Claudia Kestermann &amp; Luise Greuel</i>	
Qualitätssicherung kriminalpolizeilicher Ermittlungspraxis durch Verhaltensbeobachtung – erste Ergebnisse aus dem Pilotprojekt BEOCOP (Abstract) .....	19
<i>Edgar W. Klinger &amp; Heiner Gehring</i>	
Vermutungen von Juristen und Rechtslaien zur Wirkung von Gerichtsfernsehen .....	223
<i>Clemens Lorei</i>	
Psychologische Faktoren der Arbeitssicherheit bei der Polizei am Beispiel Schutzwesten .....	239
<i>Clemens Lorei</i>	
Ergebnisse einer Bürgerbefragung zur Akzeptanz von polizeilichen Kontroll- und Eigensicherungsmaßnahmen .....	253
<i>Friedrich Lösel &amp; Thomas Bliesener</i>	
Entwicklung und Persönlichkeit von Hooligans: Komparative Fallanalysen (Abstract) .....	22
<i>Friedrich Lösel &amp; Nicole Raichle</i>	
Kann man durch das Wissen über Realkennzeichen glaubhafter lügen? (Abstract) .....	22
<i>Charlotte Mohrbach</i>	
Explorationstechnik – Einhaltung von Vernehmungsprinzipien zur Gewinnung valider Daten .....	267
<i>Andreas Mokros &amp; Stephan Harbort</i>	
Signaturen von Sexualmördern (Abstract) .....	23
<i>Katrin Müller, Renate Volbert &amp; Detlef Busse</i>	
Der Verlauf von Strafverfahren wegen sexuellen Mißbrauchs an Kindern – Ergebnisse einer Aktenanalyse (Abstract) .....	24
<i>Susanna Niehaus</i>	
Zum Problem der Trefferquoten bei der Anwendung inhaltlicher Glaubhaftigkeitsmerkmale (Abstract) .....	24
<i>Margit Oswald, Stefan Klug, Jörg Hupfeld &amp; Ute Gabriel</i>	
Dimensionale Struktur der Reaktionen auf normabweichendes Verhalten .....	273

<i>Kerstin Panhey, Günter Köhnken &amp; Frank Eggert</i>	
Emotionen im Prozeß der Glaubwürdigkeitsbeurteilung .....	289
<i>Ernst Plaum</i>	
Multimethodale Diagnostik bei forensisch-psychologischen Begutachtungen .....	297
<i>Anne Reichert &amp; Wolfgang Bilsky</i>	
Kriminalität – Ursachenzuschreibung und Strafhärte: Eine Untersuchung aus der Sicht juristischer Laien .....	303
<i>Judith Rickers &amp; Susanna Niehaus</i>	
Stilistische Kriterien in der Glaubhaftigkeitsdiagnostik: Ein Linguistisches Modell als Analyseinstrument zur Beurteilung kindlicher Zeugenaussagen (Abstract) ....	27
<i>Anja Schmidt, Jürgen Hoyer &amp; Heike Kunst</i>	
Diagnostik psychischer Störungen bei Sexualstraftätern im Maßregelvollzug .....	321
<i>Vera Schneider &amp; Klaus-Peter Dahle</i>	
Entwicklungspsychologisch bedeutsame Faktoren in der Kriminalprognose (Abstract) .....	28
<i>Nadja Schreiber &amp; Janat Fraser Parker</i>	
“Einladen zur Spekulation” – Anforderungscharakteristika versus Quellenkonfusi- on in Falschaussagen von Kindern (Abstract) .....	29
<i>Stefan Schulz-Hardt, Eberhard Höfer &amp; Günter Köhnken</i>	
Schuldig bei Verdacht: Wie konfirmatorisches Hypothesentesten zur fälschlichen Beschuldigung wegen sexuellen Kindesmißbrauchs führt .....	335
<i>Peter Steck</i>	
Tötung in Verbindung mit sexueller Gewalttat .....	349
<i>Barbara Trinkl, Elena Guberova &amp; Siegfried L. Sporer</i>	
Der Ausländereffekt bei türkischen und deutschen Kindern: Ergebnisse einer neuen Matching-Aufgabe (Abstract) .....	31
<i>Renate Volbert, Julia Braun, Yvonne Gretenkord, Maren Teske &amp; Sandra Wilma-Mews</i>	
Konstanz in erlebnisbasierten und erfundenen Aussagen (Abstract) .....	31
<i>Renate Volbert &amp; Yvonne Rutta</i>	
Verbesserung der Inhaltsqualität von Falschaussagen durch Training (Abstract) .....	32
<i>Arne Weigold &amp; Dirk Wentura</i>	
Wiedererkennen des Täters oder Identifizieren des Verdächtigten? Experimentelle Evidenz für einen Bias in der Zeugengegenüberstellung .....	355

## **Was wissen Jugendliche über Sexualdelikte? – Eine Untersuchung zu Skripten von sexuellem Mißbrauch an Kindern**

*Heike Argstatter & Sandra Loohs*  
*Lehrstuhl Psychologie II, Universität Regensburg*

Untersuchungen (z. B. von Pezdek & Holst, 1992; Greenberg, Westcott & Bailey, 1998) haben gezeigt, daß Versuchspersonen Skripten zu unterschiedlichen Delikten besitzen können, ohne jemals selbst in eine Straftat involviert gewesen zu sein. Bei der Begutachtung der Glaubhaftigkeit von Zeugenaussagen ist eine zentrale Frage, inwieweit Zeugen in der Lage sind, aufgrund ihres Alltagswissens vermeintlich erlebnisbezogene Berichte zu konstruieren. Speziell bei den Aussagen Jugendlicher zu sexuellem Mißbrauch wird häufig argumentiert, die Zeugen könnten Inhalte aus Jugendzeitschriften in ihre Angaben übernommen haben.

Versuchspersonen der vorliegenden Untersuchung waren 84 Jugendliche (Schüler der siebten Jahrgangsstufe, davon 21 Hauptschüler, 31 Realschüler und 32 Gymnasiasten). Die Schüler wurden anonym mit Hilfe eines Fragebogens zu ihren Einschätzungen von Ausmaß und Umständen von sexuellem Mißbrauch sowie zu ihren diesbezüglichen Informationsquellen befragt. Darüber hinaus wurden die Jugendlichen gebeten, eine ihrer Ansicht nach typische „Mißbrauchssituation“ zu konstruieren.

Die Angaben der Jugendlichen wurden mit den aus der Literatur bekannten Daten abgeglichen, ebenso mit den Ergebnissen einer Textanalyse der Serie „Mißbraucht - auch Du?“, die 1996 in der Zeitschrift „Bravo Girl“ abgedruckt worden war.

Die Ergebnisse zeigen, daß die Schüler mit ihren Einschätzungen einzelner Fakten wie Prävalenzrate, Dunkelziffer oder Tätergruppe größtenteils korrekt lagen, verglichen mit den aus der forensischen Forschung bekannten Daten, während sich im Vergleich zu den Informationen aus „Bravo Girl“ Diskrepanzen zeigten. Die Mehrzahl der Jugendlichen war nicht in der Lage, aufgrund ihres Faktenwissens ein „Mißbrauchsskript“ zu konstruieren oder deliktsspezifische Vorgehensweisen von Tätern zu beschreiben. Dies betraf auch die Jugendlichen, die angaben, aus Jugendzeitschriften Informationen zu sexuellem Mißbrauch bezogen zu haben.

Die Ergebnisse werden im Hinblick auf ihre Bedeutung für die Glaubhaftigkeitsbegutachtung von Zeugenaussagen und die Prävention von sexuellem Mißbrauch diskutiert.

## **Gründe für das Verschweigen sexuellen Mißbrauchs – Ergebnisse einer Analyse der Aussagen von Opferzeugen**

*Katrin Behruzi & Udo Undeutsch*  
*Psychologisches Institut, Universität Köln*

Die Deliktsspezifität einer Aussage gilt als eines der Glaubhaftigkeitsmerkmale bei der Glaubhaftigkeitsbegutachtung. Als ein solches deliktsspezifisches Verhaltensmuster des Opfers bei sexuellem Mißbrauch kann das Verschweigen des Mißbrauchs angesehen werden, das Kinder zum Teil über Jahre beibehalten. Die Basis dieses Verhaltensmusters sind die Gründe für das Verschweigen, also solche Argumente, die beim Abwägen negativer Konsequenzen eines Bekanntwerdens des Mißbrauchs ins Gewicht fallen und das Kind dazu bringen, das Geschehene bewußt und absichtsvoll zu verbergen. Im vorliegenden Artikel geht es um die verschiedenen Gründe, die Kinder veranlassen, keiner dritten Person etwas über einen stattgefundenen sexuellen Mißbrauch zu sagen. Diese Argumente können einerseits vom Täter geäußert und andererseits vom Kind selbst überlegt werden. Auf

einer Datenbasis von 64 – sowohl von erfahrenen Sachverständigen als auch vom Gericht – als glaubhaft eingestuften Fällen sexuellen Mißbrauchs wurde ein Kategoriensystem zur Erfassung von Gründen für das Verschweigen erstellt. Es differenziert zwischen Schweigegeboten, die entweder vom Täter ausgesprochen wurden oder die sich das Kind selbst auferlegt hat. Insgesamt 24 verschiedene Gründe für das Verschweigen konnten unterschieden werden, wobei von einzelnen Zeugen bis zu fünf verschiedene Gründe geschildert wurden. Die verschiedenen Schweigegebote werden jeweils inhaltlich beschrieben. Unter Verwendung des gewonnenen Kategoriensystems werden zudem statistische Analysen durchgeführt, bei denen die Gründe für das Verschweigen zu folgenden weiteren Variablen in Beziehung gesetzt werden: Alter des Kindes bei der letzten von ihm berichteten Tat, Geschlecht des Kindes, Täter-Opfer-Beziehung in ihrem formalen Verwandtschafts-/Bekanntschaftsgrad und die Häufigkeit der Vorfälle.

### **Zum Einfluß personaler Ressourcen auf den Zusammenhang zwischen Substanzgebrauch und Delinquenz**

*Doris Bender<sup>1</sup>, Friedrich Lösel<sup>1</sup> & Thomas Bliesener<sup>2</sup>*

*<sup>1</sup>Institut für Psychologie, Universität Erlangen-Nürnberg, <sup>2</sup>Institut für Psychologie, Universität Kiel*

Substanzgebrauch und Delinquenz werden oft zu einem Syndrom des Problemverhaltens zusammengefaßt. In einer früheren Arbeit zeigte sich aber, daß das Familienklima und die Gleichaltrigen-gruppe auf diesen Zusammenhang einen Einfluß haben (Lösel & Bliesener, 1998). Insbesondere bei Mädchen hatte das Familienklima – je nach Ausprägung – eine wesentliche Risiko- und Schutzfunktion. In einem analogen Design mit typischen und untypischen Gruppen aus einer Stichprobe von über 1100 Jugendlichen soll nun der Einfluß von personalen Ressourcen geprüft werden. Dabei handelt es sich um Merkmale der sozialen Kompetenz und des Coping-Verhaltens, die teilweise über Selbstauskünfte und teilweise über Lehreranteile erfaßt werden.

### **Konfliktverhalten und Konfliktwahrnehmung von Tätern und Opfern der Gewalt unter Jugendlichen**

*Thomas Bliesener & Friedrich Lösel*

*Institut für Psychologie, Universität Kiel, Institut für Psychologie, Universität Erlangen-Nürnberg*

In den vergangenen Jahren sind national wie international einige zum Teil groß angelegte Umfragestudien zur Gewalt unter Schülern durchgeführt worden. Diese lieferten wichtige Daten zur Prävalenz schulischer Gewalt und konnten eine Reihe von Zusammenhängen der Dissozialität bzw. Viktimisierung mit verschiedenen sozialen und personalen Merkmalen aufdecken. Weniger aussagekräftig sind derartige Befragungsdaten jedoch hinsichtlich des konkreten Verhaltens von Tätern und Opfern in Konflikten, da hier verschiedene psychologische Prozesse wie z.B. Selbstdarstellungsbedürfnisse, Attributionsfehler, mangelnde Reflexion eigenen Verhaltens usw. die Selbstauskünfte der Person verzerren können. Für die Prävention und Intervention der Gewalt unter Schülern ist jedoch gerade das Verständnis des konkreten Verhaltens notwendig. Dies gilt besonders für das Konfliktverhalten der Opfer schulischer Gewalt. Die vorliegende Studie widmet sich deshalb der Analyse des Verhaltens und der sozialen Wahrnehmung von Tätern und Opfern schulischer Gewalt in einer Konfliktsi-

tuation. Die Jugendlichen dieser Untersuchung waren in einer Voruntersuchung nach ihrer Selbst- bzw. Lehrerbeurteilung entweder als Täter, Opfer oder hinsichtlich Schülersgewalt unauffällige Schüler klassifiziert worden. 102 männliche Jugendliche im Alter zwischen 14 und 17 Jahren wurden jeweils in einem PC-gestützten Konfliktspiel gegen einen bisher unbekanntem Spielpartner beobachtet und anschließend interviewt. Reaktionszeiten und gewählte Reaktionsalternativen in verschiedenen Phasen des Konfliktspiels wurden aufgezeichnet und analysiert. Im Beitrag werden Ergebnisse zu differentiellen Reaktionsmustern von Tätern und Opfern berichtet. Konsequenzen für die Prognose, Prävention und Intervention jugendlicher Dissozialität und Viktimisierung im schulischen und außerschulischen Kontext werden diskutiert.

### **Deliktmerkmale und Strafzumessung bei einem Vergewaltigungsdelikt: Vergleich einer Fragebogenstudie mit einer Conjointanalyse**

*Michaela Brocke, Judith Bus, Heinz Holling & Wolfgang Bilsky*  
*Psychologisches Institut IV, Universität Münster*

Unterschiede in der Strafzumessung sind in der Vergangenheit verschiedentlich in Fragebogenstudien untersucht worden, in denen den Befragten (zumeist Laien) systematisch variierte Deliktmerkmale mittels Vignetten vorgegeben wurden. Datenerhebung und Kontrolle personenbezogener 'Störvariablen' (z.B. individuell unterschiedlicher Attributionsstile, Strafhärteeinstellungen, etc.) erfolgten dabei typischerweise durch einen kovarianzanalytischen Plan ohne Meßwiederholung (between-subjects-design). In der hier vorgestellten Studie wurde, in Ergänzung eines solchen konventionellen Ansatzes (Bus, 2001), in einem zweiten Schritt eine adaptive Conjointanalyse (vgl. Holling, Jütting & Großmann, 2000) durchgeführt. Dieser liegt ein Meßwiederholungsdesign (within-subjects-design) zugrunde, das die Kontrolle individueller Beurteilermerkmale gewährleistet; eine gesonderte Erfassung derartiger Merkmale erscheint daher nicht erforderlich. Zu klären war, inwieweit sich die Ergebnisse der konventionellen Fragebogenstudie (between design) von denen der Conjointanalyse (within design) unterscheiden. Gegenstand der Untersuchung war die Strafzumessung von Laien bei einem fiktiven Vergewaltigungsdelikt in Abhängigkeit von zwei Opfermerkmalen und einem Merkmal der Täter-Opfer-Beziehung. Bei der Untersuchung mit dem konventionellen (between subjects) Design zeigten sich bei einer univariaten Varianzanalyse mit der Strafzumessung als abhängige Variable und den Deliktmerkmalen als unabhängige Variablen keine Effekte. Erst durch zusätzliche Aufnahme von Beurteilermerkmalen konnte in einer Kovarianzanalyse ein signifikanter Einfluß eines der drei Deliktmerkmale gesichert werden. Die Ergebnisse der Conjoint-Analyse zeigten hingegen, daß jedem der drei Merkmale ein signifikantes Gewicht bei der Strafzumessung zukommt. Darüber hinaus ließen sich bei den Befragten clusteranalytisch Subgruppen identifizieren, innerhalb derer den Deliktmerkmalen teilweise unterschiedliche Bedeutung beigemessen wird. Insgesamt erwies sich die Conjointanalyse als Methode, die zur Untersuchung des Zusammenhangs von Deliktmerkmalen und Strafzumessung geeignet ist.

## **Psychologische Sachverständigengutachten im Entlassungsverfahren von geistig abnormen zu-rechnungsfähigen Rechtsbrechern in Österreich**

*Christine Brugger*

*Psychologisches Institut, Universität Wien*

Im Rahmen des forensischen Schwerpunktes in der Lehr- und Forschungspraxis des Ordinariats Klinische Psychologie der Universität Wien entstand diese Arbeit in Form einer Diplomarbeit, die zusammen mit Dr. Werdenich (Leiter der Justizanstalt Favoriten) konzipiert wurde. Die psychologischen Gutachten im Entlassungsverfahren aus der Maßnahme für geistig abnorme Rechtsbrecher wurden mit quantitativen und qualitativen Verfahren bezüglich ihrer inneren Struktur untersucht und psychiatrischen Sachverständigengutachten gegenübergestellt. Insgesamt wurden 32 psychologische und 31 psychiatrische Sachverständigengutachten untersucht. Nach einer theoretischen Aufarbeitung des österreichischen Maßnahmenrechtes und der Durchführung des Maßnahmenvollzuges in Österreich werden im empirischen Teil die Gutachten inhaltsanalytisch untersucht. Hier findet eine kritische Auseinandersetzung mit den verwendeten Verfahren, der Durchführung der Untersuchung, aber auch mit Fragen der Qualitätssicherung im Begutachtungsprozeß statt. Die Arbeit dokumentiert die gegenwärtige Begutachtungspraxis im Entlassungsverfahren in Österreich sowie damit verbundene Komplikationen und Unklarheiten.

## **Die Generierung empirischer Täterprofile: Eine post-hoc Klassifikation am Beispiel der Tötung des Intimpartners**

*Thomas P. Busch & O. Berndt Scholz*

*Psychologisches Institut, Universität Bonn*

Ein regelgeleitetes wissenschaftlich-induktives Vorgehen zur Generierung von psychologischen Täterprofilen wird vorgestellt und am Beispiel der Tötung des Intimpartners praktisch demonstriert. Dazu wurde ein Kategoriensystem zur inhaltsanalytischen Erfassung von Täterprofilen bei Tötungsdelikten konstruiert, welches sich aus 45 erlebens- und verhaltensbezogenen Merkmalen in sieben psychologischen Kategorien zusammensetzt. Jedes dieser 45 Merkmale weist einen mehr oder weniger direkten Bezug zur Tathandlung auf. Die Ausprägung der sieben Kategorien ermöglicht eine post-hoc Klassifikation der Täter in deliktspezifische Täterprofile. In die Analyse gingen 47 polizeiliche Erstaussagen zu Tötungsdelikten ein. Aufgrund der Ergebnisse einer durchgeführten Cluster- und Diskriminanzanalyse ließ sich zeigen, dass es bei der Tötung des Intimpartners nicht einen einzigen, umschriebenen Tätertyp gibt, sondern mindestens zwei deliktspezifische Tätergruppen unterschieden werden können. Statistisch bedeutsame Unterschiede ( $p < .05$ ) zeigten sich in folgenden Bereichen: „Aggressivität in der Vorgeschichte“, „Verhalten in Konfliktsituation“, „Planungsniveau der Tat“ und „Kontrolle bei der Tatausführung“. Die Validierung der generierten Täterprofile an Einzelfällen zeigte auf kategorialer Ebene in keinem Fall eine hinreichende Übereinstimmung zwischen Beziehungstätern und Tätern, welche ein Tötungsdelikt begingen, bei dem nicht der Intimpartner das Opfer war. Die vorgestellten Täterprofile können somit für die hier untersuchte Stichprobe als deliktspezifisch für Beziehungstäter angesehen werden. Die Ergebnisse der vorliegenden Studie machen augenfällig, dass ein gemäß den definierten Regeln geleitetes Vorgehen geeignet ist Täterprofile auf der Basis quantifizierbarer und empirisch fundierter Merkmale zu generieren.

## **Das Kind als Zeuge: Einfluss des Erzählverbots auf die Wiedergabeleistung von Vorschulkindern**

*Andrea Bütikofer & Margit Oswald  
Institut für Psychologie, Universität Bern*

Im Zusammenhang mit kindlichen Zeugenaussagen und Suggestionseffekten sind in der Literatur vor allem Studien beschrieben, die suggestive Effekte von Befragungsmethoden untersuchen. Insbesondere bei sexuellem Missbrauch an Kindern kann eine Beeinflussung aber bereits im unmittelbaren Zusammenhang mit der Tat, also lange vor der Zeugenbefragung erfolgen. Solche Suggestionseffekte wurden bisher noch wenig untersucht. In einem Experiment (49 Kinder im Alter von 4 ½ bis 7 Jahren) wurde überprüft, ob sich ein *Erzählverbot* auf die mündliche Wiedergabe eines Geschehens auswirkt. Zu diesem Zweck wurde ein Film mit vier einzelnen Szenen gedreht, in denen ein Zauberer bestimmte Dinge vorführt. In der ersten Szene zieht sich der Protagonist um, in den darauffolgenden zwei Szenen führt er Tricks vor und in der letzten Szene zerbricht er ein wertvolles Glas. Nach der ersten und letzten Szene spricht der Zauberer ein Erzählverbot aus, die Kinder werden gebeten, nicht weiterzuerzählen, dass er (a) sich die neu angezogenen Kleider illegal ausgeliehen und (b) ein Glas zerbrochen hat. Zur experimentellen Variation wurden die beiden Verbotssequenzen bei der Hälfte der Kinder herausgeschnitten. Fünf Wochen nach Präsentation des Films wurden die Kinder einzeln befragt. Die Wiedergabeleistungen wurden anhand von drei Variablen erfasst: „Erzählleistung“ (Menge aller korrekt genannten Haupt- und Teilhandlungen), Anzahl beanspruchter Hilfeleistungen und Anzahl gemachter Fehler. Die Ergebnisse zeigen einen signifikanten Effekt der unabhängigen Variable „Erzählverbot“ hinsichtlich der erteilten Hilfestellungen bei den Verbotssequenzen: Die Kinder, die ein Erzählverbot erhalten, benötigen für die Wiedergabe bedeutend weniger Hilfestellungen als diejenigen ohne Erzählverbot.

## **Zur prognostischen Validität des HCR-20 und des LSI-R bei deutschen Strafgefangenen**

*Klaus-Peter Dahle  
Institut für Forensische Psychiatrie, Freie Universität Berlin*

Die derzeit wohl einflussreichsten Arbeiten zur Entwicklung effizienter Methoden für die Prognose von Rückfallkriminalität stammen von nordamerikanischen Arbeitsgruppen. Hierzu gehören etwa die Arbeitsgruppe um Christopher Webster, die mit dem HCR-20 ein Instrument zur Vorhersage insbesondere von Gewalttaten entwickelte, wie auch die Arbeitsgruppe um Don Andrews, die mit dem Level of Service Inventory (LSI-R) ein Instrument zur allgemeinen Rückfallprognose und zur Feststellung des Behandlungsbedarfs von Straftätern vorgelegt hat. Auch im deutschen Sprachraum wird in jüngerer Zeit zunehmend die Anwendung dieser im Ausland bewährten Instrumente gefordert, obwohl grundsätzliche Bedenken gegen eine ungeprüfte Übertragung bestehen. Tatsächlich stehen systematische Untersuchungen der Anwendbarkeit an unselektierten deutschen Populationen weitgehend aus, Überprüfungen der tatsächlichen Vorhersagekraft an deutschsprachigen Stichproben fehlen bislang gänzlich. Im vorgesehenen Beitrag werden erste Ergebnisse einer Studie vorgestellt, die den Nutzen der Instrumente an deutschen Strafgefangenen untersucht. Es handelt sich um eine Teilstudie aus dem Berliner CRIME-Projekt, in dessen Rahmen seit 1976 die Lebensbiographien ehemaliger Gefangener des Berliner Strafvollzugs untersucht werden. Berichtet wird über die Erprobung der deutschen Version der HCR-20 (Version Haina) und des LSI-R an einer Teilstichprobe von n=200 ehemaligen

Strafgefangenen aus dem Hauptprojekt. Der Katamnesezeitraum umfasst rund 20 Jahre nach Haftentlassung der Probanden, so dass auch die Potenz der Instrumente zur Vorhersage langfristiger Rückfallkarrieren untersucht werden konnte.

## **Sozialtherapeutische Einrichtungen im Justizvollzug – aktueller Stand und Entwicklungen**

*Rudolf Egg*

*Kriminologische Zentralstelle, Wiesbaden*

Mit dem „Gesetz zur Bekämpfung von Sexualdelikten und anderen gefährlichen Straftaten“ vom 26. Januar 1998 rückte eine bereits in den 60er Jahren propagierte und dann doch nur halbherzig umgesetzte Behandlungsmöglichkeit für Straftäter wieder in den Blickpunkt der kriminalpolitischen Diskussion: die sozialtherapeutischen Anstalten und Abteilungen des Justizvollzuges. Die Verlegung in solche behandlungsorientierten Einrichtungen war zwar schon bisher möglich, allerdings ausschließlich freiwillig und mit Zustimmung des Anstaltsleiters. Nach dem neuen Gesetz sieht der geänderte § 9 StVollzG ab dem Jahre 2003 für Sexualstraftäter mit mehr als 2 Jahren Freiheitsstrafe eine Verlegung in die Sozialtherapie nunmehr als Regelfall vor. Diese Regelung stellt die Landesjustizverwaltungen, die Justizpraxis und nicht zuletzt auch die im Strafvollzug tätigen Psychologinnen und Psychologen vor neue Herausforderungen. Bundesweit führten die neuen Bestimmungen bereits zu zahlreichen Diskussionen, Planungen und Vorhaben, aber auch zu offenen Fragen und Klärungsbedarf.

Die Kriminologische Zentralstelle in Wiesbaden begleitet die Entwicklung der sozialtherapeutischen Einrichtungen durch eine seit 1997 jährlich durchgeführte Stichtagserhebung zur Belegungssituation, zu den Insassen (Alter, Delikte, Strafen, Lockerungen) sowie zum Personal (Stellenzahl, Berufsgruppen, Supervision). Im Jahre 2001 wurde diese Stichtagserhebung erstmalig durch eine Umfrage zur Erfüllung der „Mindestanforderungen an Organisationsform, räumliche Voraussetzungen und Personalausstattung sozialtherapeutischer Einrichtungen“ ergänzt. Des Weiteren wurde der aktuelle Stand zur Umsetzung der neuen gesetzlichen Bestimmungen sowie die zukünftig vorgesehenen Schritte durch eine Umfrage unter den Landesjustizverwaltungen erhoben.

Der Beitrag berichtet über die Ergebnisse dieser Datenerhebungen und diskutiert unterschiedliche Möglichkeiten der Behandlung von Sexualstraftätern auch vor dem Hintergrund der im Rahmen einer umfassenden empirischen Studie zur Rückfälligkeit einzelner Tätergruppen ermittelten Erkenntnisse.

## **Bindungsstile bei Straftätern – Konstruktvalidität eines neuen Verfahrens**

*Johann Endres & Wilfried Hommers*

*Justizvollzugsanstalt Straubing*

Bindungsstile von Erwachsenen als relativ stabile Merkmale des Verhaltens in Partnerschaften gehen nicht nur in vielfältiger Weise in die Definition von verschiedenen Persönlichkeitsstörungen ein. Aufgrund ihrer in neueren Untersuchungen aufgedeckten Beziehungen zur Affektregulation und zum Konfliktverhalten läßt sich auch postulieren, daß sie in je spezifischer Weise für unterschiedliche Formen von Kriminalität bedeutsam werden. Mit dem Bindungsstil-Inventar, einem neu entwickelten Fragebogen, der mit 63 Items Aspekte eines sicheren, eines besorgten, eines distanzierten und eines



furchtsamen Bindungsstils erfaßt, wurden eine Stichprobe aus der Allgemeinbevölkerung (N = 262) sowie eine Stichprobe von Straftätern (ungefähr N = 70), vor allem Sexualstraftätern, untersucht. Vorgestellt werden erste Auswertungen zu den Gütekriterien des Verfahrens, zur Konstruktvalidität und zu Unterschieden zwischen den Stichproben.

### **Psychologische Rhetorik für Juristen – Evaluation eines Trainingsprogramms und die Wirkung der rhetorischen Gestaltung von Verteidigungsplädoyers auf richterliche Strafurteile**

*Birte Englich*

*Lehrstuhl für Psychologie II, Universität Würzburg*

Im Gegensatz zur klassischen Rhetorik bietet die Sozialpsychologie empirisch fundierte Erkenntnisse zur Wirksamkeit sprachlicher Präsentationstechniken. Entsprechend müsste eine psychologisch fundierte Rhetorik ein nützliches Mittel zur gezielten Darstellung juristischer Sachverhalte sein. Im Rahmen eines eineinhalb- bis zweitägigen Trainingsprogramms wurde Juristen ein Überblick über sozialpsychologische Theorien und Modelle gegeben, aus denen sich wertvolle Anregungen für argumentative Vorträge oder Redebeiträge ableiten lassen. Am Beispiel der sprachlichen Gestaltung eines Verteidigungsplädoyers sollten die Trainingsteilnehmer die Umsetzung der sozialpsychologisch fundierten Empfehlungen üben. In einem Vortest–Nachtest-Design mit Kontrollgruppe wurde der Lerneffekt dieses verbalen Trainingsprogramms evaluiert. Das Trainingsprogramm erweist sich im Rahmen dieser ausführlichen Evaluation als sehr erfolgreich. Die Teilnehmer, vornehmlich Staatsanwälte, Richter und Rechtsreferendare, bewerten das Trainingsprogramm nicht nur positiv, es zeigen sich auch konkrete Lerngewinne sowohl auf der Wissens- als auch auf der Verhaltensebene. Die trainierten rhetorischen Mittel werden nach dem Training deutlich besser erkannt und im Rahmen einer Plädoyergestaltung vermehrt umgesetzt. Schließlich erzielen rhetorisch geschulte Plädoyers auch die erwartete Wirkung auf Juristen: *Rhetorisch trainierte Verteidigungsplädoyers zu dem gleichen strafrechtlichen Fall erzielen deutlich niedrigere Strafurteile und deutlich häufiger eine Strafaussetzung zur Bewährung als rhetorisch untrainierte Verteidigungsplädoyers.* Somit stellt auch die rhetorische Gestaltung von Redebeiträgen vor Gericht einen nicht zu vernachlässigenden Einfluss auf richterliche Strafurteile dar, was gerade auch angesichts der hohen Disparität richterlicher Strafurteile ein bedeutsamer Befund ist.

### **Vergleich von kriminellen Handlungen mit und ohne Gewaltanwendung hinsichtlich Tatplanung, Tathergang und Nachtatverhalten**

*Matthias Fischer*

*Fachgruppe Psychologie, Universität Konstanz*

Simons (1988) erstellte handlungstheoretische Modelle für die Trennungstat und die Bereicherungstat und interpretierte Tötungsdelikte als Folge misslungener Problemlösungen. Volbert (1990) sowie Rehm und Servay (1989) machten deutlich, dass die Planung von Bereicherungsdelikten häufig undetailliert ist, wobei beim Täter dennoch meist die Ansicht vorhanden sei, die Tat weitgehend geplant zu haben. Diese und andere Untersuchungen (wie etwa Burgheim 1993, Erez, 1980, Steck et al., 1997) verweisen auf die Notwendigkeit der Einbeziehung handlungstheoretischer Konzepte zur Erhellung kriminalpsychologischer Probleme.

In dieser Studie werden handlungstheoretische Modelle herangezogen, um Erklärungsansätze für kriminelle Handlungen mit Gewaltanwendung zu gewinnen. Dazu erfolgt ein Vergleich von 50 Tätern, die zu einer kriminellen Handlung ohne Gewaltanwendung (Betrug, schwerer Diebstahl) und 110 Tätern, die zu einer kriminellen Handlung mit Gewaltanwendung (Raub, sexuelle Aggressionstat, sonstige Aggressionstat) mit einem halbstandardisierten Interviewleitfaden befragt wurden. Zusätzlich wurden die Gerichtsurteile eingesehen. Es wurden behaviorale, motivationale und volitionale Daten für die Tatanlaufzeit (Zeitraum die letzten vier Wochen vor der Tat), die Vorszene (wenige Stunden vor der Tat) den Tathergang und das Nachtatverhalten des Täters erfasst. Bei der Tatanlaufzeit werden besondere Lebensereignisse, Verhaltensveränderungen, ökonomische Situation, Alkohol- und Drogenkonsum sowie Zielsetzung, Abwägen und Planung der Tat seitens des Täters analysiert. Bei der Vorszene werden insbesondere konfliktverschärfende Ereignisse und Handlungen, beim Tathergang Handlungsbarrieren und Täter-Opfer-Interaktion, beim Nachtatverhalten die Organisiertheit des Verhaltens untersucht. Zusätzlich werden die Tätergruppen hinsichtlich biographischer Merkmalskomplexe varianzanalytisch verglichen. Die Ergebnisse zeigen, dass zwischen den Tätergruppen in den meisten untersuchten Merkmalen und Merkmalskomplexen signifikante Unterschiede bestehen.

### **Experimentelle Testung der Neutralisationstheorie abweichenden Verhaltens**

*Immo Fritsche*

*Institut für Psychologie, Universität Magdeburg*

Die Neutralisationstheorie – erstmals formuliert von den amerikanischen Soziologen Sykes & Matza (1957) – besitzt eine ungebrochene Popularität bei der Erklärung normwidrigen Verhaltens. Die Grundannahme, Personen brächen gesellschaftliche Normen nicht aus einem Mangel an Normakzeptanz, sondern weil sie in der Lage sind, diese von ihnen selbst geteilten Normen mithilfe von Rechtfertigungen und Entschuldigungen situativ zu neutralisieren, wurde auch über den strafrechtlich relevanten Bereich hinaus auf andere Normverletzungen übertragen. Trotz des hohen Verbreitungsgrades der Theorie war es der psychologischen und soziologischen Forschung bislang nicht möglich, die kausale Wirkung von Neutralisationen auf Verhalten angemessen zu belegen, da zumeist lediglich Korrelationsstudien mit Selbstberichtsdaten vorliegen. In der vorgestellten Studie wurde erstmals ein laborexperimenteller Zugang zum Nachweis der Wirkung von Neutralisationen auf beobachtetes Verhalten gewählt. Die experimentelle Manipulation bestand in der Hemmung bzw. der Förderung der Neutralisierungsfähigkeit einer Versuchsperson, das anschließende normwidrige Verhalten (Benutzen von Einweg-Getränkedosen) wurde im Labor beobachtet. Die Erhebung zahlreicher Kontrollvariablen, wie z.B. die Normbindung der Vpn, ermöglicht die Untersuchung der Bedingungen der Neutralisations-Verhaltens-Wirkung. Es werden die Ergebnisse zweier unabhängiger Experimente (N = 80, N = 54) vorgestellt, von denen lediglich *eines* die Kernannahme der Theorie bestätigt. Die Voraussetzungen des Neutralisationseffekts werden diskutiert.

## **Psychologie der Eigensicherung**

*Uwe Füllgrabe*

*Bildungsinstitut der Polizei Niedersachsen, Hann. Münden*

Der Beitrag behandelt ein wichtiges, aber wissenschaftlich bisher kaum untersuchtes Gebiet: die Psychologie der Eigensicherung.

Die Ermordung von Polizisten, Geiselnahmen in Gefängnissen u.ä. zeigen, dass Personen, die im polizeilichen und forensischen Bereich arbeiten, ein besonders hohes Gefährdungspotenzial durch gewaltbereite Personen haben. Die Praxis zeigt aber auch, dass es große individuelle Unterschiede in der Survivability gibt, d.h. der Fähigkeit, gefährliche Situationen rechtzeitig zu erkennen, zu vermeiden und zu überleben. Aus der Analyse konkreter Vorfälle lassen sich gemäß der Methode der kritischen Vorfälle die einzelnen Komponenten dieser Fähigkeit ermitteln: aktiver vs. passiver Lebensstil, synergistischer Denkstil vs. ICH-Zentrierung, gelassene Wachsamkeit vs. Angst, Humor, nichtsprachliche Signale der Selbstsicherheit usw. Diese Komponenten stehen in Verbindung mit bekannten psychologischen Konstrukten, z.B. Bindungsstil, Monitoring/Blunting, Kontrollüberzeugung.

Survivability kann auch gemäß einer zwischenmenschlichen Spieltheorie und Waddingtons epigenetischer Landschaft betrachtet werden, wobei die Praxis zeigt, dass Personen, die gemäß der TIT FOR TAT-Strategie handelten, Gewalt vermieden, Geiselnahmen überstanden und selbst in gewaltbereiten Umgebungen ein kooperatives Beziehungsgeflecht aufbauen konnten.

Der Beitrag zeigt die praktische Umsetzung der Survivability in einem Modell, das die Bewältigung verschiedener Gefährdungsstufen behandelt :

1. Eigensicherung zur Verhinderung des Umkippen des nichtaggressiven Zustandes einer Situation in Gewalt durch ein „Gefahrenradar“ (wobei die Gefahrenwahrnehmung durch Gefühle oder voreilige kognitive Festlegungen beeinträchtigt werden kann), Reaktionsbereitschaft usw.
2. Krisenbewältigung (bei Schusswechseln usw.) durch vorherige „Stressimpfung“, Abrufen automatisierter Verhaltensweisen
3. Bei schweren Verletzungen und Bedrohungen des Lebens: Aktivierung des psychologischen Immunsystems und schrittweises Handeln zum Entkommen aus der Gefahrensituation

## **Kriminalitätsfurcht und Strafhärte: Es kommt darauf an, ob man an „Prävention durch Strafe“ glaubt**

*Ute Gabriel & Colette Marti*

*Institut für Psychologie, Universität Bern*

Ob die Forderung nach harten Strafen Ausdruck von Kriminalitätsfurcht ist, ist eine in der kriminologischen Forschung seit langem diskutierte Frage. Die diesbezüglichen Forschungsergebnisse sind allerdings ambivalent: Manche Untersuchungen zeigen, dass die Furchtsameren härtere Strafen fordern (z.B. Sprott & Doob, 1997), in anderen Studien hingegen zeigt sich kein Zusammenhang zwischen Kriminalitätsfurcht und Punitivität (z.B. Ouimet & Coyle, 1991).

Ziel dieses Beitrages ist es nun nicht, für oder gegen diesen Zusammenhang zu argumentieren, sondern ihn mittels der Unterscheidung zwischen einer verhaltenskontrollierenden und einer ausgleichenden Wirkung von Strafen (z.B. Vidmar & Miller, 1982) zu differenzieren: Wenn eine Person davon überzeugt ist, mittels Bestrafungen könne das *zukünftige Verhalten* des Täters (und Anderer) *kontrolliert* werden, dann könnte Furcht tatsächlich ein Grund dafür sein, vergleichsweise harte Stra-

fen zu fordern und auf diese Weise die Quelle der Furcht zu verringern. Ist eine Person dagegen davon überzeugt, dass das Geschehene durch eine Bestrafung *ausgeglichen* werden könnte, dann sollte sie darum bemüht sein, die Höhe der geforderten Strafe der Tat anzupassen. Ihre persönliche Furcht davor, selbst Opfer zu werden, sollte hierauf keinen (unmittelbaren) Einfluss nehmen.

Per Fragebögen wurden 236 Personen u.a. zu ihrer Kriminalitätsfurcht und ihrer Einstellung gegenüber Strafen befragt. Punitivität wurde anhand von Fallvignetten erhoben. Erwartungskonform ist der Zusammenhang zwischen Kriminalitätsfurcht und Strafhärte in der Gruppe der Personen, die von einer verhaltenskontrollierenden Wirkung einer Bestrafung überzeugt sind, signifikant höher als bei den Personen, die von einer ausgleichenden Wirkung überzeugt sind. Strafhärte kann demnach insbesondere dann ein Ausdruck von Kriminalitätsfurcht sein, wenn man an die präventive Wirkung von Strafen glaubt.

### **Empirisch fundierte Prognosestellung im Maßregelvollzug nach § 63 StGB – EFP-63**

*Lutz Gretenkord*

*Klinik für Forensische Psychiatrie, Haina*

Bei Patienten des psychiatrischen Maßregelvollzugs (§ 63 StGB) ist jährlich zu entscheiden, ob die weitere Unterbringung noch erforderlich ist. Wesentliches Kriterium ist die Frage, ob weitere Straftaten zu erwarten sind. Die zu diesem Zweck erstellten Gutachten stützen sich in aller Regel nicht auf empirische Untersuchungsergebnisse; es gibt nur wenige einschlägige Untersuchungen, diese sind nicht gut zugänglich und nicht praxisgerecht aufbereitet. Am Beispiel einer Patientengruppe der Klinik für forensische Psychiatrie Haina (188 männliche Probanden, die 1977-1985 entlassen wurden, durchschnittlich acht Jahre Katamnesezeit) wird demonstriert, wie Ergebnisse von Follow-up-Studien für die Praxis nutzbar gemacht werden können. Von den Variablen, die einen signifikanten Zusammenhang mit der Gewaltrückfälligkeit aufweisen, wurden zehn ausgesucht, die zuverlässig zu erheben und interpretierbar sind, und die sich auch in anderen Untersuchungen als rückfallrelevant erwiesen haben. Die in diesen enthaltene Information ließ sich mit einer multiplen logistischen Regression auf vier Variablen reduzieren (Vorliegen einer Persönlichkeitsstörung, gewaltsames Vordelikt, Gewalttätigkeit während des Maßregelvollzugs und Alter bei Entlassung). Die Rückfallwahrscheinlichkeiten für die einzelnen sich hieraus ergebenden Gruppen, die zwischen 6% und 65% liegen, werden in Tabellenform dargestellt. Sie können als Basisraten bei Prognosebegutachtungen verwendet werden.

### **Gedächtnispsychologische Grundlagen der Glaubhaftigkeitsbegutachtung – Ansätze zu einer theoretischen Fundierung der “Undeutsch-Hypothese”**

*Luise Greuel*

*Institut für Psychologie und Kognitionsforschung, Universität Bremen*

Grundlage der aussagepsychologischen Glaubhaftigkeitsbegutachtung ist bekanntermaßen die “Undeutsch-Hypothese”, die letztlich die Distinktheit erlebnisbasierter Aussagen postuliert. Trotz über vier Jahrzehnten erfolgreicher Anwendung in der forensischen Praxis und kodifizierter Anerkennung dieses Grundpostulats durch die höchstrichterliche Rechtsprechung steht eine theoretische Begründung der “Undeutsch-Hypothese” nach wie vor aus.

Ein Grund für dieses Theoriedefizit mag darin bestehen, daß man sich in den vergangenen Jahrzehnten primär mit der experimentellen Validierung *singulärer* Qualitätsmerkmale erlebnisfundierter Aussagen und weniger mit den zugrunde liegenden psychologischen Prozessen befaßt hat. Ausnahmen bilden hier jene Konzeptionen, die unter Einbeziehung sozial-kommunikativer Aspekte von Täuschung oder kognitionspsychologischer Aspekte der Wirklichkeitswahrnehmung oder Realitätsüberwachung auf spezifische Charakteristika erlebnisbasierter Aussagen abheben, doch ist auch diesen Modellen letztlich die Fokussierung der phänomenalen Merkmalsebene gemeinsam. Auch sie fragen primär danach, *wie* sich Repräsentationen unterschiedlicher Wirklichkeitsstufen unterscheiden, während das *Warum* eher nachrangig behandelt wird.

Im Rahmen dieses Überblicksreferats werden Theorien und Befunde aus der Gedächtnis- und Narrationsforschung vorgestellt, wonach sich die "Undeutsch-Hypothese" sowohl über die strukturelle als auch die phänomenale Distinktheit des autobiographischen und Handlungsgedächtnisses post hoc theoretisch fundieren läßt. Unter Berücksichtigung sozial-konstruktivistischer Modelle wird darüber hinaus erörtert, daß Aussagen als verbale Rekonstruktionen subjektiver Erlebnismomente nicht losgelöst von identitäts- und bedeutungsstiftenden Prozessen innerhalb eines spezifischen kulturellen Kontextes verstanden werden können. Auf dem Hintergrund eines ganzheitlichen Verständnisses von Aussage und Aussagebegutachtung wird schließlich die Frage zu diskutieren sein, inwieweit die Aussagepsychologie als solche zur kritischen Selbstreflexion disziplinimmanenter Aussage- und Narrationstechniken und damit zur Analyse ihres eigenen Beitrags zur Aushandlung und Konstruktion prozessualer Wirklichkeiten bereit ist.

### **Etikettierungen im Jugendstrafvollzug? Wider die Gewißheiten des Labeling-Ansatzes**

*Werner Greve & Dirk Enzmann*

*Kriminologisches Forschungsinstitut Niedersachsen, Hannover*

Seit den vielzitierten Thesen von Goffmann, die „totale Institution“ Gefängnis zerstöre die Identität und insbesondere das Selbstwertempfinden der Gefangenen, gehört der „Labeling-Effekt“ institutionalisierter Strafsanktionen zum common sense der Kriminologie. Demgegenüber gibt es nur sehr wenige Studien, die mit einem längsschnittlichen Untersuchungsansatz die individuellen und sozialen Folgen einer Haftstrafe für Gefangene über die bloße Legalbewährung hinaus empirisch untersucht haben. Diese Forschungslücke ist insbesondere für jugendliche und heranwachsende Straftäter, die eine Jugendstrafe verbüßen, ein ernsthaft bedenkliches Defizit, weil hier die individuelle erzieherische Wirkung der Sanktion ausdrückliche Legitimation der Sanktion ist.

Vorgestellt werden Befunde eines laufenden Längsschnittprojektes zu den Folgen der Jugendstrafe für die persönliche und soziale Identität, die unter dem Gesichtspunkt der identitätsdestruktiven und stigmatisierenden Wirkung des Gefängnisses diskutiert werden. Tatsächlich zeigt sich ein sehr differenziertes Bild, sowohl im interindividuellen Vergleich als auch in einer intraindividuellen Perspektive im Hinblick auf verschiedene abhängige Variablen (z.B. Selbstwert, Depressivität, Prisonisierung, Normorientierung). Konsequenzen für die Interventionsangebote im Jugendstrafvollzug werden abschließend diskutiert.

## **Prognose von Täterverhalten bei Geiselnahmen**

*Everhard von Groote*

*Polizeifortbildungsinstitut, Neuss*

In Geiselnahmen stellt sich für die Polizei das Problem, unter oft minimalen Informationen Entscheidungen über Interventionen mit weitreichenden Konsequenzen treffen zu müssen. Besondere Probleme ergeben sich in der Prognose von Täterverhalten im weiteren Tatverlauf. Die psychologischen Entsprechungen dieser polizeilichen Probleme werden hier betrachtet, wobei die Prognose von Täterverhalten als Sonderfall psychologischer Diagnostik verstanden wird. Angelehnt an die Verhaltensgleichung von Westhoff und Kluck (1991) werden Variablengruppen untersucht, die Hinweise zur Prognose von Täterverhalten erstens in Bezug auf Fragestellungen zu bevorstehender Gewaltanwendung und zweitens zu Suizidalität des Täters geben. Das aus diesen Überlegungen resultierende Instrument ist als PC-gestütztes System zur Entscheidungsunterstützung realisiert. Erläutert wird der Aufbau und die Arbeitsweise dieses Programms. Neben Problemen auf der Seite der Informationserhebung und der Dateneingabe sind die interne Verrechnung der Werte und die Art der Ergebnisdarstellung von besonderem Interesse. In dem hier vorgestellten PC-System erfolgt die Informationserhebung und Dateneingabe in Form von Einschätzungen über die Ausprägung relevanter Variablen. Diese Einschätzungen werden gewonnen erstens aufgrund von Analysen bisherigen Täterverhaltens, zweitens aus Hinweisen aus der Kommunikation mit dem Täter sowie drittens aus weiteren Ermittlungsergebnissen. Die so entstehenden Werte werden nach Regeln, die durch Expertenbefragungen ermittelt wurden, verarbeitet. Für die Ergebnisausgabe wird ein kategoriales System vorgeschlagen, das neben der Prognose von Gewaltbereitschaft und Suizidalität auch eine Information über die Sicherheit der Prognose gibt. Weiterhin kann der Nutzer den Weg, auf dem die Prognose vom System erstellt wurde, nachvollziehen. Eine Validierung des Systems ist aufgrund der geringen Fallzahlen schwierig. Es werden Ansätze zur Überprüfbarkeit der Testgütekriterien diskutiert.

## **Maßgeschneiderte psychometrische Verfahren für die familienrechtliche Begutachtung**

*Wilfried Hommers*

*Lehrstuhl für Psychologie I, Universität Würzburg*

Untersucht werden die Erscheinungsformen zweier diagnostischer Strategien in der familienrechtlichen Begutachtung, wobei eine Beschränkung auf den Fall-Typ zweier streitender Eltern erfolgt.

Ad 1. „Maßgeschneidert“: Die unzulängliche Erfüllung der Strategie maßgeschneiderter Diagnostik wurde in der Literatur bemängelt, da in der sorgerechtlichen Begutachtung Verfahren angewendet wurden, die aus anderem diagnostischen Kontext stammten, z.B. der Erziehungsberatung. Daher stellte sich die Frage, wodurch sich „maßgeschneiderte“ Verfahren in der familien- oder sorgerechtlichen Diagnostik auszeichnen. Diesbezüglich werden bestehende Verfahren analysiert. Als Fazit ergibt sich, dass in dem behandelten Fall-Typ ein besonderer Fall der Vergleichsdiagnostik vorliegt.

Ad 2. „Psychometrisch“: Die allgemein anerkannte Strategie der psychometrischen Verfahren besitzt einen geradezu katastrophalen Stellenwert in der familienrechtlichen Begutachtung. Diese Strategie hat weiterhin durch die Kindschaftsrechtsreform besonderen Stellenwert erlangt. Die Kindschaftsrechtsreform hat die geforderte diagnostische Schlussweise für die sorgerechtliche Begutachtung verändert. Statt aufgrund der multi-trait-multi-method-Befunde den vermeintlich besseren Elternteil angesichts einer quasi zwangsweise anstehenden Entscheidung für die Alleinige Elterliche Sorge eines

Elternteils auszumachen, geht es nun um die Prüfung der Nullhypothese einer im Prinzip trotz Trennung oder Scheidung fort-bestehenden Gemeinsamen Elterlichen Sorge. Die psychodiagnostische Alternative tritt nur ein, wenn die Befunde dafür Anlass geben, von der Widerlegung der Nullhypothese der Gemeinsamen Elterlichen Sorge auszugehen. Die damit verschobene Beweislast könnte sich in einem veränderten statistischen Risiko für ein Präferenzergebnis zugunsten eines Elternteils in der durchzuführenden sorgerechtlichen Vergleichsdiagnostik ausdrücken. Wenn früher etwa auf dem 20%-Niveau eine festgestellte Präferenz zählen durfte, weil man Begründungen in eine Richtung finden musste, könnte eine Präferenzfeststellung nun nur noch mit dem 5%-Niveau begründet werden, weil erst dann nach allgemein verbreiteten Maßstäben die Nullhypothese nicht mehr zutrifft.

Die sich ergebenden Qualitätsanforderungen an eine sorgerechtliche Vergleichsdiagnostik betreffen die psychometrischen Eigenschaften der einzelnen primären Präferenzmaße, die auf Prüfungen von Differenzen beruhen, und die sekundäre integrierende Gesamtbewertung der Einzelbefunde. Es stellt sich daher die Frage, ob und inwiefern die verwendete psychologische Verfahren zur Prüfung und eventuellen Widerlegung der Nullhypothese geeignet sind. Diese Fragestellung soll an dem Beispiel der Erfassung der emotionalen Beziehungen zwischen Kind und Eltern verdeutlicht werden.

Ein Verfahren für Kinder und ein Verfahren für Eltern wurden bezüglich Reliabilität und Validität untersucht. Dabei stellte sich heraus, dass für die angestrebte Vergleichsdiagnostik in besonderer Weise das Problem der Reliabilität von Differenzen bedeutsam wird.

### **Zum Einfluss des Tatmotivs auf eine Voraussetzung der strafrechtlichen Verantwortungsreife**

*Wilfried Hommers & Martin Lewand  
Lehrstuhl für Psychologie I, Universität Würzburg*

Die Entwicklung der Unrechtserkenntnisfähigkeit aus der Sozialbindung könnte als Orientierung an der Ordnungsgefährdung operationalisiert werden. Zwei Querschnittsstudien mit dem Ansatz des Gedanken-Paradigmas untersuchten die Wirkung des Tatmotivs auf die Berücksichtigung der Ordnungsgefährdung im Urteil. 263 Probanden zwischen 7 und 16 Jahren beurteilten auf der Grundlage des Tatmotivs (altruistisch bzw. egoistisch) und der moralischen Gedanken eines Einbrechers dessen verdiente Strafe. Die Gedanken wurden systematisch variiert und bezogen sich inhaltlich auf die ersten vier Kohlbergstufen (Entdeckungsfahr, Meinung der Frau, vermutetes Handeln eines Freundes, Ordnungsgefährdung) und auf die moralische Pflicht der Wiedergutmachung. Das psychometrisch gesicherte Unterscheiden der variierten Gedanken und die Dominanz eines dieser Effekte wurden als Zeichen von moralischem Verständnis gewertet. Bereits 10-Jährige waren unabhängig vom Tatmotiv zur Beachtung der Variation von Gedankeninhalten über die Ordnungsgefährdung fähig. Weiterhin dominierte ab einem Alter von 10 Jahren die Beachtung der Ordnungsgefährdung alle anderen Kohlbergstufeninhalte. Bei 14-Jährigen dominierte zudem die Ordnungsgefährdung auch die Wiedergutmachungsabsicht. Die prozentuale Beachtung des Gedankeninhalts über das vermutete Handeln des Freundes war zu Lasten der Meinung der Frau bei der Beurteilung eines egoistischen Tatmotivs größer als bei der Beurteilung eines altruistischen Tatmotivs. Die Implikationen de lege ferenda und de lege lata wären zu diskutieren.

## **Soziale Unterstützung im Strafvollzug: Der Einfluss sozialer Beziehungen auf das Befinden und die Normorientierung junger Männer**

*Daniela Hossler*

*Kriminologisches Forschungsinstitut Niedersachsen, Hannover*

Befunde über erhöhte Prävalenzraten psychischer Störungen und ein hohes Stresserleben bei Inhaftierten werfen die Frage auf, welche Faktoren dazu beitragen können, die psychischen Belastungen der Inhaftierten zu reduzieren und negativen Haftfolgen entgegenzuwirken. Die vorliegende Studie untersucht, welchen Einfluss verschiedene Formen sozialer Unterstützung (wahrgenommene Unterstützung, erhaltene Unterstützung, Netzwerkkontakte) und die sozialer Beziehungen im Vollzug auf die psychische Befindlichkeit (Selbstwert, Depressivität) und die Normorientierung von Inhaftierten im Jugendstrafvollzug nehmen. Mittels mündlicher standardisierter Interviews wurden 1998 insgesamt 295 Jugendliche und junge Erwachsenen im Alter von 16 bis 24 Jahren in fünf norddeutschen Jugendvollzugsanstalten befragt. Die Ergebnisse zeigen, dass die wahrgenommene Unterstützung und enge Beziehungen zu den Mitinsassen mit niedrigerer Depressivität und höherem Selbstwert in Zusammenhang stehen. Ein häufiger Unterstützungserhalt und konfliktträchtige Beziehungen zum Vollzugspersonal gehen mit einem schlechteren Befinden einher. Während sich Kontakte mit Familienangehörigen und ein positives Anstaltsklima in einer höheren Normorientierung der Inhaftierten bemerkbar machen, stehen enge Beziehungen zu den Mitinsassen und konfliktträchtige Beziehungen zu den Vollzugsbediensteten mit einer geringeren Normorientierung in Zusammenhang. Die Ergebnisse verdeutlichen die ambivalenten Auswirkungen sozialer Unterstützung und bieten Ansatzpunkte für die Praxis im Jugendstrafvollzug und die Strafvollzugsforschung.

## **Dimensionen retributiver Gerechtigkeit**

*Jörg Hupfeld*

*Institut für Psychologie, Universität Bern*

Während zur Frage der Dimensionalität distributiver Gerechtigkeitsüberlegungen empirische Untersuchungen vorliegen (vgl. Reis, 1984), fehlen bislang vergleichbare Strukturanalysen für den Bereich der retributiven Gerechtigkeit. Bei der Bewertung einer Straftat können z.B. relativistische oder universalistische Standards angewendet werden (vgl. Forsyth, 1980), die Überlegungen können primär fürsorge- oder gerechtigkeitsorientiert sein (vgl. Gilligan, 1982) und sich auf die Absichten des Straftäters oder die Konsequenzen der Straftat beziehen (vgl. Haidt & Baron, 1996). Berichtet werden die empirischen Befunde einer schriftlichen Befragung zu Reaktionen der Bevölkerung auf Rechtsverstöße. Jede der 691 teilnehmenden Person beurteilte anhand eines standardisierten Fragebogens eine von insgesamt drei Fallgeschichten (Raubüberfall, Körperverletzung, Betrug). Im Rahmen der Untersuchung wurde zum einen der Frage nachgegangen, inwieweit die in der Forschung diskutierten Gerechtigkeits- und Moraloientierungen tatsächlich empirisch voneinander abgrenzbar sind bzw. welche allgemeine Struktur ihnen zugrunde liegt. Zum anderen wurden Zusammenhänge zwischen den grundlegenden Gerechtigkeitsdimensionen und Sanktionspräferenzen (Sanktionshärte, Akzeptanz alternativer Rechtsfolgen) analysiert.



## **Die Verwendung ereigniskorrelierter Potentiale im Rahmen der Psychophysiologischen Aussagebegutachtung**

*Prisca Jager & O. Berndt Scholz*

*Psychologisches Institut, Universität Bonn*

Anliegen: Die Forschung im Rahmen der Psychophysiologischen Aussagebegutachtung untersucht neben herkömmlichen Verfahren (Kontrollfragen- und Tatwissenverfahren) die Möglichkeit der Verwendung ereigniskorrelierter Potentiale zur Täuschungsaufdeckung. Exemplarisch wird die Leistungsfähigkeit der P300-Komponente, die im Oddball-Paradigma auf seltene, bedeutsame Reize hin auftritt, zur Täuschungsaufdeckung erprobt. Annahme: Es wird angenommen, dass die P300 bei zu Recht beschuldigten Probanden bei Testreizen eine größere Amplitude aufweist als bei irrelevanten Reizen. Die maximale Amplitude der P300, die bei einem zu Unrecht beschuldigten Probanden auf die Testreize hin auftritt, soll erwartungsgemäß kleiner oder gleich derjenigen bei irrelevanten Reizen sein. Stichprobe: Dreiundvierzig männliche Probanden im Alter zwischen 20 und 43 Jahren nahmen an der Untersuchung teil. Methodik: Im Rahmen des „Bonner Szenarios“ wurden 23 Probanden bezüglich eines Films, der ein Sexualdelikt zeigt, aufgrund schlechter Beobachtungsleistung zu Recht beschuldigt, sich inadäquat mit den Filminhalten beschäftigt zu haben. Weitere 20 Probanden wurden zu Unrecht beschuldigt. Ein EEG wurde von Fz, Cz und Pz abgeleitet. Seltene, aufgabenrelevante Informationen aus dem Film stellten die Zielreize dar. Neue Items wurden als irrelevante Reize verwendet. Die sogenannten Testreize, die in Bezug auf die Schuldbedingung diagnostischen Wert hatten, stellten tatrelevantes Wissen dar. Ergebnis: Ein Bootstrap-Algorithmus, mit dessen Hilfe Differenzen zwischen Amplitudenmaxima verglichen wurden, erbrachte Indizes, anhand derer nicht sicher zwischen den Schuldbedingungen differenziert werden konnte ( $AUC \pm SE = .538 \pm .096$ ). Schlussfolgerungen: Es zeigte sich, dass anhand der P300-Amplituden keine korrekte Klassifikation vorgenommen werden konnte. Es wird diskutiert, inwieweit die sexuellen Inhalte des Stimulusmaterials die Ergebnisse verzerren.

## **Qualitätssicherung kriminalpolizeilicher Ermittlungspraxis durch Verhaltensbeobachtung – erste Ergebnisse aus dem Pilotprojekt BEOCOP**

*Claudia Kestermann & Luise Greuel*

*Institut für Psychologie und Kognitionsforschung, Universität Bremen*

Aus der polizeilichen Praxis wird zunehmend ein Bedarf an psychologischer Expertise hinsichtlich der Entwicklung wissenschaftlich fundierter Methoden und Instrumente zur Sicherung der Qualität kriminalpolizeilichen Handelns bekundet; dies gilt insbesondere für die Beobachtung, Begleitung und Analyse von Strafprozessen als einem zusätzlichen Erkenntnisfeld im Bereich der Organisierten Kriminalität und bei weiteren komplexen Straftatbeständen. Das Pilotprojekt BEOCOP (Beobachtung und Codierung von Prozeßverläufen), welches durch das EU-Programm *Falcone* co-finanziert wird, greift dieses Anliegen auf mit dem Ziel, ein standardisiertes, modulares Beobachtungsinstrument zu entwickeln. Zur Realisierung dieses Projektziels werden zunächst Aktenanalysen abgeschlossener und in diesem Kontext relevanter Fälle und darüber hinaus Interviews mit Experten aus Kriminalpolizei und Staatsanwaltschaft durchgeführt, um anhand der Ergebnisse die zentralen Bedingungsgrößen für Bedeutung, Gerichtsverwertbarkeit und -relevanz bestimmter Ermittlungsmaßnahmen, ihrer Durchführung und Dokumentation zu identifizieren. Die resultierenden Erkenntnisse finden Eingang in die

Entwicklung von Beobachtungsmodulen, die sich zum einen auf die Selbstbeobachtung (im Sinne einer systematischen Selbstkontrolle) beziehen und zum anderen auf die Beobachtung von Strafverfahren mit Konzentration auf die Psychodynamik zwischen den Prozeßbeteiligten und die gerichtliche Beweiswürdigung. Die ersten Ergebnisse dieses Projektes werden dargestellt, die bis dato entwickelten Beobachtungsmodule präsentiert und hinsichtlich ihrer methodischen Fundierung, ihrer kriminaltaktischen Relevanz sowie ihrer praktischen Implikationen diskutiert.

## **Vermutungen von Juristen und Rechtslaien zur Wirkung von Gerichtsfernsehen**

*Edgar W. Klinger & Heiner Gehring*

*Fachbereich Psychologie und Gesundheitswissenschaften, Universität Osnabrück*

Das Für und Wider einer Fernsehberichterstattung aus Gerichtsverhandlungen wird seit etwa 50 Jahren in rechtswissenschaftlichen Fachpublikationen im deutschen Sprachraum diskutiert. Die Entscheidung des Bundesverfassungsgerichts zur Zulassung von Gerichtsfernsehen hat zu Beginn dieses Jahres die Diskussion neu entfacht und in die Tagespresse getragen. In der Diskussion überwiegen Stellungnahmen von Juristen, in denen das Gerichtsfernsehen abgelehnt wird. Begründet wird diese Position zumeist mit der Vermutung, dass das Wissen, gefilmt zu werden, das Verhalten der Beteiligten im Gerichtssaal beeinflusse und damit die Wahrheitsfindung erschwere. Die Gerichtsverhandlung verkomme zu einer Show-Veranstaltung, in der Beschuldigte und Zeugen sich gehemmt fühlen müssten, über Einzelheiten ihres Privatlebens zu berichten, und in der sich die Beschuldigten an den Pranger gestellt sähen. Diese Vermutungen klingen plausibel, sind indes nicht durch empirische Studien belegt. In welchem Maße geben nun diese öffentlich vorgetragenen Stellungnahmen das Meinungsbild von Richtern, Rechtsanwälten und Rechtslaien zu Wirkungen von Gerichtsfernsehen wieder? Der Beitrag stellt die Ergebnisse von Umfragen vor, in denen nach solchen Einschätzungen gefragt wurde. Sie zeigen, dass Juristen negative Effekte eines Gerichtsfernsehens in stärkerem Maße vermuten als Rechtslaien. Ferner erwarten die befragten Juristen deutlichere Auswirkungen auf das Verhalten anderer Verhandlungsbeteiligter als auf das eigene Verhalten. Die Einschätzungen deutscher Juristen werden jenen amerikanischer Juristen gegenübergestellt, die seit etwa zehn Jahren Erfahrungen mit der Arbeit des Gerichtsfernsehensenders „Court-TV“ sammeln konnten. Abschließend werden Forschungsperspektiven skizziert, deren Umsetzung zur Gewinnung gesicherter Erkenntnisse über die Wirkungen von Gerichtsfernsehen beitragen kann.

## **Ergebnisse einer Bürgerbefragung zur Akzeptanz von polizeilichen Kontroll- und Eigensicherungsmaßnahmen**

*Clemens Lorei*

*Redaktion Polizei und Wissenschaft, Frankfurt*

Polizeiliches Handeln im Kontakt mit Bürgern und Rechtsbrechern erfordert mitunter eine Gradwanderung zwischen Bürgerfreundlichkeit und notwendigen Maßnahmen, um die Sicherheit der eingesetzten Polizisten zu gewährleisten. Infolge der Tötung von 8 Polizisten im Jahr 2000 wurden sogenannte "amerikanische Verhältnisse" erwogen, das heißt Kontrollbedingungen und Maßnahmen, die einer verstärkten Sicherheit der eingesetzten Polizeikräfte bewirken sollen. Unabhängig von einer

taktischen Bewertung dieser Eigensicherungsmaßnahmen scheint es im Sinne sowohl der Bürgerfreundlichkeit als auch der Kooperationsbereitschaft von bürgerlicher Seite aus wichtig, die Akzeptanz von entsprechenden polizeilichen Kontroll- und Eigensicherungsmaßnahmen zu erfragen. Es wurde in einer nicht repräsentativen Befragung von 431 hessischen Bürgerinnen und Bürgern versucht, mittels Fragebogen eine Bewertung polizeilicher Kontroll- und Eigensicherungsmaßnahmen, die Einstellung gegenüber der Polizei und die psychologische Reaktanz zu erfassen. Die Ergebnisse zeigen, dass Bewertungen verschiedener Maßnahmen interindividuell unterschiedlich variieren. Die mittleren Beurteilungen zeigen trotz überwiegend positiver Einstellung gegenüber der Polizei teilweise deutliche Ablehnung. Fast alle Kontroll- oder Eigensicherungsmaßnahmen bei Fußgängerkontrollen werden geringfügig, aber statistisch signifikant negativer bewertet als bei Fahrzeugkontrollen. Insgesamt zeigten sich höchstens geringe korrelative Zusammenhänge zwischen verschiedenen Maßnahmen und der Einstellung gegenüber der Polizei und der Reaktanzneigung. Es wird deshalb daraus geschlossen, dass die polizeilichen Maßnahmen nicht auf Grund einer allgemeinen Haltung bewertet werden. Der Versuch, die verschiedenen Maßnahmen faktorenanalytisch zu gruppieren, legt nahe, dass neben einer starken allgemeinen Bewertung 4 Gruppen von polizeilichen Handlungsweisen existieren, die sich inhaltlich mehr oder minder gut interpretieren lassen und Rückschlüsse über die Beurteilungen erlauben.

## **Psychologische Faktoren der Arbeitssicherheit bei der Polizei am Beispiel Schutzwesten**

*Clemens Lorei*

*Redaktion Polizei und Wissenschaft, Frankfurt*

Polizeiliche Einsätze sind seit jeher und global mit dem Risiko verbunden, von Rechtsbrechern verletzt oder getötet zu werden. Als im Jahr 2000 acht Polizisten durch Schuss- oder Stichwaffen im Dienst ermordet wurden, führte dies zu einer entsprechenden Diskussion über Möglichkeiten zur Steigerung der Sicherheit von Polizeibeamten. Im Sinne der Schutzzielhierarchie stellt die Nutzung von Schutzwesten zur Verringerung der Gefährdung eine für die Polizei naheliegende Option dar. Dabei lassen sich verschiedene psychologische Einflussfaktoren auf das Trageverhalten annehmen, die aus industriellen Bereichen bereits bekannt sind. In einer bundesländerübergreifenden, nicht repräsentativen Befragung von 1679 Polizistinnen und Polizisten mittels Fragebogen wurde versucht, Faktoren aus den Bereichen persönliche Gewalterfahrung, Gefahrenkognition, Funktionalitätsüberzeugung, Annahmen über Gefährdung auf Grund einer Schutzweste, Tragekomfort, Verfügbarkeit und Sicherheitskultur zu erfassen. Die Ergebnisse zeigen, dass sowohl die angegebene Nutzung der Schutzausrüstung als auch die erfragten Einflussfaktoren interindividuell deutlich variieren. Es konnten geringe bis mittlere korrelative Zusammenhänge zwischen verschiedenen Faktoren und dem angegebenen Trageverhalten gefunden werden. Besonders variiert die empfundene Beweglichkeitseinschränkung, der Aufwand, der notwendig ist, um diese Ausrüstung zu tragen, und das Trageverhalten der Kollegen mit der Nutzung. Aus den Ergebnissen wird geschlussfolgert, dass das Tragen von Schutzwesten bei der Polizei multifaktoriell beeinflusst wird und auf komplexen und individuellen Entscheidungen basiert. Entsprechend wird ein mehrdimensionales psychologisches Vorgehen zur Steigerung der Tragegeräte empfohlen.

## **Entwicklung und Persönlichkeit von Hooligans: Komparative Fallanalysen**

*Friedrich Lösel & Thomas Bliesener*

*Institut für Psychologie, Universität Erlangen-Nürnberg, Institut für Psychologie, Universität Kiel*

Hooligan-Probleme haben international grosse Aufmerksamkeit erreicht. Dabei ist nicht nur die Berichterstattung in den Massenmedien stark an Stereotypen orientiert. Auf Grund von Zugangsproblemen zu diesem Feld gibt es nur wenig Daten, die über lokale sportsoziologische Analysen mehr oder weniger aggressiver Fangruppen hinausgehen. Vor diesem Hintergrund wurden in dem vom Bundesinnenministerium geförderten Projekt 33 bundesweit rekrutierte Personen des "Harten Kerns" der Hooligan/Szene eingehend untersucht. Unter anderem führten wir ausführliche Interviews, Persönlichkeits- und Leistungstests sowie klinisch-psychologische Screenings durch. Die Daten wurden im Sinne einer systematischen komparativen Kasuistik aufbereitet. Trotz der bei manchen Merkmalen erheblichen Unterschiede zwischen den Hooligans zeigt sich insgesamt ein deutlich überwiegendes Entwicklungsmuster mit Belastungen und Problemen in der Herkunftsfamilie, in der Schule, im Beruf, in der Persönlichkeit, im Substanzenmissbrauch und in der Kriminalität. Die Ergebnisse sprechen dafür, die Phänomene des Hooliganismus nicht zu mystifizieren und die Prävention und Intervention stärker an der allgemeinen Längsschnittforschung über die Gewalt junger Menschen zu orientieren.

## **Kann man durch das Wissen über Realkennzeichen glaubhafter lügen?**

*Friedrich Lösel & Nicole Raichle*

*Institut für Psychologie, Universität Erlangen-Nürnberg*

Obwohl die Analyse inhaltlicher Realkennzeichen glaubhafter Aussagen ursprünglich für kindliche Zeugen entwickelt wurde, hat man sie international teilweise auch auf Erwachsene übertragen. Dabei stellt sich in besonderem Masse die Frage, inwieweit Personen mit Kenntnis der Realkennzeichen glaubhafter lügen können. In unserem Beitrag wird dieser Frage nachgegangen.

In einem mehrfaktoriellen Versuchsplan wurden wahre und erfundene Erlebnisse von Personen präsentiert, die ein Training über Realkennzeichen erhielten oder keine Kenntnisse in diesem Bereich hatten. Die auf Video aufgezeichneten Geschichten wurden von Laien, Psychologiestudenten und Experten hinsichtlich ihrer Glaubhaftigkeit beurteilt. Die Experten waren am besten in der Lage, die falschen Geschichten untrainierter Erzähler zu identifizieren. Jedoch wurden sowohl bei den Laien als auch den Experten die falschen Aussagen der trainierten Erzähler für glaubwürdiger befunden als die falschen Geschichten der untrainierten. Die Trefferquoten der Laien und Psychologiestudenten unterschieden sich zumeist nicht signifikant. Auch bei den wahren Geschichten bestanden insgesamt weniger Unterschiede. Die Ergebnisse werden hinsichtlich der empirischen Validität der Realkennzeichen und der forensischen Praxis diskutiert.

## **Explorationstechnik – Einhaltung von Vernehmungsprinzipien zur Gewinnung valider Daten**

*Charlotte Mohrbach*

*Gesellschaft für Forensische Aussagepsychologie, Bonn*

Mit der BGH-Entscheidung vom 30.07.1999 sind die methodischen und formalen Standards festgeschrieben worden, die bei der aussagepsychologischen Glaubhaftigkeitsbegutachtung beachtet werden müssen. In Veröffentlichungen darüber ist insbesondere das Prinzip der Hypothesenprüfung verknüpft mit einer Forderung nach einem lückenlosen Transkript der auf Tonträger festgehaltenen Exploration zur Sache in den allgemeinen Fokus der Aufmerksamkeit geraten. Dabei ist unserer Ansicht nach ein wesentliches Instrument der aussagepsychologischen Begutachtung weitgehend aus dem Blickfeld gerückt: die Technik der *aussagepsychologischen Exploration zur Sache*. Die Analyse des uns vorliegenden Gutachtenmaterials, welches ausnahmslos vollständige Wortprotokolle (Tonbandtranskripte) enthielt und somit formal den geforderten Standards entsprach, ergab immer wiederkehrende Mängel, mit denen sich dieser Beitrag kritisch auseinandersetzt (z.B. ergebnisorientierte Steuerung des Explorationsverlaufs, mangelnde Flexibilität des Befragers, Evozieren und selektives Verstärken von Realzeichen mit der Folge der zirkulären Argumentation im Befund). Es werden Aspekte herausgestellt, an denen sich eine regelrechte Sachexploration zur Vermeidung von Fehlentscheidungen orientieren muß (Einstellung auf entwicklungspsychologische und kognitive Bedingungen beim Befragten; Berücksichtigung von Gedächtnisgesetzmäßigkeiten; Umsetzung vernehmungpsychologischer Grundprinzipien). Dabei wird der Nutzen eines intensiven Trainings anhand der empirisch gewonnenen Erfahrungen aus Vernehmungstrainings mit Kriminalbeamten exemplifiziert. Fazit: Intensives Training und ständige Selbstkontrolle sind die wesentlichen Voraussetzungen für eine adäquate und valide Datenerhebung. Die Aufzeichnung auf Tonträger dient somit nicht nur der Transparenz, sondern sollte auch zur kritischen Selbstkontrolle des Gutachters genutzt werden. Das Ziel ist eine ausgefeilte Befragungstechnik – die Kunst, eine Aussage zu gewinnen, die eine verwertbare Grundlage für die Anwendung der kriterienorientierten Aussageanalyse abgibt.

## **Signaturen von Sexualmördern**

*Andreas Mokros & Stephan Harbort*

*Bergische Universität - Gesamthochschule Wuppertal und Polizeipräsidium Düsseldorf*

Der Begriff der *Signatur* bezeichnet im Hinblick auf sexuelle Tötungsdelikte individuelle und ritualhafte Handlungssequenzen, die innerhalb der Tatbegehung keinen strategischen Charakter besitzen. Solche Verhaltensweisen, wie etwa das Zufügen von Brand- oder Bißverletzungen, gehen über das zur Begehung des Mordes Notwendige hinaus. Weil der *Modus operandi* als situationsabhängig gilt, gewinnt das Konzept der Signatur im Hinblick auf die Identifikation von Straftatserien (*linking*) und die Vorhersage wahrscheinlicher Tätereigenschaften (*profiling*) an Bedeutung. Die Verfahrensakten solcher Probanden (N = 76) werden inhaltsanalytisch ausgewertet, die in Deutschland (einschließlich der ehemaligen DDR) im Zeitraum von 1915 bis 2000 wegen mindestens zwei unabhängig voneinander begangener sexuell motivierter Tötungsdelikte abgeurteilt worden sind. Tathandlungen, Hintergrundeigenschaften und sexualanamnestische Informationen werden extrahiert. In einer Diskriminanzanalyse werden diejenigen Extremgruppen miteinander verglichen, die durchgängig bzw. in der Mehrzahl ihrer Delikte eine Signatur aufweisen (n = 30) und die in keinem ihrer Delikte eine Signatur zeigen (n = 28). Basierend auf den 26 Variablen über Hintergrundeigenschaften (alle binär kodiert bis

auf Alter zum Zeitpunkt der ersten Tat und IQ) erlaubt die Diskriminanzfunktion eine korrekte Gruppenzuordnung für 57 Probanden (= 98,3%). Eine Betrachtung der Diskriminanzkoeffizienten zeigt, dass folgende Variablen in besonderem Maße bei den Signatur-Tätern auftreten: Berufstätigkeit zum Zeitpunkt der Taten, Sadismus im Sinne einer Störung der Sexualpräferenz, Vorstrafe(n) für Körperverletzungs- und Rohheitsdelikte. Folgende Variablen kommen besonders häufig bei den Tätern ohne Signatur vor: einen Lebenspartner zu haben, eine Berufsausbildung abgeschlossen zu haben, sozial integriert zu sein. Die Implikationen für die Vorhersage von Tätereigenschaften aus Tatmerkmalen bei sexuellen Tötungsdelikten werden diskutiert.

## **Der Verlauf von Strafverfahren wegen sexuellen Mißbrauchs an Kindern –Ergebnisse einer Aktenanalyse**

*Katrin Müller<sup>1</sup>, Renate Volbert<sup>2</sup> & Detlef Busse<sup>2</sup>*

*<sup>1</sup>Cornell University, Ithaca, USA, <sup>2</sup>Institut für Forensische Psychiatrie, Freie Universität Berlin*

Vorgestellt werden Ergebnisse einer Analyse aller Ermittlungs- und Strafverfahren wegen sexuellen Mißbrauchs von Kindern, die 1991 von der Staatsanwaltschaft Berlin eingetragen wurden; es handelt sich dabei um 1085 Verfahren. Die sich in dieser Hellfeld-Stichprobe ergebenden Charakteristika der Täter-Opfer-Beziehungen und der Tathandlungen werden referiert und entsprechende qualitative Unterschiede zu Ergebnissen aus Dunkelfeldbefragungen diskutiert. In Dunkelfelduntersuchungen werden mehr schwerwiegende und multiple sexuelle Übergriffe durch dem Opfer nahestehende Personen beschrieben, während bei den polizeilichen Anzeigen Einzeldelikte durch unbekannte Täter dominieren. Zu einer strafrechtlichen Sanktion führten rund 12 % aller analysierten Anzeigen; bezogen auf die Gruppe der ermittelten Tatverdächtigen endeten 28 % der Verfahren mit einer strafrechtlichen Sanktionierung. Wenn es zu einer Verurteilung kam, war die Freiheitsstrafe die am häufigsten ausgesprochene Sanktion, dabei wurde der Strafraum etwa zu einem Drittel ausgeschöpft.

## **Zum Problem der Trefferquoten bei der Anwendung inhaltlicher Glaubhaftigkeitsmerkmale**

*Susanna Niehaus*

*Institut für Psychologie, Universität Potsdam*

Vorliegende Untersuchungen zur Einschätzung des Wahrheitsgehaltes von Aussagen kommen überwiegend zu dem Ergebnis, dass Beurteiler selten dazu in der Lage sind, überzufällig hohe Trefferquoten zu erzielen. Auch Untersuchungen, in denen die Beurteiler explizit dazu angehalten waren, ihre Entscheidung auf der Grundlage der merkmalsorientierten Inhaltsanalyse zu treffen, führten kaum zu zufriedenstellenden Ergebnissen. Als Erklärung dieses Phänomens wurden methodische Argumente wie beispielsweise ein zu geringer Umfang des Stimulusmaterials angeführt. Möglicherweise bieten aber auch unzureichende Trainingsmaßnahmen in der Anwendung der merkmalsorientierten Inhaltsanalyse und daraus folgende Mißverständnisse und Anwendungsfehler eine Erklärung. Im Rahmen einer größer angelegten Studie zur differentiellen Validität der merkmalsorientierten Inhaltsanalyse wurde u.a. auch der Frage nachgegangen, inwieweit umfangreiche Trainingsmaßnahmen das Verständnis der Beurteiler erhöhen und in der Folge auch zu höheren Trefferquoten beitragen können. In einem mehrwöchigen Intensivtraining erlernten 9 Beurteiler die Anwendung der merk-

malsorientierten Inhaltsanalyse. Die Effektivität des Trainings wurde mittels anwendungsorientierter Tests kontrolliert und durch Inter- und Intra-Rater-Reliabilität dokumentiert. An die Intensivtrainingsphase schloß sich eine Auswertungsphase an, in der die Beurteiler neben der Einschätzung der Aussagequalität kindlicher Zeugenaussagen unterschiedlichen Wahrheitsgehaltes (N = 160) die Aufgabe hatten, jeweils eine abschließende Glaubhaftigkeitsentscheidung zu treffen. Diskutiert werden die erzielten Trefferquoten der trainierten Beurteiler vor dem Hintergrund des Trainingserfolgs, der Reliabilität und Validität der erhobenen inhaltlichen Merkmale sowie der Trefferquoten naiver Beurteiler und diskriminanzanalytischer Auswertungen.

### **Dimensionale Struktur der Reaktionen auf normabweichendes Verhalten**

*Margit Oswald, Stefan Klug, Jörg Hupfeld & Ute Gabriel*  
*Institut für Psychologie, Universität Bern*

Normabweichendes, speziell strafrechtswidriges Verhalten löst bei Personen in der Regel nicht nur negativ gefärbte Emotionen und moralische Reaktionen aus, sondern auch Vorstellungen darüber, wie und mit welchen Zielen auf den Normbruch reagiert werden sollte. Wie diese verschiedenen Reaktionen in Beziehung zueinander stehen, wurde bisher kaum untersucht. In einer Fragebogenstudie wurden Personen befragt, wie sie auf konkrete Straftatschilderungen reagieren und zwar hinsichtlich ihrer Bereitschaft zur Normrelativierung (Forsyth, 1985), der wahrgenommenen Bedrohung des gesellschaftlichen Zusammenhalts (Tyler & Boeckmann, 1997), der präferierten Strafziele (Endres, 1992), der geforderten Strafhärte (Oswald, 1994) und der eingenommenen Mikro- oder Makro-Perspektive (Miller & Vidmar, 1981). Der Fragebogen wurde einer Zufallsstichprobe deutschsprachiger EinwohnerInnen der Stadt Bern (CH) zugeschickt. Bei einer Rücklaufquote von 45,2 % konnten 357 Fragebögen ausgewertet werden. Die Ergebnisse der Studie geben Auskunft über bestehende Unterschiede der geforderten Strafhärte, vor allem aber über die dimensional Strukturen, die unabhängig von der Strafhärte dimension die Reaktionen auf normabweichendes Verhalten bestimmen. Eine zusätzlich in Erlangen (D) durchgeführte Parallelstudie belegt die Stabilität der gefundenen Strukturen.

### **Emotionen im Prozeß der Glaubwürdigkeitsbeurteilung**

*Kerstin Panhey, Günter Köhnken & Frank Eggert*  
*Institut für Psychologie, Universität Kiel*

Dieser Beitrag beschäftigt sich mit der Rolle von Emotionen in der Übermittlung und Rezeption wahrer und falscher Aussagen. Zumeist wird Emotion – in Form von Nervosität oder Aufgeregtheit – als ein Hinweis für die Unglaubwürdigkeit einer Aussage gehalten. Ausgehend von der Theorie der kognitiven Kapazität von Ellis und Ashbrook wurde eine Hypothese untersucht, nach der Emotionen vom Sender zur Tarnung falscher Aussagen eingesetzt werden können: Indem der Sender den Empfänger durch induzierte Emotionen beschäftigt, stört er dessen rationale Analyse des Gehörten und erhöht die Chancen, daß seine Lüge unentdeckt bleibt. In der Studie sahen und/oder hörten 58 Versuchspersonen auf Video aufgenommene wahre (d.h. tatsächlich erlebte) und falsche (d.h. ausgedachte) Geschichten. Darüber hinaus wurde eine Gruppe (N=17) kanadischer Versuchspersonen ohne Deutschkenntnisse

untersucht. Mit Hilfe eines selbstentwickelten Fragebogens wurden globale und differenzierte Einschätzungen der Emotionalität von Sender und Empfänger und Fragen zur Glaubwürdigkeit der Geschichte erhoben. Die Ergebnisse zeigen einen moderierenden Einfluß von Emotionen auf die Glaubwürdigkeit der Geschichten. So wurden wahre Geschichten als emotionaler eingeschätzt als falsche. Die Emotionalität des Empfängers korrelierte hoch mit der des Senders und war nicht unabhängig vom Glaubwürdigkeitsurteil. Auch in der Untergruppe fremdsprachiger Versuchspersonen, die die Geschichten inhaltlich nicht verstehen konnten, bestätigte sich, daß emotionalere Aussagen mit höherer Wahrscheinlichkeit für wahr gehalten werden. Die Ergebnisse bestätigen zwar nicht, daß Emotionalität von Sendern zur Tarnung falscher Aussagen eingesetzt wird. Es konnte aber nachgewiesen werden, daß Emotionen die Darstellung wahrer und falscher Aussagen durch den Sender und ihre Bewertung durch den Empfänger beeinflussen.

### **Multimethodale Diagnostik bei forensisch-psychologischen Begutachtungen**

*Ernst Plaum, Katholische Universität Eichstätt*

Ungeachtet einer praktischen Realisierung wird in der Klinischen Psychologie seit langem eine multimethodale (und "multimodale") Diagnostik gefordert. Bei forensischen Begutachtungen überwiegen offenbar Vorgehensweisen, die den Schwerpunkt auf ganz bestimmte methodische Zugänge legen, im Extremfall sich auf nur einen Ansatz, etwa psychodiagnostische Interviews, beschränken. Wenn es um die Untersuchung von Straftätern geht, hat man mit massiven Verfälschungstendenzen zu rechnen; so wäre es z.B. naiv, anzunehmen, ausschließlich über Gespräche und eventuell Fragebogenverfahren könnten beispielsweise Aggressionsneigungen valide erfaßt werden. Auch wenn Gutachter in subjektiver Selbstüberschätzung meinen, den Wahrheitsgehalt der Äußerungen ihrer Probanden beurteilen zu können, bleibt das Problem der personalen Realität derselben bestehen. Eine multimethodale Diagnostik mag zwar ebenfalls keine ideale Lösung bringen, verringert aber die Wahrscheinlichkeit von Fehleinschätzungen. Zum einen bietet bereits der dabei erforderliche Zeitaufwand eine gewisse Gewähr dafür, daß die Probanden systematische Verfälschungen nicht konsistent durchzuhalten vermögen. Zum anderen sollte eine multimethodale Strategie auch Verfahren einschließen, die einen eher indirekten Zugang und geringere Verfälschungsmöglichkeiten bieten; hier sind die umstrittenen "projektiven" Techniken zu nennen. Zweifellos wäre es unverantwortlich, allein oder schwerpunktmäßig auf Grund derselben Entscheidungen zu treffen; bei der diagnostischen Urteilsbildung bietet jedoch ein multimethodales Vorgehen nach dem - häufig mißverstandenen - Konvergenz-Divergenz-Prinzip die Chance, besser gesicherte Ergebnisse zu erhalten. Anhand empirischer Untersuchungsergebnisse läßt sich diese Auffassung begründen.

### **Kriminalität – Ursachenzuschreibung und Strafhärte: Eine Untersuchung aus der Sicht juristischer Laien**

*Anne Reichert & Wolfgang Bilsky  
Psychologisches Institut IV, Universität Münster*

Ursachenzuschreibung und Strafzumessungspräferenzen juristischer Laien wurden im Rahmen einer Interviewstudie deliktenspezifisch untersucht. Ausgehend von einer Sekundäranalyse der von Görden



(1996) mit einem umfangreichen Fragebogen erhobenen kriminalitätsbezogenen Ursachenzuschreibungen wurde zunächst ein Instrument entwickelt, das es gestattet, für Diebstahls- und Körperverletzungsdelikte zwischen person-, interaktions- und systembezogenen Attributionen zu differenzieren (Bilsky, Oswald & Marciszewski, 1989). In der anschließenden Hauptuntersuchung mit insgesamt 211 Befragten unterschiedlichen Alters und unterschiedlicher Berufszugehörigkeit wurden im ersten Teil auf der Grundlage vorgegebener Sanktionsalternativen individuelle Strafzumessungspräferenzen erhoben, die sich auf kurze Beschreibungen (Vignetten) eines Diebstahls- und eines Körperverletzungsdeliktes bezogen; zwei Tätermerkmale, Vorstrafe und sozioökonomischer Status, wurden hierbei im Rahmen eines 'between-subjects-design' systematisch variiert. Im zweiten Teil erfolgte die Erhebung kriminalitätsbezogener Ursachenzuschreibungen mit Hilfe des zuvor entwickelten Instruments; jeweils die Hälfte der Befragten äußerte sich hierbei zur Entstehung von Diebstahls- und Körperverletzungsdelikten. Im abschließenden dritten Teil bestand die Aufgabe der Befragten darin, die im ersten Teil vorgegebenen Sanktionsalternativen nach ihrer subjektiv wahrgenommenen Strafhärte zu ordnen. Hinsichtlich der Deliktspezifität von Ursachenzuschreibungen ergaben sich für Körperverletzungsdelikte höhere personbezogene Attributionen als für Diebstahlsdelikte. Darüber hinaus konnte gezeigt werden, daß ein deliktspezifischer Zusammenhang zwischen den Ursachenzuschreibungen einer Person und der von ihr je präferierten Strafhärte existiert. Schließlich bestätigte die kovarianzanalytische Datenauswertung sowohl für Diebstahl als auch für Körperverletzung den Einfluß der untersuchten Tätervariablen auf die individuell präferierte Strafhärte.

### **Stilistische Kriterien in der Glaubhaftigkeitsdiagnostik: Ein Linguistisches Modell als Analyseinstrument zur Beurteilung kindlicher Zeugenaussagen**

*Judith Rickers & Susanna Niehaus*

*Institut für Psychologie, Universität Osnabrück, Institut für Psychologie, Universität Potsdam*

Die Methode der Wahl in der Glaubhaftigkeitsdiagnostik kindlicher Zeugenaussagen ist zur Zeit die Kriterienorientierte Inhaltsanalyse nach Steller und Köhnken (1989). Insbesondere die Unterscheidung wahrer Aussagen von falschen Aussagen mit Erlebnishintergrund gestaltet sich jedoch weiterhin als schwierig (Niehaus, 1999). Aus diesem Grund wurde eine Erweiterung des inhaltlichen Kriterienkatalogs um stilistische Kriterien erwogen. Ziel war es, ein Linguistisches Kategorienmodell aus der Attributionsforschung (Semin & Fiedler, 1988) und den daraus abzuleitenden Abstraktheitsgrad verbaler Aussagen auf seine Tauglichkeit für die Differenzierung kindlicher Zeugenaussagen mit unterschiedlichem Wahrheitsgehalt zu überprüfen.

In den Arbeiten von Semin und Fiedler (z.B. 1991) hat sich eine recht einfache Unterscheidung verschiedener Wortklassen, die als Prädikate in Beschreibungen von Verhalten auftreten können, als vielversprechend für die quantitative Analyse sprachlichen Materials erwiesen. Die Wortklassen, deren Einteilungsmerkmale auf linguistischen Kriterien beruhen und somit unabhängig vom semantischen Inhalt der Prädikate sind, bilden eine Hierarchie aufsteigender Abstraktheit.

Im Rahmen einer Pilotstudie wurden einer wahrheitsgemäß aussagenden Gruppe kindlicher Verkehrsunfallopfer (*Erlebnisgruppe*, N=10) zwei Gruppen Falschaussagender (jeweils N=10) gegenübergestellt mit dem Ziel, unterschiedliche Wahrnehmungsgrundlagen zu operationalisieren. Die *Phantasiegruppe* verfügte über keinerlei eigene Unfallerfahrungen, während die Kinder der Gruppe *Nacherzählung mit Unfallerfahrung* über eigene Unfallerfahrungen und die eines Unfallkindes verfügten. Die drei Aussagebedingungen wurden bezüglich verbaler Fähigkeiten, Alter, Geschlecht und

Nationalität parallelisiert. Die verschrifteten Interviews wurden von zwei Raterinnen ausgewertet, die bezüglich der Gruppenzugehörigkeit blind waren.

Die Ergebnisse der Varianzanalyse und anschließender ROC-Analysen (Receiver-Operating-Characteristic) zu ausgewählten Kontrasten sollen bezüglich ihrer potentiellen Bedeutung für die Glaubhaftigkeitsdiagnostik diskutiert werden.

### **Diagnostik psychischer Störungen bei Sexualstraftätern im Maßregelvollzug**

*Anja Schmidt<sup>1</sup>, Jürgen Hoyer<sup>1</sup> & Heike Kunst<sup>2</sup>*

*<sup>1</sup>Institut für Psychologie, Technische Universität Dresden, <sup>2</sup>Justizvollzugsanstalt Kassel II*

Strukturierte Untersuchungen zur Diagnostik psychischer Störungen bei Sexualstraftätern im Maßregelvollzug fehlen bisher im deutschen Sprachraum fast völlig. In der vorliegenden Studie wurden auf der Basis des DSM-IV 30 Sexualstraftäter mit einer Paraphilie (DSM-IV 302.xx) und 25 Sexualstraftäter mit einer Impulskontrollstörung n.n.b. (DSM-IV 312.30) in zwei Maßregelvollzugseinrichtungen im Hinblick auf die Aktual- und Lebenszeitprävalenz ausgewählter Achse-I-Störungen und Persönlichkeitsstörungen untersucht. Die Diagnostik der psychischen Störungen erfolgte unter Verwendung des Diagnostischen Kurz-Interview bei psychischen Störungen (Mini-DIPS) und des Strukturierten Klinischen Interview für DSM-IV Achse-II (SKID-II). Sechsendneunzig Prozent (N = 53) der Sexualstraftäter wiesen mindestens eine Achse-I-Störung oder eine Persönlichkeitsstörung als Lebenszeitdiagnose neben der Paraphilie oder der Impulskontrollstörung n.n.b. auf. Angststörungen bildeten mit 69% in der untersuchten Stichprobe die häufigste Störungsgruppe; ähnlich häufig waren Affektive Störungen (58%) und Substanzbezogene Störungen (56%). Paraphile wiesen besonders häufig die Lebenszeitdiagnose der Sozialphobie auf und Impulskontrollgestörte die Lebenszeitdiagnose des Schweren Depressiven Syndrom. Siebzig Prozent (N = 21) der Paraphilen und 48% (N = 12) der Impulskontrollgestörten hatten mindestens eine Cluster B-Persönlichkeitsstörung bzw. Cluster C-Persönlichkeitsstörung. Die Ergebnisse unterstreichen, dass eine umfassende Diagnostik psychischer Störungen bei Sexualstraftätern im Maßregelvollzug eine wesentliche Voraussetzung für eine effektive Therapie und Rückfallprävention dieser kleinen, aber besonders gefährlichen Tätergruppe ist. Auf dieser Basis sollten sich nach weiterer strukturierter Diagnostik störungsspezifisch relevante Behandlungselemente zur Therapie paraphiler und impulskontrollgestörter Sexualstraftäter ableiten lassen.

### **Entwicklungspsychologisch bedeutsame Faktoren in der Kriminalprognose**

*Vera Schneider & Klaus-Peter Dahle*

*Institut für Forensische Psychiatrie, Freie Universität Berlin*

Die zunehmende Bedeutung entwicklungspsychologischer Konzepte und Perspektiven hat die Erforschung normabweichenden und delinquenten Verhaltens erheblich befruchtet. Gleichwohl ist festzustellen, daß der Schwerpunkt der bisherigen Forschungsbemühungen sich – trotz des Anspruchs einer „life-span-developmental“ Perspektive – bislang weitgehend auf frühe Entwicklungsstadien beschränkte. Dies gilt auch für den wichtigen Anwendungsbereich der Entwicklung von Prognosemethoden zur Vorhersage von Rückfalldelinquenz. Hier zeichnet sich in jüngerer Zeit ein Trend in Richtung kriteriengebundener bzw. statistischer Prognosemethoden ab, die vielfach jedoch relativ starre

Prädiktoren beschreiben und unabhängig von Alter und Lebensphase das Rückfallrisiko erfassen sollen.

Besitzen bestimmte Rückfallprädiktoren zu unterschiedlichen Lebensphasen im Erwachsenenalter eine unterschiedlich starke prognostische Relevanz? Welche Merkmale sind dies und gibt es andererseits Risikofaktoren, die sich unabhängig von der Lebensphase der Betroffenen als prognostisch valide erweisen?

Im Rahmen der Berliner CRIME-Studie, in der die Biographien von 397 ehemaligen Strafgefangenen seit 1976 weiterverfolgt werden, wird derzeit die Beziehung zwischen diesen entwicklungspsychologischen Faktoren einerseits und Prädiktoren für den Verlauf krimineller Karrieren andererseits untersucht. Ausgehend von Theorien der Entwicklungspsychologie des Erwachsenenalters sowie gängiger Prognosekriterien, die in klinischen und statistischen Kriminalprognosen verwendet werden, ist hierbei die Fragestellung forschungsleitend, ob und inwieweit die Vorhersage der zukünftigen Legalbewährung von Straftätern zum Zeitpunkt der Haftentlassung vom Entwicklungsstadium abhängt.

Vorgelegt werden erste Ergebnisse aus einem Teilprojekt der Berliner CRIME-Studie, das sich insbesondere mit der Erforschung dieser Fragestellung beschäftigt. Langfristiges Ziel ist hierbei die Verfeinerung kriminalprognostischer Methoden, indem zusätzlich entwicklungspsychologische Aspekte Eingang in die Beurteilung zukünftiger Rückfallrisiken finden.

### **“Einladen zur Spekulation” – Anforderungscharakteristika versus Quellenkonfusion in Falschaussagen von Kindern**

*Nadja Schreiber & Janat Fraser Parker*

*Florida International University, Miami*

Im Rahmen verschiedener Massenbeschuldigungsverfahren der vergangenen zwanzig Jahre wurde wiederholt auf suggestive Befragungsmethoden hingewiesen. In der empirischen Literatur wurde gezeigt, dass verschiedene dieser suggestiven Befragungstechniken Zeugenaussagen von Kindern verfälschen können. In einer ersten empirischen Überprüfung der Interviewtechnik “Einladen zur Spekulation” konnte gezeigt werden, dass in Abhängigkeit vom Messzeitpunkt Kinder im Vorschulalter mehr hoch erwartbare falsche Antworten produzieren, wenn sie eingeladen werden, zu einem Originalereignis zu spekulieren (Schreiber, Wentura & Bilsky, im Druck). Die vorliegende Untersuchung überprüft, ob den Falschaussagen eine echte Erinnerungsverfälschung im Sinne einer Quellenkonfusion zugrunde liegt, oder ob die Falschaussagen durch Anforderungscharakteristika der Interviewsituation zu erklären sind. Darüberhinaus wird untersucht, ob “Einladen zur Spekulation” eine ebenso grosse Wirkung auf die Erinnerung interaktiver gegenüber nicht interaktiven Handlungen und in einer anderen Altersgruppe hat. Hundertvierzig Kindern im Alter von 5/6 und 8/9 Jahren wurden Clownshows gezeigt, in denen ein Clown hoch und wenig wahrscheinliche interaktive und nicht interaktive Handlungen mit alltäglichen Gegenständen präsentierte. Zwei Wochen nach der Clownshow wurden die Kinder zu einer Hälfte der interaktiven und nicht interaktiven Items eingeladen zu spekulieren. Drei Wochen später wurden die Kinder gebeten, die Clownshow zu erinnern. Zu jedem der erinnerten Items wurde mit dem Kind ein Quellenüberwachungstest durchgeführt (“logic-of-opposition-technique”; Lindsay, 1990). Wir erwarten mehr Falschaussagen zu den Items, zu denen eingeladen wurde zu spekulieren, zu den nicht interaktiven Items und in der jüngeren Altersgruppe. Ebenso erwarten wir weniger Falschaussagen nach der Aufforderung zur Quellenüberwachung, was für einen Anteil an Anforderungscharakteristika an den Falschaussagen spricht.

## **Schuldig bei Verdacht: Wie confirmatorisches Hypothesentesten zur fälschlichen Beschuldigung wegen sexuellen Kindesmißbrauchs führt**

*Stefan Schulz-Hardt<sup>1</sup>, Eberhard Höfer<sup>2</sup> & Günter Köhnken<sup>2</sup>*

*<sup>1</sup>Institut für Psychologie, Universität München, <sup>2</sup>Institut für Psychologie, Universität Kiel*

Im Verlauf der letzten 15 Jahre hat eine Reihe spektakulärer Gerichtsverfahren wegen vermeintlichen sexuellen Kindesmißbrauchs stattgefunden (z.B. der Montessori-Prozeß), in denen Kinder durch suggestive Befragungen zu geradezu obskuren Falschbeschuldigungen gebracht wurden. Diese suggestiven Befragungen fanden statt, weil die beteiligten Personen (Eltern, Mitarbeiter von Kinderschutzorganisationen, Ermittlungsbeamte) schon frühzeitig von der Schuld der verdächtigten Person(en) überzeugt waren, ohne daß entsprechende Belege für diese Annahme vorlagen. Für das Phänomen solcher realitätsfremder „Vorverurteilungen“ haben wir auf der Fachgruppentagung Rechtspsychologie in Erlangen 1999 einen Erklärungsansatz vorgestellt, der auf sozialpsychologischen Überlegungen zu individueller und kollektiver Informationsverarbeitung aufbaut. Dieser Erklärungsansatz soll im vorliegenden Vortrag weiterentwickelt und durch jüngst durchgeführte Studien illustriert werden. Unserem Ansatz zufolge kann die Überprüfung eines Mißbrauchsverdachts als Testung einer sozialen Hypothese verstanden werden. Wie aus der sozialpsychologischen Forschung bekannt ist, verläuft soziales Hypothesentesten unter bestimmten Bedingungen *confirmatorisch*, d.h. es kommt zu einer erhöhten Wahrscheinlichkeit, die Hypothese fälschlicherweise gegen die vorliegende Evidenz zu bestätigen. Wie wir herausarbeiten, liegen im Fall des sexuellen Mißbrauchsverdachts genau diese Bedingungen – wie z.B. emotionalisierende Wirkung der Hypothese, uneindeutige Evidenz, Fehlen von Kennzeichen für das Vorliegen der Alternativhypothese (Nicht-Mißbrauch) – vor. In mehreren Experimenten zeigen wir, wie der bloße Verdacht des sexuellen Kindesmißbrauchs dazu führen kann, eine harmlose kindliche Schilderung einer Zaubervorstellung als Mißbrauchsbeleg zu interpretieren und den darin geschilderten Zauberer für einen Sexualstraftäter zu halten.

## **Tötung in Verbindung mit sexueller Gewalttat**

*Peter Steck*

*Fachbereich Psychologie, Universität Konstanz*

Die Studie bezieht sich auf diskriminative Merkmale von Tötungen oder Tötungsversuchen, die in unmittelbarem zeitlichem Zusammenhang mit Vergewaltigung oder sexueller Nötigung eines weiblichen Opfers ausgeführt werden. Es wird angenommen, daß die zeitliche Kontingenz der beiden Delikte einem funktionalem Zusammenhang entspricht. Die Tötungshandlung wird dabei als Reaktion auf einen vorausgegangenen Konflikt verstanden. Unter dieser Annahme werden besondere Lebensereignisse der Tatanlaufzeit (Zeitraum bis vier Wochen vor der Tat) auf Seiten des Täters sowie Merkmale der Tatvorszene und des Tatablaufes erfaßt. Eigentümlichkeiten tödlich verlaufender sexueller Gewalttaten sollen auf diskriminanzanalytischem Wege sichtbar gemacht werden. Den nach dem logistischen Modell ausgeführten Analysen liegen die dichotomisierten Daten von 19 Fällen eines Tötungsdelikts in Verbindung mit einer Straftat gemäß §§ 177 und 178 StGB (alte Fassung) sowie von 48 Fällen sexueller Gewalttaten der gleichen Art ohne Tötung zugrunde. Die tatbezogenen diskriminanzanalytischen Vergleiche zwischen beiden Tätergruppen werden ergänzt durch Vergleiche auf der biographischen Ebene.

Die Datenanalyse weist folgende Merkmale als typisch für den tödlichen Verlauf der sexuellen Gewalttat aus: Die Tat ist relativ kurzfristig geplant (nicht mehr als zwei Tage), die Tat wird während der Tagesstunden ausgeführt, der Täter erinnert für die Tatzeit Eindrücke der Bedrohung und der Frustration, zwischen Täter und Opfer kommt es vor der Tötung zu einer tätlichen Auseinandersetzung, es findet keine Penetration statt.

### **Der Ausländereffekt bei türkischen und deutschen Kindern: Ergebnisse einer neuen Matching-Aufgabe**

*Barbara Trinkl<sup>1</sup>, Elena Guberova<sup>1</sup> & Siegfried L. Sporer<sup>2</sup>*

*<sup>1</sup> Institut für Psychologie, Universität Wien, <sup>2</sup> Institut für Psychologie, Universität Giessen*

Der sog. Ausländereffekt bezeichnet das differentiell schlechtere Wiedererkennen von Gesichtern einer anderen ethnischen Gruppe im Vergleich zu Gesichtern der eigenen ethnischen Gruppe. In der angloamerikanischen Literatur wurde der Ausländereffekt wiederholt nachgewiesen, vor allem mit Gesichtern von Schwarzen, Orientalen und Weissen (im physiognomischen Sinn). Ergebnisse zu anderen ethnischen Gruppen, insbesondere zu Mittel- und Südeuropäern bzw. mit Nicht-Erwachsenen fehlen. Um den Ausländereffekt mit Kindern zu testen, wurde eine neue Gesichter-Zuordnungsaufgabe konstruiert, in der Kinder Zielgesichter anderen Gesichtern zuordnen sollten. Abhängige Variable war die dafür benötigte Zeit. Die Untersuchung wurde an fünf Wiener Hauptschulen mit 256 Kinder der 5. bis 8. Klasse (128 türkische, 128 österreichische Kinder) durchgeführt. Ausgehend von dem In-group/Out-group-Modell von Sporer (2001) wurde postuliert, dass der Ausländereffekt bereits auf der Wahrnehmungsebene, und nur bei deutschen, nicht aber bei türkischen Kindern, anzutreffen sein sollte. Die Ergebnisse entsprachen voll den theoretischen Erwartungen. Türkische Kinder waren schneller als österreichische Kinder bei Gesichtern ihrer eigenen ethnischen Gruppe, während sich bei den Gesichtern von Deutschen keine Unterschiede ergaben. Implikationen für Identifizierungsaussagen von Kindern werden diskutiert.

### **Konstanz in erlebnisbasierten und erfundenen Aussagen**

*Renate Volbert, Julia Braun, Yvonne Gretenkord, Maren Teske & Sandra Wilma-Mews*

*Institut für Forensische Psychiatrie, Freie Universität Berlin*

Sowohl bei richterlichen wie auch bei psychologischen Beurteilungen der Glaubhaftigkeit einer Aussage kommt der Aussagekonstanz bei wiederholter Befragung eine wichtige Rolle zu. Es wird dabei davon ausgegangen, daß aufgrund gedächtnispsychologischer Gegebenheiten Konstanz in bestimmten Aspekten auch über längere Erinnerungsintervalle hinweg zu erwarten ist (beispielsweise Schilderung des zentralen Kerngeschehens, Benennung der beteiligten Personen), während in anderen Aspekten Inkonzistenzen auftreten (beispielsweise Angaben zur Reihenfolge bei mehreren Handlungssequenzen oder zum Schmerzerleben). Bisher liegen jedoch kaum empirische Überprüfungen zu diesen Annahmen vor. Weder sind die Zuordnungen von Items zu der Kategorie der "erwarteten Konstanz" bzw. der "erwarteten Inkonzistenzen", wie sie beispielsweise von Arntzen (1993) formuliert werden,

<sup>1</sup> Arntzen, F. (1993). *Psychologie der Zeugenaussage. System der Glaubwürdigkeitsmerkmale*. (3. Aufl.), München: Beck.

empirisch überprüft, noch ist systematisch untersucht worden, ob Unterschiede bezüglich der Konstanz zwischen erlebnisbasierten und erfundenen Aussagen existieren. In einer Untersuchungsreihe wurde diesen Fragen nachgegangen.

Untersucht wurden sowohl Aussagen von Erwachsenen wie von kindlichen Versuchspersonen. Dreißig Kinder zwischen sieben und neun Jahren berichteten ein erlebnisbasiertes und ein erfundenes Ereignis und wurden nach acht Wochen sowie nach einem Jahr erneut hierzu befragt. Ebenso schilderten 15 Erwachsene ein wahres und ein erfundenes Ereignis und wurden nach einem Intervall von acht Wochen erneut interviewt. In jedem Interview wurde jede Sinneinheit einer Unterkategorie der Gruppe "erwartete Konstanz" bzw. "erwartete Inkonzanzen" zugeordnet und anschließend im Vergleich zum vorherigen Befragungszeitpunkt als Übereinstimmung, einfache Ergänzung, qualifizierte Ergänzung oder Widerspruch kodiert. Informationen, die nur im vorherigen und nicht im späteren Interview genannt wurden, wurden entsprechend als Auslassungen kodiert. Unterschieden wurde zusätzlich, ob die jeweilige Information im freien Bericht bzw. auf offene Fragen produziert wurde oder ob sie erst auf spezifische Fragen erfolgte.

Bei allen Vergleichen zeigten sich erwartungsgemäß mehr Übereinstimmungen in der Kategorie der "erwarteten Konstanz" als in der Kategorie der "erwarteten Inkonzanzen". Wahre und falsche Aussagen wiesen unterschiedliche Konstanzmuster auf: Bei allen Vergleichen traten in unwahren Aussagen signifikant mehr Widersprüche auf. Qualifizierte Ergänzungen fanden sich dagegen (zumindest nach den 8-Wochen-Intervallen) signifikant häufiger in wahren Aussagen. Nicht bestätigt werden konnte hingegen in diesen Arbeiten die Auffassung, daß ein Muster von differenzierter Inkonzanz (Übereinstimmungen in der Kategorie der "erwarteten Konstanz"; inkonzante Angaben in der Kategorie der "erwarteten Inkonzanzen") ein Glaubhaftigkeitsmerkmal darstellt. Entsprechende Muster traten nämlich in ähnlichem Ausmaß in wahren und in falschen Aussagen auf. In beiden Kategorien konstante Aussagen fanden sich dagegen besonders häufig in wahren, in beiden Kategorien inkonzante Angaben dagegen besonders häufig in erfundenen Darstellungen.

## **Verbesserung der Inhaltsqualität von Falschaussagen durch Training**

*Renate Volbert & Yvonne Rutta*

*Institut für Forensische Psychiatrie, Freie Universität Berlin*

Es wurde die Frage geprüft, ob die inhaltliche Qualität von Falschaussagen durch Vermittlung von Wissen über die aussagepsychologische Methodik und insbesondere über die inhaltlichen Qualitätsmerkmale so verbessert werden kann, daß keine Differenzen zwischen erlebnisbasierten und erfundenen Aussagen mehr zu finden sind. Ziel der Untersuchung war nicht in erster Linie, Maßstäbe zu erhalten für wohl eher wenig praxisrelevante Fälle, in denen eine Aussage durch ein aussagepsychologisches Training gezielt vorbereitet worden sein könnte. Intention war vielmehr, die theoretische Implikation zu prüfen, dass es die kognitive Leistungsfähigkeit eines Zeugen überschreiten würde, in einer erfundenen Aussage ähnlich viele Qualitätsmerkmale zu produzieren wie in einer selbst erlebten.

Hierzu wurden je eine wahre und eine erfundene Aussage zu einem autobiographisch relevanten Thema mit negativer Tönung von 16 Experten und 16 Laien erhoben. Bei den Experten handelte es sich um Psychologiestudenten, die gerade ein einsemestriges Seminar zum Thema "Glaubwürdigkeitsbegutachtung" absolviert hatten, in dem die gesamte Methode sowie die einzelnen Qualitätsmerkmale ausführlich dargestellt und an Praxisfällen demonstriert und eingeübt worden waren. Die Kontrollgruppe bestand ebenfalls aus Psychologiestudenten im Hauptstudium, die beabsichtigten, Veranstaltungen in Rechtspsychologie zu besuchen, bislang aber noch keine Kenntnis der aussage-

psychologischen Methodik erworben hatten. Anhand der Transkripte der Aussagen wurden inhaltliche Qualitätsmerkmale von zwei gutachterlich tätigen Experten kodiert.

Während bei der Kontrollgruppe die Qualitätsmerkmale zwar keineswegs durchweg signifikant zwischen wahren und falschen Aussagen differenzierten, lagen hier doch alle Mittelwertunterschiede in der erwarteten Richtung. Dagegen fand sich in der Expertengruppe ein umgekehrter Trend: Mittelwerte waren in den erfundenen mindestens mit wahren Aussagen vergleichbar, wenn sich signifikante Unterschiede fanden, waren diese erwartungskonträr, lagen also in erfundenen Aussagen signifikant höher als in wahren. Einzige Ausnahme bildete das Merkmal "Eingeständnis von Erinnerungslücken". Die Ergebnisse werden diskutiert im Hinblick auf ihre theoretischen Implikationen sowie ihre praktische Bedeutung. In theoretischer Hinsicht ist insbesondere zu erörtern, ob die mehrfach belegten qualitativen Unterschiede zwischen wahren und falschen Aussagen weniger auf kognitiven Überforderungen bei der Produktion von komplexen Falschaussagen basieren, sondern stärker auf von dem aussagepsychologischen Ansatz abweichende implizite Theorien zur Glaubwürdigkeitsbeurteilung zurückzuführen sind, bzw. ob die Annahme der kognitiven Überforderung nur bis zu einem bestimmten Niveau intellektueller Kapazitäten auf Seiten der Aussagenden gilt. Im Hinblick auf die praktische Bedeutung ist zu betonen, daß in realen Fällen die Qualitätsanalyse in einen komplexen Beurteilungsprozeß eingebunden ist und auch andere Analyseeinheiten (z.B. Konstanzanalyse) zur Verfügung stehen. Dennoch sind mögliche Grenzen der merkmalsorientierten Inhaltsanalyse zu diskutieren.

### **Wiedererkennen des Täters oder Identifizieren des Verdächtigen? Experimentelle Evidenz für einen Bias in der Zeugengegenüberstellung**

*Arne Weigold & Dirk Wentura*

*Psychologisches Institut IV, Universität Münster*

In einem Experiment (N=70) wurde die Frage untersucht, in welchem Maße ein Tatverdächtiger in einer Wahlgegenüberstellung nur aufgrund seines Status als Verdächtiger identifiziert werden kann. Für den Verdächtigen ist die Situation – im Gegensatz zu den Distraktorpersonen – durch eine hohe persönliche Relevanz des Ausgangs der Wahlgegenüberstellung gekennzeichnet: Wird er durch einen Zeugen identifiziert, ist er schwer belastet; wird er nicht identifiziert, ist er in der Regel deutlich entlastet. Diese Relevanz kann zum einen zu Zeichen von Anspannung und erhöhtem arousal führen, die möglicherweise wahrnehmbar sind. Zum anderen führt sie zu dem Versuch, sich möglichst unauffällig zu verhalten, was ebenfalls zu wahrnehmbaren Unterschieden zu den Distraktorpersonen führen könnte. Um im Labor diese psychologischen Merkmale des Verdächtigen nachzustellen, wurde in insgesamt sechs Wahlgegenüberstellungen jeweils ein „Verdächtiger“ per Los bestimmt, der die Chance hatte, einen hohen Geldbetrag zu gewinnen, wenn er es schafft, nicht durch Beobachter identifiziert zu werden. Videoaufnahmen dieser Gegenüberstellungen wurden Versuchspersonen mit der Aufgabe präsentiert, jeweils die Person mit der Aussicht auf den Geldgewinn zu identifizieren. Es zeigte sich erwartungskonform, daß diese „Verdächtigen“ mit einer um 10% gegenüber einer Kontrollbedingung (Videoaufnahmen derselben Wahlgegenüberstellungen vor der Auslösung des „Verdächtigen“) erhöhten Wahrscheinlichkeit identifiziert werden.





# Was wissen Jugendliche über Sexualdelikte? - Eine Untersuchung zu Skripten von sexuellem Mißbrauch an Kindern

Heike Argstatter & Sandra Loohs

Lehrstuhl Psychologie II der Universität Regensburg

## 1 Einleitung

### 1.1 Hintergründe

Sexueller Missbrauch ist ein sehr komplexes Themengebiet. In Anlehnung an Amann & Wipplinger (1997) können jedoch einige wesentliche Zusammenhänge zwischen Wissen einerseits und daraus resultierendem Handeln bzw. Verhalten andererseits angenommen werden. Zentrale Aspekte bezüglich des Wissens um sexuellen Mißbrauch sind Wissen zur Definition, zur Epidemiologie

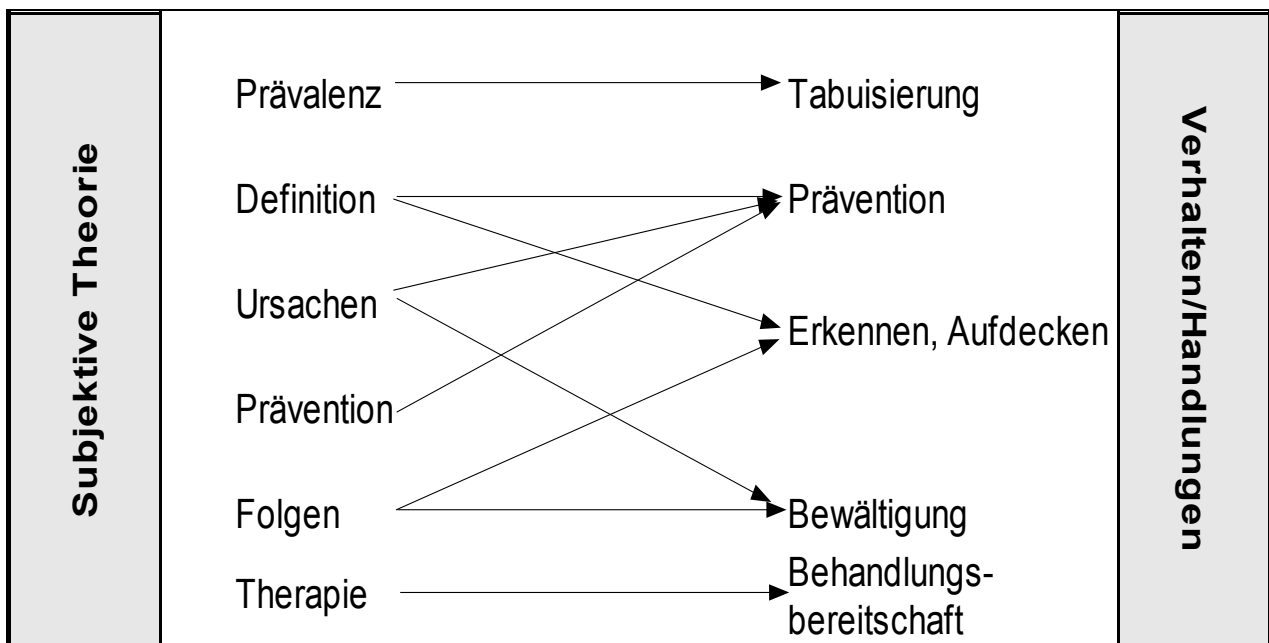


Abbildung 1: Zusammenhänge zwischen Wissen und Handlungskonsequenzen (Amann & Wipplinger, 1997, S. 775)

(insbesondere der Prävalenz), zu den Ursachen, zur Prävention, zu Folgen und zu möglichen therapeutischen Interventionen.

Das Wissen um die Häufigkeit und die Ursachen von sexuellem Mißbrauch sowie die vorkommenden Taten hängt direkt mit der Tabuisierung zusammen – und je bekannter das Ausmaß und die Umstände von sexuellem Mißbrauch sind, desto höher ist die Wahrscheinlichkeit des Intervenierens. Sind präventive Möglichkeiten bekannt, kann zumindest im persönlichen Umfeld neuen Übergriffen vorgebeugt werden. Mit größerem Wissen um mögliche Folgen von sexuellem Mißbrauch sowie um therapeutische Möglichkeiten kann Betroffenen effektiver geholfen und gegebenenfalls die Bewältigung von eigenen Mißbrauchserlebnissen unterstützt werden.

#### 1.1.1 Definition

Juristisch gehört sexueller Mißbrauch als Offizialdelikt zu den „Straftaten gegen die sexuelle Selbstbestimmung“ und wird in §174 (sexueller Mißbrauch von Schutzbefohlenen) und §176 StGB (sexueller Mißbrauch von Kindern) geregelt.

Der vorliegenden Arbeit liegt eine (psychologische) Definition von Bange und Deegener (1996) zugrunde:

„Sexueller Mißbrauch an Kindern ist jede sexuelle Handlung, die an oder vor einem Kind entweder gegen den Willen des Kindes vorgenommen wird oder der das Kind aufgrund körperlicher, psychischer kognitiver oder sprachlicher Unterlegenheit nicht wesentlich zustimmen kann. Der Täter nutzt seine Macht- und Autoritätsposition aus, um seine eigenen Bedürfnisse auf Kosten des Kindes zu befriedigen“ (S. 105).

Während in der juristischen Definition von sexuellem Mißbrauch die vorkommenden Handlungen bewußt allgemein formuliert sind und damit der Interpretation des Gerichts Spielraum läßt, ist es im Kontext einer wissenschaftlichen Untersuchung wichtig, die möglichen vorkommenden sexuellen Handlungen zu erfassen und zu kategorisieren. In Anlehnung an Bange und Deegener (1996, S. 134 ff.) können dazu vier Kategorien angenommen werden, die von weniger intensivem über intensiven bis zu sehr intensivem sexuellen Mißbrauch reichen und zusätzlich noch sexuellen Mißbrauch ohne Körperkontakt umfassen (siehe Tabelle 1).

Tabelle 1: Kategorien zur "Intensität" des sexuellen Mißbrauchs nach Bange und Deegener (1996, S. 135)

Sehr intensiver sexueller Mißbrauch	Intensiver sexueller Mißbrauch	Weniger intensiver sexueller Mißbrauch	Sexueller Mißbrauch ohne Körperkontakt
<ul style="list-style-type: none"> <li>• Versuchte oder vollendete Vergewaltigung (vaginal, oral, anal)</li> <li>• Opfer mußte Täter anal/oral befriedigen</li> </ul>	<ul style="list-style-type: none"> <li>• Aktive oder passive Masturbation</li> <li>• Täter zeigte/faßte an Genitalien</li> <li>• Opfer mußte Genitalien zeigen/anfassen</li> </ul>	<ul style="list-style-type: none"> <li>• Täter versuchte, an Genitalien zu fassen</li> <li>• Täter faßt an Brust des Opfers</li> <li>• Zungenküsse</li> </ul>	<ul style="list-style-type: none"> <li>• Exhibitionismus</li> <li>• Opfer muß Pornos anschauen</li> <li>• Täter beobachtet Opfer, wenn es nackt ist (beispielsweise beim Baden)</li> </ul>

### 1.1.2 Epidemiologie

Von besonderem Interesse für die Epidemiologie sind die Maße Prävalenz, Anzeigenrate, Dunkelziffer und Zeit von der Tat bis zur Anzeige.

Insgesamt kann man sagen, daß bei vorsichtigen Schätzungen jede 4. - 5. Frau und etwa jeder 8. Mann mindestens ein Erlebnis hatten, das als „sexueller Mißbrauch“ bezeichnet werden kann. Bei retrospektiven Befragungen hat sich ergeben, daß nur ein geringer Teil der Mißbrauchsfälle tatsächlich zu einer Anzeige geführt hat, „und daß in der Regel entweder besonders schwerer Mißbrauch (beispielsweise langdauernder Vater-Tochter-Inzest) angezeigt wird, oder Ereignisse, bei welchen unbekannte Personen (beispielsweise Exhibitionisten) beteiligt waren“ (Ernst, 1997, S. 57). Tendenziell gilt: der Mißbrauch wird entweder sofort mitgeteilt oder die Opfer schweigen mehr als ein Jahr, manchmal sogar Jahrzehnte (Bange & Deegener, 1996). Dunkelzifferschätzungen reichen von 60.000 Fällen pro Jahr (BKA, 1997) bis 300.000 Fällen (Kavemann & Lohstötter, 1986, S. 28).

### 1.1.3 Täter und Opfer

Statistische Angaben zu den Tätern (Polizeiliche Kriminalstatistiken) und empirische Untersuchungen (Bange & Deegener, 1996, S. 49f.) zeigen, daß die Täter bei sexuellem Mißbrauch von Kindern (nach §176 StGB) zu rund 98% Männer sind. Über 75% der Täter sind älter als 21 Jahre (d.h. der Altersunterschied zwischen Opfer und Täter beträgt mehr als 5 Jahre), woraus folgt, daß immerhin bis zu einem Viertel der Täter etwa Gleichaltrige sind! Aufgeschlüsselt nach Tätergruppen, stammt etwa ein Viertel der Täter aus der Familie des Opfers (Vater, Stiefvater, Großvater, Bruder, Onkel, Cousin...), etwa die Hälfte aus der näheren Umgebung (Bekannte, Freunde, Lehrer etc.) und nur etwa ein Viertel sind völlig Fremde (vgl. Kavemann & Lohstötter, 1986; Bange & Deegener, 1996).

Wie schon die Heterogenität der Tätergruppen andeutet, existiert kein eindeutiges Profil von Mißbrauchstätern, sie sind vielmehr erschreckend "normal" (vgl. Kavemann & Lohstötter, 1986, S.

97). „Das Einzige das alle Täter miteinander verbindet, ist ein geringes Selbstwertgefühl und eine Persönlichkeit voller Schamgefühle“ (Marquit, 1984, S. 124). Die Motivation für sexuellen Mißbrauch ist also nicht die sexuelle Befriedigung, sondern vielmehr die Erfüllung nicht-sexueller Bedürfnisse.

Die Opfergruppe umfaßt sowohl Mädchen als auch Jungen. Das Verhältnis Jungen - Mädchen liegt bei etwa 1:3. Die Betroffenen stammen aus allen Altersstufen, vom Kleinkind bis zum Erwachsenen, im Mittel sind die betroffenen Mädchen zu Beginn des Mißbrauchs 9,6 - 11,4 Jahre, die betroffenen Jungen 9,8 - 12 Jahre alt (Bange & Deegener, 1996).

#### 1.1.4 Ursachen

Die Ursachen für sexuellen Mißbrauch sind vielfältig. Genau wie es verschiedene Definitionen gibt, abhängig von der "Weltanschauung" der jeweiligen Forscher, findet man auch ideologisch geprägte Erklärungsmodelle. Nach Jönsson „läßt sich sexueller Mißbrauch auf mindestens drei verschiedenen Ebenen erklären: einer Individuum-zentrierten, einer familien-zentrierten und einer gesellschaftlich-zentrierten Ebene“ (Jönsson, 1997, S. 26ff.). Das „Drei - Perspektiven - Modell“ von Brockhaus & Kolshorn (1993) versucht „eine Integration von soziologischen Fragestellungen, die eher auf die Gesellschaft fokussieren“ mit „psychologischen Analysen, die das Augenmerk auf das Individuum richten“ (Brockhaus & Kolshorn, 1993, S. 220).

#### 1.1.5 Folgen

Die Diagnose von sexuellem Mißbrauch ist nicht einfach: Zum einen ist die Definitionsfrage problematisch (je nach Definitionsweite und -kriterien werden andere Auffälligkeiten in Zusammenhang mit sexuellem Mißbrauch gebracht), zum anderen gibt es kein eindeutiges Meßinstrument (vgl. Bange & Deegener, S. 62). Zudem gibt es, abgesehen von direkten körperlichen Folgen, kein eindeutiges Symptom oder Syndrom, das für sexuellen Mißbrauch typisch ist. Die Ursachen für auftretende Probleme und Auffälligkeiten können auch andere psychopathologische Befunde sein.

Konkrete Folgen können körperlich, psychosomatisch, psychisch, sozial oder sexuell sein. Die folgende Tabelle gibt eine Übersicht über die wichtigsten Folgen.

Tabelle 2: Folgen von sexuellem Mißbrauch (nach Bange & Deegener, 1996, S. 77ff. und Moggi, 1998, S. 188ff.)

Körperliche Folgen	Verletzungen im Anal- und Genitalbereich (beispielsweise Bluten, Hämatome, Scheiden- und Anlrisse, Fremdkörper in Scheide oder After) und Geschlechtskrankheiten, Schwangerschaft
Psychosomatische Folgen	Kopf-, Hals-, Magen- und Unterleibsschmerzen ohne erkennbare organische Ursachen, Eßstörungen, Schlafstörungen (Alpträume), Erstickungsanfälle, Sprachstörungen
Psychische Probleme	Niedriges Selbstwertgefühl, Depression, Angst, regressives Verhalten, Autoaggression (Selbstverletzungen, Suizidgedanken, -versuche)
Soziale Auffälligkeiten	Schulversagen oder extreme Leistungssteigerung, Verhaltensauffälligkeiten (Aggression, Hyperaktivität), Konzentrationsstörungen,
Sexuelle Auffälligkeiten	Exzessive Beschäftigung mit der Sexualität; Angst vor sexuellen Dingen, sexuelle Dysfunktionen

## 1.2 Strafprozessurale Aspekte

Wird ein Fall von sexuellem Mißbrauch zur Hauptverhandlung zugelassen, ist die oberste Aufgabe der Richter die sogenannte „Wahrheitserforschungspflicht“. Dabei muß das Gericht alle zur Wahrheitsfindung in Frage kommenden Beweismittel heranziehen.

Da es bei sexuellem Mißbrauch häufig keine „Sachbeweise“ (etwa eine Tatwaffe, gesicherte Spuren und dergleichen) sondern nur „persönliche Beweismittel“, insb. Zeugenaussagen gibt, steht die Frage nach der Glaubhaftigkeit der Zeugenaussage im Vordergrund.

Diese Frage wird in der Regel durch ein Gutachten eines psychologischen Sachverständigen geklärt.

### 1.2.1 Glaubhaftigkeit

Ausgehend von der sogenannten „Undeutsch Hypothese“ (Undeutsch, 1967), der zufolge sich wahre und erfundene Aussagen in inhaltlichen Eigenschaften in quantitativer und qualitativer Ausprägung unterscheiden, versucht die psychologische Glaubhaftigkeitsbegutachtung zwischen erlebnisbegründeten und intentional täuschenden oder durch Dritte induzierte Zeugenaussagen zu differenzieren.

Zur Überprüfung des Realitätsgehalts von Zeugenaussagen wurden verschiedene Systeme von Glaubhaftigkeitsmerkmalen erstellt. Aufbauend auf solchen Systemen von Undeutsch (1967), Arntzen (1970, 1993), Trankell (1971) und Littmann und Szewczyk (1983) entwickelten Steller und Köhnken die sogenannten Realkennzeichen (Steller & Köhnken, 1989).

Diese Realkennzeichen beziehen sich jeweils auf die Analyse einer einzelnen Aussage. Neben den allgemeinen Merkmalen wie dem allgemeinen Aufbau der Aussage, den speziellen Inhalten und inhaltlichen Besonderheiten sind vor allem deliktspezifische Aussageelemente wichtige Indikatoren für die Glaubhaftigkeitsbeurteilung. Deliktspezifität liegt „in qualifizierter Ausprägung [...] dann vor, wenn ein Zeuge ein komplexes deliktspezifisches Verhaltensmuster beschreibt, ohne über die erforderlichen deliktspezifischen Kenntnisse zu verfügen, oder wenn von sozial etablierten Stereotypen abweichende, aus psychologischer Sicht gleichwohl für den fraglichen Deliktbereich spezifische Angaben gemacht werden“ (Greuel, 1998, S. 99).

Arntzen (1993) hat durch die Analyse von rund 43.000 Fallbeispielen eine Reihe von delikttypischen Zeichen für sexuellen Mißbrauch gefunden. Je nach Alter des Opfers verändern sich die Annäherungs- und Geheimhaltungsstrategien. Tabelle 3 gibt einen Überblick über das altersabhängige Täterverhalten.

Tabelle 3: Altersabhängiges Täterverhalten bei sexuellem Mißbrauch nach Arntzen (1993)

	<b>Kleinkinder (0 - 6)</b>	<b>Schulkinder (7 - 9)</b>	<b>Vorpubertät (9 - 14)</b>
<b>Annäherung</b>	Ausnutzen einer günstigen Gelegenheit (Spiel-, Pflegesituation); Anlocken durch Genußmittel	Gespräch (Aufklärung, „Komplimente“, Witze), medizinische Erklärungen; materielle Anreize	Ausnutzung sexueller Neugierde, sexuell-erotische Verführung ==> Abhängigkeitsverhältnis; Gespräche
<b>Situationen</b>	Beiläufiger Charakter im Spiel oder der Pflege	Häusliches Umfeld, Schulweg, Schwimmbad, Kino	Reisen, Besuche, Übernachtung, Feiern, räumliche Enge im Schlafzimmer
<b>Häufigkeit</b>	Wenige Kontakte, selten langfristig	Langfristige Beziehungen	Charakter eines Liebesverhältnisses, sexuelle Bindung, sowohl langfristig, als auch nur bei Gelegenheit
<b>Tätergruppe</b>	Enger Umkreis (Hausmitbewohner, Nachbarn)	Verwandtschaft, Nachbarn	Alle möglichen
<b>Geheimhaltung</b>	Explizites Schweigegebot, durch Geschenke „erkauft“	Drohungen (beispielsweise körperliche Strafen), materielle Nachteile	explizites Schweigegebot, Drohungen mit sozialen Nachteilen (beispielsweise Heim, Gefängnis)
<b>Psychische Folgen</b>	Scham, Ekel	Gefühlsambivalenz	Gefühlsambivalenz

### 1.3 Gedächtnispsychologische Aspekte

Bei einer Zeugenaussage handelt es sich um einen Abrufprozeß aus dem Langzeitgedächtnis. Ereignisse und Begebenheiten, die selbst erlebt wurden, sind dabei im episodischen Gedächtnis gespeichert, während erfundene Berichte dem semantischen Gedächtnis, welches das „Weltwissen“ in Form von Schemata und Skripts umfaßt, entstammen. „Wenn jemand einen nicht selbst wahrgenommenen Sachverhalt schildert, steht ihm als Grundlage das abstrakte Schemawissen selbst zur Verfügung. Inkonsistente und irrelevante Details, die bei einer realen Wahrnehmung als „Anhang“ zum Schema gespeichert werden, finden sich hier nicht.“ (Steller und Volbert, 1999, S. 5). Das bedeutet, daß jede Zeugenaussage zunächst einmal auf einem allgemeinen Schema (vgl. Schank & Abelson, 1977) basiert.

Will man den Verlauf eines Ereignisses erfassen, muß der Schemabegriff allerdings etwas präzisiert werden, was durch die Einführung von sogenannten 'Skripts' geschieht. Ein 'Skript' bezeichnet die Repräsentation von Ereignissen anhand von typisierten Verlaufsschemata, die mit Hilfe der Skriptanalyse untersucht werden können. „Die Erfassung des Verlaufs von Ereignissen gibt Aufschluß über die raum-zeitliche Sequenzierung von Handlungen, über Aktivitäten der beteiligten Personen und über einbezogene Objekte. Die Organisation von Skripts zeigt, wie Ereignisabläufe erinnert werden. Das Wissen über Ablaufsequenzen impliziert sowohl Kontingenz- als auch Zielerwartungen“ (Dreher, 1999, S. 639).

Entgegen den Erwartungen von Schank und Abelson (1977) existieren Skripts auch für Ereignisabläufe ohne Erlebnisbezug, wie Fisher-Holst & Pezdek (1992) sowie Greenberg et al. (1998) am Beispiel eines bewaffneten Überfalls zeigen konnten.

Immer wieder wird in der Literatur darauf hingewiesen, daß Zeugen schemakonsistente bzw. schemadiskrepante Aussagen machen, ohne daß ein solches Schema je erfaßt worden wäre (vgl. Greuel, 1998, S.118; Ceci, 1995, S.43).

Auch in Bezug auf deliktspezifische Merkmale als Qualitätsmerkmal erlebnisfundierter Aussagen wird darauf hingewiesen, die Aussagen müßten sich von sozial etablierten Stereotypen unterscheiden lassen (vgl. oben).

Für die Bewertung der Deliktspezifität einer Aussage haben sich in der Praxis zwar einige Standards (beispielsweise Erwähnung eines Schweigegebots) eingebürgert, doch es existiert keine fundierte Erhebung, auf welche „sozial etablierten Stereotype“ von sexuellem Mißbrauch Bezug genommen werden kann.

## 2 Methode

### 2.1 Fragestellung

Auf Basis der vorgestellten Hintergründe zu sexuellem Mißbrauch, gedächtnis- und aussagepsychologischen Grundlagen untersucht die vorliegende Arbeit die Wissensbasis von Jugendlichen über sexuellen Mißbrauch anhand eines Fragebogens. Besonderer Wert wird dabei darauf gelegt, daß die Jugendlichen nicht über einen eigenen Erlebnishintergrund von sexuellem Mißbrauch verfügen und ihr Wissen daher aus dem semantischen Gedächtnis konstruieren müssen. Zunächst interessieren dabei die zentralen Aspekte des Wissens über sexuellen Mißbrauch, wie Wissen zu Epidemiologie, Täter-Opfer-Konstellation, Definition, Intervention und Folgen von sexuellem Mißbrauch. Außerdem soll herausgefunden werden, aus welchen Quellen die Jugendlichen ihre Informationen beziehen und schließlich ist noch von Interesse, inwiefern sie in der Lage sind, diese Informationen in ein Skript umzusetzen.

## 2.2 Versuchspersonen und Vergleichsmaterialien

An der Fragebogenuntersuchung haben 136 Schüler im Alter von 15 Jahren ( $\bar{X} = 15,36$ ) aus 3 Regensburger Schulen, einer Hauptschule, einer Realschule und einem Gymnasium teilgenommen. Das Alter von ca. 15 Jahren wurde deshalb gewählt, weil sich die meisten von sexuellem Mißbrauch Betroffenen in diesem Alter an die Öffentlichkeit wenden und in der Folge die meisten Gerichtsverhandlungen mit damit verbundenen Zeugenaussagen stattfinden. In einer Zusatzerhebung wurden 50 Studierende verschiedener Fachrichtungen im Alter von 19 Jahren befragt.

Die Fragebögen wurden zusammen mit unbeschrifteten Umschlägen ausgegeben. Dadurch wurde die Option eröffnet, den Fragebogen unbeantwortet zurückzugeben, ohne daß nachzuprüfen war, wer aktiv an der Untersuchung teilgenommen hat. Fragebögen wurden ausgeschlossen, wenn sie nicht oder offensichtlich nicht ernsthaft ausgefüllt waren, oder die VP angab, selbst von sexuellem Mißbrauch betroffen gewesen zu sein.

Die folgende Tabelle gibt einen Überblick über die Verteilung.

Tabelle 4: Teilnehmende an der Fragebogenerhebung

	Hauptschule			Realschule			Gymnasium			Schulen Gesamt			Studierende		
Ausgeteilte Fragebögen	36			50			50			136			50		
Rücklauf	21			46			43			110			28		
Leer	0			3			9			14			0		
Nicht ernsthaft ausgefüllt	0			5			1			6			0		
Selbst betroffen	0			4w	3m	7	1w	0m	1	5w	3m	8	1w	0m	1
Ausgewertete Fragebögen	12w	9m	21	16w	15m	31	14w	18m	32	42w	42m	84	22w	5m	27

Die bei den Jugendlichen erhobenen Daten wurden mit den Ergebnissen einer Textanalyse von 56 BravoGirl - „Erlebnisberichten“ der Serie „Mißbraucht, auch Du?“ (Bravo-Girl, 1996) sowie den in empirischen Studien (Bange & Deegener, 1996; Amann & Wipplinger, 1997; Polizeiliche Kriminalstatistik des BKA, Fassungen von 1998, 1999 und 2000) festgestellten Daten verglichen. Besonderes Gewicht liegt bei den Informationsquellen auf dem Vergleich mit den BravoGirl-Artikeln, da Bravo häufig als „Hauptinformationsquelle“ für Jugendliche bezeichnet wird.

## 3 Ergebnisse

Da im Fragebogen jeweils mehrere Fragen zu ähnlichen Aspekten gestellt wurden, wurden die Ergebnisse in fünf Komplexe zusammengefaßt: Definition (inclusive Epidemiologie), Skriptabfrage, Hilfeleistung, Folgen und Information.

### 3.1 Definition von sexuellem Mißbrauch

Die Jugendlichen gehen von einer relativ engen Definition von sexuellem Mißbrauch aus, da für sie Mißbrauch überwiegend erst bei „intensiven“ (d.h. Berührungen, Masturbation) oder „sehr intensiven“ (Penetration) Handlungen beginnt. In der BravoGirl-Analyse ist dieser Trend noch stärker ausgeprägt: 80% der Handlungen sind „intensiv“ oder „sehr intensiv“.

#### 3.1.1 „Epidemiologie“

Die Jugendlichen schätzen die Prävalenzrate auf durchschnittlich 24%. Der gemittelte Literaturwert (aus vier Erhebungen von Bange (Dortmund), 1992; Raupp/Eggers (Essen), 1993; Burger et al. (bundesweit), 1993; Richter (Hamburg), 1995, nach Bange & Deegener, 1996) liegt bei 27%, was eine gute Übereinstimmung darstellt.

Die Anzeigenrate wird von den Jugendlichen auf 16% geschätzt. Der Literaturwert (BKA, 1997) liegt bei 28%.

Die Schätzung der Dunkelziffer durch die Jugendlichen beträgt 8,77, d.h. auf jede bekannte Tat sollen rund 9 Taten kommen, die nicht öffentlich werden. Der empirische Vergleichswert wurde aus 3 verschiedenen Angaben errechnet. Ausgehend von rund 16.000 Anzeigen jährlich (PSK, 1997, 1998, 1999) ergeben sich Dunkelzifferzahlen von 3,75 (BKA, 1997), 5 (Unicef, 1999) oder 19 (Kavemann & Lohstötter, 1984), woraus sich ein arithmetisches Mittel von 9,25 ergibt, das sehr gut mit den Angaben der Jugendlichen korrespondiert.

Durchschnittlich schätzen die Jugendlichen die Dauer bis zur Ersteinlassung auf 12,12 Monate. Der Literaturwert (Dortmunder Studie, 1990, Bange & Deegener, 1996) liegt bei 12,2 Monaten, die durchschnittlichen Schätzungen und der Literaturwert unterscheiden sich also nicht.

Der wichtigste Grund für das Verschweigen von sexuellem Mißbrauch ist den Jugendlichen zufolge „Angst“ (75%), gefolgt von „Scham“ (45%) und „Drohungen des Täters“ (14%).

### 3.2 „Skriptabfrage“

Die Protagonisten, d.h. die an einem sexuellen Mißbrauch beteiligten Personen, sind ein männlicher Täter, der mindestens 5 Jahre älter als das weibliche, etwa zehnjährige Opfer ist. Täter und Opfer sind miteinander bekannt, da der Täter Vater, Freund oder Bekannter des Opfers ist.

Um das „übliche“ Skript eines sexuellen Mißbrauchs zu ermitteln, wurden die genannten Handlungen in 35 Kategorien kodiert (im Anhang findet sich eine Erläuterung für die Kategorien). In die Auswertung gingen dabei nur Ereignisse ein, die jeweils von mindestens 25% der Versuchspersonen genannt wurden. In Tabelle 5 sind die Ereignisse in der zeitlichen Ordnung aufgelistet, die ihrem mittleren Rangordnungsplatz entspricht. Als Qualitätsindikatoren der Skripte wurde die Gesamtzahl der genannten Items, die durchschnittliche Itemzahl und die Uniformität, d.h. das Verhältnis der Gesamtanzahl der Ereignisse je Skript zu Ereignissen, die nur von einer Person genannt wurden angegeben. Diese Angaben sind in der zweiten Zeile von Tabelle 5 aufgeführt.

Tabelle 5: Skriptabfrage. In Anlehnung an Greenberg et al. (1998) wurde als Maßstab für die Wichtigkeit der einzelnen Handlungen die relative Häufigkeit, mit denen sie genannt wurden, angelegt: **VERSALIEN** (75 - 100% Nennungen) = zentrale Ereignisse, *Kursiv* (45 – 75% Nennungen) = dezentrale Ereignisse, normal (25 – 45% Nennungen) = periphere Ereignisse

	<b>Innerfamiliäres Mißbrauchs-Skript EXPERTEN</b>	<b>Innerfamiliäres Mißbrauchs-Skript SCHULE</b>	<b>Innerfamiliäres Mißbrauchs-Skript UNI</b>	<b>Innerfamiliäres Mißbrauchs-Skript UNI + SCHULE</b>	<b>Innerfamiliäres Mißbrauchs-Skript BRAVO</b>	<b>Fremdtäter-Skript SCHULE</b>
	N=11; 159/7; 4,4%	N=11; 134/8; 5,97%	N=24; 221/7; 3,16%	N=35; 308/6; 1,95%	N=39; 316/7; 2,22%	N=29; 287/9; 3,14%
1	In Zimmer gehen	Erstkontakt	In Zimmer gehen	<i>Gelegenheit</i>	Erste Annäherung	<i>Erstkontakt</i>
2	Absperren	<i>Erste Annäherung</i>	<i>Gelegenheit</i>	In Zimmer gehen	<b>GELEGENHEIT</b>	Opfer wird verfolgt
3	Gespräch	<i>Gelegenheit</i>	<i>Tatort</i>	Gespräch	<b>TATORT</b>	<b>ÜBERFALL/ ENTFÜHRUNG</b>
4	Annäherung	Gespräch	<i>Gespräch</i>	<i>Tatort</i>	<b>KÜSSEN</b>	Gespräch
5	<b>TATORT</b>	<i>In Zimmer gehen</i>	Aufklärung	Entkleiden	Entkleiden	Fesseln
6	Porno	<i>Weitere Annäherung</i>	Entkleiden	Weitere Annäherung	<b>TÄTER BERÜHRT OPFER</b>	<i>Tatort</i>
7	<i>Gewalt</i>	Vertrauen	<b>TÄTER BERÜHRT OPFER</b>	<b>TÄTER BERÜHRT OPFER</b>	<i>Opfer muß Täter berühren</i>	Täter berührt Opfer
8	Weigerung	<i>Entkleiden</i>	Opfer muß Täter berühren	Wehren	<i>Wehren</i>	<i>Entkleiden</i>
9	<b>TÄTER BERÜHRT OPFER</b>	Tatort	<i>Negative Gefühle</i>	<i>Negative Gefühle</i>	Gewalt	<i>Wehren</i>
10	Negative Gefühle	<i>Wehren</i>	Täter beteuert Liebe	Fügung	Tat "normal"	<b>GEWALT</b>
11	<i>Wehren</i>	<b>GEWALT</b>	Fügung	<i>Penetration</i>	Orale Manipulation	<b>PENETRATION</b>
12	<i>Drohung</i>	<b>TÄTER BERÜHRT OPFER</b>	<i>Penetration</i>	Befriedigung/ Orgasmus	<i>Fügung</i>	<i>(Ermordung)*</i>
13	Orale Manipulationen	Fügung	Befriedigung/Orgasmus	<i>Drohung</i>	<i>Penetration</i>	Drohung
14	<b>ENTKLEIDEN</b>	<b>DROHUNG</b>	<i>Schweigegebot</i>	<i>Schweigegebot</i>	Schweigegebot	Hilferufe des Opfers
15	<i>Fügung</i>	Negative Gefühle	Drohung	Tat „normal“		Schweigegebot
16	Penile, vaginale Penetration	<b>PENETRATION</b>	Tat „normal“	Täter verläßt den Tatort		<i>Täter verläßt den Tatort</i>
17	Befriedigung/ Orgasmus	<i>Tat „normal“</i>				Opfer bleibt zurück
18	<i>Schweigegebot</i>	<i>Schweigegebot</i>				<i>Täter verläßt den Tatort</i>
19	Anziehen	<i>Täter verläßt den Tatort</i>				*Optional: entweder Ermordung oder weiter mit den Folgeschritten des Skripts



### 3.3 Wissen der Jugendlichen zur „Hilfeleistung“: Eingreifsituationen und Ansprechpartner

Die Mehrheit der Jugendlichen (70%) gibt an, daß sie prinzipiell eingriffen, wenn sie von sexuellem Mißbrauch erführen; wobei rund 14 % der Jugendlichen "immer" oder "sofort" bereit seien, aktiv zu werden und 17% erst, wenn sie Zeuge des Geschehens würden. Weniger als 10 % würden Hilfe holen, wenn die Betroffenen von dem Mißbrauchsgeschehen erzählen (8%) oder explizit um Hilfe bitten (7%) würden.

Die Jugendlichen erwähnen im wesentlichen 3 Interventionsmöglichkeiten: Die Mißbrauchsopfer direkt ansprechen (27%), Anzeige erstatten (25%) oder externe Hilfe holen (20%).

Die Frage nach den Ansprechpartnern wurde zwei Mal gestellt: Bei der Eigeneinschätzung der Ansprechpartner mußten die Jugendlichen sich vorstellen, in der Situation eines Mißbrauchsopfers zu sein, bei der Fremdeinschätzung sollten die Jugendlichen einschätzen, welche Personengruppen die „üblichen“ Ansprechpartner von Mißbrauchsopfern sind. In Tabelle 6 sind die Antworten nochmals aufgeschlüsselt dargestellt.

Tabelle 6: Ansprechpartner

	<b>Fremdeinschätzung</b> (N=167 Nennungen)	<b>Eigeneinschätzung</b> (N=97 Nennungen)	<b>Bravo</b> (N=44 Nennungen)	<b>Literatur*</b> (N=166 Nennungen)
<i>Verwandt</i> <sup>a</sup>	72%	48%	57%	35%
<i>Freund/in</i> <sup>a</sup>	70%	24%	21%	49%
<i>Professionelle</i> <sup>a</sup>	35%	23%	11%	4%
<i>Sonstige</i> <sup>a</sup>	20%	2%	0%	5%
<i>Niemand</i> <sup>a</sup>	0%	8%	17%	35%
a: Mehrfachnennungen möglich				*Nach: Bange und Deegener, 1996, Dortmunder Studie, 1990.

### 3.4 Wissen der Jugendlichen zu den „Folgen“

Über 90% aller Versuchspersonen können sich vorstellen, daß sich Kinder anders verhalten können, wenn sie Opfer von sexuellem Mißbrauch geworden sind, und alle untersuchten BravoGirl-Artikel enthalten Berichte über Folgen. Demgegenüber wurde empirisch mehrfach ein Anteil von 1/3 symptomfreier Mißbrauchs betroffener festgestellt (vgl. Kendall-Tacket et al., 1993). Dichotomisiert man die freien Antworten, geben die Jugendlichen zu 89%, die BravoGirl-Artikel sogar zu 97% „externalisierende“<sup>1</sup> Verhaltensweisen als Folgen an und nur zu 11% bzw. 3% „internalisierende“<sup>2</sup> Verhaltensweisen, während die Metastudie von Kendall-Tacket et al. (1993) ein ausgeglichenes Verhältnis aufweist.

### 3.5 Angaben der Jugendlichen zu „Informationen“

Insgesamt wurden 208 Antworten gegeben, was im Mittel 2,48 Informationsquellen pro Jugendlichen entspricht, der Median liegt bei 3. Dabei sind die (Massen-)Medien, vor allem das Fernsehen (84% der Nennungen) und Zeitschriften (22%), die wichtigsten Informationsquellen. Personale Informationsquellen, wie Lehrer/Lehrerinnen (22%), Eltern (16%) oder Freunde/Bekannte (13%) sind demgegenüber von sekundärer Bedeutung. Diese Ergebnisse bestätigen eine Untersuchung von Knappe und Selg (1993), in der *Eltern* als die wichtigsten Informationsquellen auch Medienberichte angaben, insbesondere Zeitungsberichte, Fernsehen und Radio.

Das Medienverhalten der Jugendlichen kann noch detaillierter ausgewertet werden. In Bezug auf den Zeitschriftenkonsum ergibt sich, daß 25% der Jugendlichen aktive Bravo-Leser sind, 34% „noch nie“ Bravo gelesen haben; die meisten Jugendlichen (42%) geben jedoch an, die Bravo „nicht mehr“ zu lesen, wobei der aktive Konsum im Mittel 2,6 Jahre (Median 2) zurückliegt. Von den

Jugendlichen, die die Bravo „nicht mehr“ lesen, geben 54% an, früher „häufig“ Bravo gelesen zu haben, 46% geben an, „selten“ Bravo gelesen zu haben.

Bezüglich des Fernsehkonsums wurde ermittelt, daß etwa 2/3 der Schüler täglich fernsehen, 26% 3-4 mal wöchentlich und nur 6% 1 mal pro Woche oder seltener und dann im Durchschnitt 2,03 Stunden vor dem Fernseher sitzen, was verglichen mit den Werten aus Umfragen zum Medienkonsum etwas untertrieben erscheint (Berliner Zeitung, 4.9.2000).

Die Benutzung des Computers hat sich flächendeckend durchgesetzt, vier von fünf Jugendlichen benutzen regelmäßig einen Computer. Wichtigster Zugang sind familiäre Computer (neben gemeinsam genutzten Rechnern besitzt bereits die Hälfte der Jugendlichen einen eigenen Computer) und Computerräume in der Schule. Das Internet wird auch rege genutzt, nur 14% der Jugendlichen geben an, gar nicht zu surfen. Allerdings sind Internetseiten mit sexuellem Inhalt (noch) nicht relevant, ein Fünftel der Befragten gibt an, einmal solche Seiten besucht zu haben, lediglich 6% taten dies öfter. Diese Ergebnisse widersprechen Befunden, daß „Sexseiten“ die mit Abstand am häufigsten angewählten Webpages sind.

### 3.6 Einflüsse von Moderatorvariablen

Als mögliche Moderatorvariable wurde der Informationsstand der Jugendlichen, operationalisiert als „Anzahl von Informationsquellen“ (Teilung nach dem Median (Median=3) der Anzahl an Informationsquellen => 32 Versuchspersonen (38%) in der Kategorie „wenig Informationen“ und 52 Versuchspersonen (62%) in der Kategorie „viele Informationen“) und „Konsum von Bravo und BravoGirl“ (Teilung nach „lese ich heute“ & „habe ich früher häufig gelesen“ => 34 Versuchspersonen (40,5%) in der Kategorie „Konsum“ sowie „habe ich noch nie gelesen“ & „habe ich früher selten gelesen“ => 50 Versuchspersonen (59,5%) in der Kategorie „kein Konsum“) untersucht.

Zwischen dem Informationsstand der Jugendlichen und den abhängigen Variablen lassen sich keine signifikanten Zusammenhänge nachweisen. Auch der Bravo-Konsum hatte keinen Einfluß auf das Antwortverhalten der Jugendlichen – mit Ausnahme der dargestellten Skripte: Obwohl sie angeben, Bravo aktiv zu lesen, schildern Leser/Leserinnen zu 90% ein „Fremdtäter-Skript“ und nur zu 10% ein "Innerfamiliäres Mißbrauch – Skript". Dagegen beträgt das Verhältnis bei Nicht-Leser/Leserinnen 70% „Fremdtäter“- zu 30% "Innerfamiliärem Mißbrauch - Skript". Da die BravoGirl-Artikel zu 97% über innerfamiliäre Mißbrauchsszenarien berichten, liegen die Leser/Leserinnen nicht näher an den Ergebnissen der BravoGirl-Analyse, sondern eher weiter davon entfernt.

## 4 Diskussion der Ergebnisse

Die Ergebnisse der vorliegenden Arbeit zeigen, daß die Schüler einzelne Fakten zu Täter-Opfer-Konstellation und Ausmaß (Prävalenzrate, Dunkelziffer) verglichen mit den aus der forensischen Literatur bekannten Daten (Bange & Deegener, 1996; Amann & Wippliner, 1997; PSK, 1999, 2000) sehr gut einschätzen konnten.

Die Aufforderung, sich den Ablauf einer typischen Situation von sexuellem Mißbrauch vorzustellen und eine Liste der Handlungen aufzustellen, führt jedoch überwiegend zu der Schilderung eines Fremdtäterüberfalls. Da kein Erlebnishintergrund vorhanden ist, muß das Wissen aus dem deklarativen Gedächtnis konstruiert werden – ein Vorgang, mit dem die Jugendlichen offensichtlich überfordert sind.

Trotzdem war es möglich, einen sinnvollen Ablauf zu ermitteln, und auch quantitativ sind die erhaltenen Skripte durchaus mit anderen Studien vergleichbar. Die kleinen N verursachten natürlich vergleichsweise geringe Gesamtitemzahlen, die Zahl der Ereignisse, die nur von einer Versuchsperson genannt wurden, liegt verglichen mit den Untersuchungen von Bower, Black & Turner (1979) und Holst & Pezdek (1992) jedoch sehr niedrig: Während Holst & Pezdek bis zu 44

einmalige Nennungen verzeichneten, liegt diese Anzahl in der vorliegenden Arbeit bei maximal 9. Auch im Vergleich der Quotienten aus Gesamtzahl der genannten Ereignisse zu Einmalnennungen liegen die Daten dieser Untersuchung im Rahmen der Ergebnisse von Bower, Black & Turner (1979) und Holst & Pezdek (1992).

Als das homogenste Skript hat sich das „Innerfamiliäre Mißbrauchs – Skript“ der Jugendlichen erwiesen. Die „Innerfamiliären Mißbrauchs - Skripte“ der Jugendlichen sind anfangs (Handlungen 1-6) eher heterogen, werden dann aber zunehmend homogen; d.h. die Jugendlichen unterscheiden sich in ihren Annahmen bezüglich der Annäherungsstrategien, die Taten selbst und ihr Fortgang wird dagegen relativ einheitlich beschrieben.

Der von Sgroi (1982) empirisch ermittelte 5-Phasen-Ablauf spiegelt sich allerdings trotz der Anweisung, die Handlungen vom Erstkontakt von Täter/Täterin und Opfer bis zum Ende des Mißbrauchs zu berücksichtigen, nicht in den Handlungsketten der Jugendlichen wider. Diese beschränken sich im wesentlichen auf die Phasen 2 (sexuelle Interaktionsphase) und 3 (Geheimhaltungsphase), wobei die high frequency-high importance events des innerfamiliären Mißbrauchs-Skripts SCHULE (GEWALT, TÄTER BERÜHRT OPFER, DROHUNG und PENETRATION) alle ein Gewaltpotential enthalten. Mit Hinblick auf die delikt-spezifischen Merkmale (Arntzen, 1983) überbetonen auch die sexuellen Interaktionen, die als typisch angegeben werden, den gewaltsamen Aspekt.

Das Experten-Skript war nicht homogener als die anderen Skripte, sondern eher inhomogener. Die Schwierigkeit für die Experten lag vermutlich darin, mit einer Vielzahl sehr unterschiedlicher Mißbrauchssituationen vertraut zu sein, so daß daraus keine „typische“ Situation abstrahiert werden konnte. Eine Expertin bemerkt auch bei der Frage nach dem Skript: „Kann ich nicht beantworten. Es gibt für mich keine „typische“ Mißbrauchssituation. Mißbrauch kann auf so vielfältige Art und Weise stattfinden“.

Die zum Vergleich durchgeführte Analyse von Artikeln der Jugendzeitschrift BravoGirl zeigt, daß die dort geschilderten „Erlebnisberichte“ eine überwiegend korrekte Täter-Opfer-Konstellation und einen kohärenten Tatverlauf präsentieren. Sie enthalten aber im wesentlichen nur die sexuellen Handlungen selbst, die „Vorgeschichte“ (d.h. wie erfolgt die Annäherung, wann und wie kommt es zum Mißbrauch?) fehlt. Die BravoGirl-Skripte sind im ersten (Handlungen 2 – 8) und letzten Teil (ab Handlung 12) sehr homogen, im mittleren Teil (Handlungen 9-11) werden die Angaben inhomogener, da hier mehr Details geschildert werden. Die wesentlichen Elemente (GELEGENHEIT, TATORT, KÜSSEN und TÄTER BERÜHRT OPFER) sind in allen BravoGirl-Artikeln vorhanden. Insgesamt spricht dieses Ergebnis für die Konstruktion der Zeitungsartikel, entweder als fiktive „Erlebnisberichte“ oder aber zumindest für eine eingehende Revidierung der ursprünglichen Erlebnisberichte.

Für die Aussagepsychologie ist von Bedeutung, daß alle Skripte ein Schweigegebot sowie Strategien, die der Verharmlosung der Tat dienen (in den Skripten zusammengefaßt als „Tat normal“) enthalten. Wurde bislang die Erwähnung eines Schweigegebots in Zeugenaussagen als delikt-spezifisches Merkmal besonderer Qualität und damit als wichtiger Indikator für die Glaubhaftigkeit einer Aussage gewertet, zeigt dieses Ergebnis, daß „Schweigegebot“ zum „gedanklichen Allgemeingut“ (Greuel, 1998, S. 100) Jugendlicher zählt. Allerdings waren die von den Jugendlichen beschriebenen Schweigegebote sehr einfach gehalten und relativ vage. Trotzdem bedeutet dies, daß andere, trennschärfere Merkmale gefunden werden müssen, wozu es weiterer Forschungsarbeit bedarf.

Die vorliegenden Ergebnisse liefern wertvolle Hinweise, um Falschaussagen besser diskriminieren zu können. Intentional falschaussagende Zeugen werden sich auf Grundlage ihres Wissens demnach vermutlich bemühen, Handlungen zu schildern, die sehr intensiv sind. In einem Umkehrschluß kann deshalb das Vorhandensein von Schilderungen weniger intensiver Handlungen für die Glaubhaftigkeit der Zeugenaussage sprechen.

Der willentliche oder versehentliche Transfer von Informationen aus Fremdquellen führt in der vorliegenden Untersuchung nicht zu einer Veränderung der Ergebnisse. Der Sonderfall „Bravo“ hat sogar konservative Auswirkungen: Die formulierten Skripte von Versuchspersonen, die angeben,

Bravo gelesen zu haben, zeigen keinerlei Affinitäten mit den Inhalten der untersuchten BravoGirl-Artikel und führen allenfalls zur Schilderung von Fremdtäterdelikten, die als Falschaussage zu identifizieren sind.

## Literatur

- Abelson, R.P. (1981). Psychological Status of the Script Concept. *American Psychologist*, 36, 715-729.
- Amann, G. & Wipplinger, R. (1997). *Sexueller Mißbrauch. Überblick zu Forschung, Beratung und Therapie*. Tübingen: dgvt Verlag.
- American Psychological Association (1998). *Diagnostisches und Statistisches Manual Psychischer Störungen DSM IV*. Göttingen: Hogrefe.
- Arntzen, F. (1983). *Psychologie der Zeugenaussage. System der Glaubwürdigkeitsmerkmale*. München: C.H. Beck'sche Verlagsbuchhandlung.
- Atkinson, R.C. & Shiffrin, R.M. (1971). The Control of Short-Term Memory. *Scientific American*, 225, 82-90.
- Bange, D. & Deegener, G. (1996). *Sexueller Mißbrauch von Kindern - Ausmaß, Hintergründe, Folgen*. Weinheim: Beltz Verlag.
- Bower, H.H., Black, J.B. & Turner, T.J. (1979). Scripts in Comprehension and Memory. *Cognitive Psychology*, 11, 177-220.
- Brockhaus, U. & Kolshorn, U. (1993). *Sexuelle Gewalt gegen Mädchen und Jungen. Mythen, Fakten, Theorien*. Frankfurt: Campus.
- Ceci, S. & Bruck, M. (1995). *Jeopardy in the Courtroom. A scientific analysis of children's testimony*. Washington: APA.
- Ernst, C. (1997). Zu den Problemen der epidemiologischen Erforschung des sexuellen Mißbrauchs. In G. Amann & R. Wipplinger (Hrsg.). *Sexueller Mißbrauch. Überblick zu Forschung, Beratung und Therapie* (S. 55-71). Tübingen: dgvt Verlag.
- Fisher-Holst, V. & Pezdek, K. (1992). Scripts for Typical Crimes and their Effects on Memory for Eyewitnesses Testimony. *Applied Cognitive Psychology*, 6, 573 - 587.
- Gomes-Schwartz, B. et al. (1990). *Child Sexual Abuse. The Initial Effects*. London: Sage Publications.
- Greenberg, M., Westcott, D. & Scott, B. (1998). When Believing is Seeing: The Effect of Scripts on Eyewitnesses Memory. *Law and Human Behavior*, 22, 685 - 694.
- Greuel, L. et al. (1998). *Glaubhaftigkeit der Zeugenaussage. Theorie und Praxis der forensisch-psychologischen Begutachtung*. Weinheim: Beltz PsychologieVerlagsUnion.
- Holst, V.F. & Pezdek, K. (1992). Scripts for typical Crimes and their Effects on Memory for eyewitness Testimony. *Applied Cognitive Psychology*, 6, 573-587.
- Jönsson, E. (1997). *Interventionen bei sexuellem Mißbrauch: ein europäischer Vergleich am Beispiel ausgewählter Länder*. Frankfurt am Main: Lang.
- Kavemann, B. & Lohstötter, I. (1986). *Väter als Täter: Sexuelle Gewalt gegen Mädchen*. Reinbek: Rowohlt.
- Knappe, A. & Selg, H. (1993). *Prävention von sexuellem Mißbrauch an Mädchen und Jungen* (Forschungsbericht). München: Bayerisches Staatsministerium für Arbeit und Sozialordnung, Familie, Frauen und Gesundheit.
- Littmann, E. & Szewczyk, H. (1983). Zu einigen Kriterien und Ergebnissen forensisch-psychologischer Begutachtung von sexuell mißbrauchten Kindern und Jugendlichen. *Forensia*, 4, 55-72.
- Markowitsch, H.J. (1992). *Neuropsychologie des Gedächtnisses*. Göttingen: Hogrefe.
- Mayring, Ph. (1988). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim: Deutscher Studien Verlag.

- Raupp, U. & Eggers, Ch. (1993). Sexueller Mißbrauch von Kindern. Eine regionale Studie über Prävalenz und Charakteristik. *Monatszeitschrift Kinderheilkunde*, 141, 316 – 322.
- Schank, R.C. & Abelson, R.P. (1997). *Scripts, Plans, Goals and Understanding*. Hillsdale, NJ: Erlbaum.
- Sgroi, S. (Ed.) (1982). *Handbook of Clinical Intervention in Child Sexual Abuse*. Lexington, MA: Lexington Books.
- Steller, M. & Volbert, R. (1999). *Wissenschaftliches Gutachten Forensisch-aussagepsychologische Begutachtung (Glaubwürdigkeitsbegutachtung)*. Im Druck.
- Steller, M. & Köhnken, G. (1989). Criteria-based statement analysis. In C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 217-245). New York: Springer.
- StGB (Strafgesetzbuch). Textausgabe mit ausführlichem Sachregister*. Stand 1. April 1999. 33. Auflage. München: dtv-Verlag.
- Trankell, A. (1971). *Der Realitätsgehalt von Zeugenaussagen. Methoden der Aussagepsychologie*. Göttingen: Vandenhoeck & Ruprecht.
- Trankell, A. (1982). *Reconstructing the Past. The Role of Psychologists in Criminal Trials*. Deventer: Kluwer.
- Undeutsch, U. (1967). Die Beurteilung der Glaubhaftigkeit von Aussagen. In U. Undeutsch (Hrsg.), *Handbuch der Psychologie, Band 11: Forensische Psychologie* (S. 26-181). Göttingen: Hogrefe.
- Whipple, G. M. (1911). The Psychology of Testimony. *Psychological Bulletin*, 8, 307-309.

Polizeiliche Kriminalstatistik im Internet: <http://www.bka.de/pks/pks1999/index2.html>

<sup>1</sup> „Externalisiert“= Aggression, Verhaltensprobleme, selbstzerstörerisches Verhalten, unangebrachtes Sexualverhalten und Externalisierung als zusammengesetztes Symptom

<sup>2</sup> „Internalisiert“= Angst, Furcht, Depression, geringer Selbstwert, somatische Beschwerden, Angst vor sexuellen Kontakten, Schul-/Lernprobleme und Internalisierung als zusammengesetztes Symptom



# Gründe für das Verschweigen sexuellen Mißbrauchs – Ergebnisse einer Analyse der Aussagen von Opferzeugen

*Katrin Behruzi & Udo Undeutsch*  
*Psychologisches Institut der Universität Köln*

Die Deliktsspezifität einer Aussage gilt als eines der Glaubhaftigkeitsmerkmale bei der Glaubhaftigkeitsbegutachtung. Als ein solches deliktsspezifisches Verhaltensmuster des Opfers bei sexuellem Mißbrauch kann das Verschweigen des Mißbrauchs angesehen werden, das Kinder zum Teil über Jahre beibehalten. Die Basis dieses Verhaltensmusters sind die Gründe für das Verschweigen, also solche Argumente, die beim Abwägen negativer Konsequenzen eines Bekanntwerdens des Mißbrauchs ins Gewicht fallen und das Kind dazu bringen, das Geschehene bewußt und absichtsvoll zu verbergen. Im vorliegenden Artikel geht es um die verschiedenen Gründe, die Kinder veranlassen, keiner dritten Person etwas über einen stattgefundenen sexuellen Mißbrauch zu sagen. Diese Argumente können einerseits vom Täter geäußert und andererseits vom Kind selbst überlegt werden. Auf einer Datenbasis von 64 – sowohl von erfahrenen Sachverständigen als auch vom Gericht – als glaubhaft eingestuften Fällen sexuellen Mißbrauchs wurde ein Kategoriensystem zur Erfassung von Gründen für das Verschweigen erstellt. Es differenziert zwischen Schweigegeboten, die entweder vom Täter ausgesprochen wurden oder die sich das Kind selbst auferlegt hat. Insgesamt 24 verschiedene Gründe für das Verschweigen konnten unterschieden werden, wobei von einzelnen Zeugen bis zu fünf verschiedene Gründe geschildert wurden. Die verschiedenen Schweigegebote werden jeweils inhaltlich beschrieben. Unter Verwendung des gewonnenen Kategoriensystems werden zudem statistische Analysen durchgeführt, bei denen die Gründe für das Verschweigen zu folgenden weiteren Variablen in Beziehung gesetzt werden: Alter des Kindes bei der letzten von ihm berichteten Tat, Geschlecht des Kindes, Täter-Opfer-Beziehung in ihrem formalen Verwandtschafts-/Bekanntschaftsgrad und die Häufigkeit der Vorfälle.

## **1 Einführung und Fragestellung**

Forensisch-psychologische Sachverständige werden vor allem zur Begutachtung der Aussagen kindlicher und jugendlicher Zeugen bei Straftaten gegen das sexuelle Selbstbestimmungsrecht und in familien- bzw. vormundschaftsrechtlichen Verfahren, bei denen der Vorwurf des sexuellen Mißbrauchs im Raum steht, herangezogen. Zur Beurteilung des Wahrheitsgehaltes von Zeugenaussagen gilt die kriteriengestützte Aussageanalyse (Undeutsch, 1967) als ein valides Beurteilungsinstrument und wird heute üblicherweise zur aussagepsychologischen Begutachtung genutzt. Undeutsch (1967) stellte die Hypothese auf, daß sich wahre (erlebnisgebundene) von falschen (nicht erlebnisgebundenen) Aussagen qualitativ unterscheiden. Ausgehend von dieser Hypothese wurden Merkmalskataloge für die Glaubhaftigkeit von Zeugenaussagen entwickelt (Undeutsch, 1967; Trankell, 1971; Arntzen, 1983; Steller & Köhnken, 1989; Greuel et al., 1998). Es handelt sich hierbei um aufgelistete Kennzeichen erlebnisfundierter Aussagen, die eine Unterscheidung von erlebnisgebundenen und nicht-erlebnisgebundenen Aussagen ermöglichen sollen. Diese Kennzeichen werden als Glaubhaftigkeitsmerkmale bezeichnet. Als eines der Glaubhaftigkeitsmerkmale gilt die Schilderung deliktsspezifischer Details. Dieses Kriterium soll in der vorliegenden Arbeit behandelt werden.

Arntzen (1983) definiert eine „deliktypische Schilderung“ wie folgt:

„Mit einer deliktypischen Schilderung ist gemeint, daß ein Zeuge ein Verhaltensmuster wiedergibt, das für ein bestimmtes Delikt typisch ist, das der Zeuge aber nicht kennen kann, ohne das beschriebene Erlebnis wirklich gehabt zu haben“ (a.a.O., S. 145).

Auch Greuel et al. (1998) weisen darauf hin, daß es nicht darum gehe,

„das Vorhandensein einzelner deliktspezifischer Angaben im Sinne eines Einzeldetails zu bestätigen, sondern um die – zumeist implizite – Skizzierung eines komplexen und übergeordneten Verhaltensmusters, über dessen Bedeutung sich der Zeuge nicht bewußt ist“ (a.a.O., S. 99).

Als für sexuellen Mißbrauch deliktspezifisch können beispielsweise Strategien, mit denen Täter versuchen, Kinder zur Teilnahme an sexuellen Handlungen zu bewegen, angesehen werden (Klaffer & Undeutsch, 2000). Immer wieder als deliktspezifisch genannt wird das Schweigegebot, das dem Kind vom Täter auferlegt werde. Die isolierte Nennung eines solchen Schweigegebots kann jedoch nicht als Glaubhaftigkeitsmerkmal gewertet werden, insbesondere wegen der weiten Verbreitung des Wissens über ein solches Schweigegebot in der Bevölkerung. So weisen beispielsweise Greuel et al. (1998) darauf hin,

„daß die Schilderung eines einzelnen deliktspezifischen Details (z.B. die bloße Angabe von Schweigegebots im Kontext eines inkriminierten Mißbrauchsdelikts) oder auch die Angabe von wenigen, nicht wechselseitig aufeinander bezogenen deliktspezifischen Details den Anforderungen nicht genügt, insbesondere dann nicht, wenn es sich aufgrund der starken Medienpräsenz der Mißbrauchsthematik um solche Angaben handelt, die zum ‚gedanklichen Gemeingut‘ gehören und von daher auch von falschaussagenden Zeugen vorzubringen wären“ (S. 99-100).

Betrachtet man jedoch das Verschweigen des Mißbrauchs als ein Verhaltensmuster, das Kinder zum Teil über mehrere Jahre beibehalten, so kann das Verhaltensmuster „Verschweigen“ als deliktspezifisch gelten. Smith et al. (2000) fanden beispielsweise in einer repräsentativen Befragung erwachsener Frauen, daß die Mehrheit den Mißbrauch nicht unmittelbar nach dem ersten Vorfall mitteilte. Mehr als die Hälfte der Frauen, die sich an ihr erstmaliges Mitteilen erinnerten, warteten sogar länger als 8 Jahre. Gegenstand der vorliegenden Arbeit sind die Gründe, die das Kind veranlassen, keiner dritten Person etwas über den sexuellen Mißbrauch zu sagen. Es geht also um Argumente, die das Kind dazu bringen, das Geschehene bewußt und absichtsvoll zu verbergen. Diese Gründe für das Verschweigen stellen die Grundlage für das Verhaltensmuster „Verschweigen“ dar.

Bullens (1995), der eine niederländische Studie von Aarts vorstellt, bei der Täter befragt wurden, berichtet darüber, daß 69% der Täter ein explizites Schweigegebot ausgesprochen hätten. 31% der Täter seien davon ausgegangen, daß das Opfer – ohne explizit zum Schweigen aufgefordert zu werden – von sich



aus schweigen würde. Will man also etwas über die Gründe für das Verschweigen des sexuellen Mißbrauchs erfahren, so darf man nicht nur die vom Täter geäußerten Schweigegebote betrachten, sondern muß auch solche Argumente berücksichtigen, die sich das Kind selbst überlegt.

Die folgenden Ausführungen verfolgen das Ziel, einen Beitrag zur Beschreibbarkeit des für sexuellen Mißbrauch deliktspezifischen Merkmals „Verschweigen“ zu leisten. Es werden Kategorien vorgestellt, die es im Sinne einer Strukturierungshilfe dem Gutachter ermöglichen sollen, die Gründe für das Verschweigen als Grundlage dieses Verhaltensmusters zu erfassen. Zudem wird erst durch eine Kategorisierung empirische Überprüfbarkeit möglich. Neben der Darstellung von Gründen für das Verschweigen werden Ergebnisse statistischer Analysen vorgestellt.

## 2 Methode

Die Datenbasis zur Identifizierung der Gründe für das Verschweigen bilden Explorationstranskriptionen und Gutachten von 64 sowohl von erfahrenen Sachverständigen als auch von Gerichten als glaubhaft eingestuften Fällen der Jahre 1990 bis 1999 aus der Gutachtenpraxis von Prof. U. Undeutsch und Frau Dipl.-Psych. Klein (Psychologisches Institut der Universität zu Köln). Es wurden zunächst Textstellen identifiziert, die Gründe für das Verschweigen beschreiben. Diese wurden zunächst danach unterschieden, ob es sich um ein vom Täter auferlegtes oder selbstaufgelegtes Schweigegebot handelt. Ähnliche Schweigegebote wurden alsdann zu Untergruppen zusammengefaßt. Diese auf induktivem Wege gewonnenen Einheiten werden jeweils beschrieben.

Der Abschnitt „Statistische Analysen“ enthält Häufigkeitsauszählungen der einzelnen Gründe für das Verschweigen. Zur weiteren Analyse wird die Anzahl genannter Schweigegebotskategorien und die Unterscheidung in vom Täter auferlegte oder selbstaufgelegte Schweigegebote verwendet. Diese Variablen werden zu folgenden Variablen in Beziehung gesetzt: Alter des Kindes bei der letzten von ihm berichteten Tat, Geschlecht des Kindes, die Täter-Opfer-Beziehung in ihrem formalen Verwandtschafts-/Bekanntschaftsgrad und die Häufigkeit der Vorfälle. Die Stärke von Zusammenhängen zwischen der Art des Schweigegebots (vom Täter auferlegt oder selbstaufgelegt) bzw. der Anzahl genannter Schweigegebote und diesen Variablen wird mittels Zusammenhangsmaßen (Cramers V, Spearmans Rangkorrelationskoeffizient  $r_s$ ) dargestellt. Zur Überprüfung der Signifikanz von Zusammenhängen werden der Pearsonsche Chi-Quadrat-Test bzw. ein auf einem angenäherten t-Wert aufbauender Test herangezogen.

## 3 Ergebnisse

### 3.1 Inhaltliche Beschreibung der gewonnenen Kategorien

#### 3.1.1 Durch den Täter ausgesprochene Schweigegebote

**Explizites Verbot, darüber mit jemandem zu sprechen:** Der Täter fordert das Kind auf, mit niemandem über die vorgefallenen sexuellen Handlungen zu

sprechen, ohne daß er negative Konsequenzen für den Fall ankündigt, daß sich das Kind seinem Verbot widersetzt.

**Explizites Verbot mit der Drohung, es gäbe sonst Ärger:** Dem Kind wird für den Fall der Nichtbeachtung des auferlegten Schweigegebots angedroht, es bekäme Ärger. Bei dieser Drohung wird nicht weiter ausgeführt, worin die angekündigten Unannehmlichkeiten konkret bestehen.

**Explizites Verbot mit Gewaltandrohung:** Der Täter spricht ein Schweigegebot aus und droht für den Fall des Nichtbeachtens mit Gewalt, beispielsweise mit Schlägen. Mitunter wird dem Schweigegebot durch das drohende Zeigen eines Messers Nachdruck verliehen. Verschiedene Täter gingen so weit, Todesdrohungen für den Fall auszusprechen, daß das Kind das Schweigegebot nicht befolgt. Die Gewaltdrohung richtet sich nicht in jedem Fall gegen das Kind, sondern kann auch gegen eine nahestehende dritte Person, beispielsweise die Mutter des Kindes, gerichtet sein. Es lassen sich auch Fälle finden, in denen der Täter mit von dritten ausgehender Gewalt droht.

**Explizites Verbot mit der Drohung negativer Konsequenzen für das Familienleben:** Der Täter droht damit, daß sich im Falle des Bekanntwerdens des Mißbrauchs die Eltern streiten würden oder die Familie auseinanderbrechen würde.

**Explizites Verbot mit Erpressung:** Eine Drohung kann auch in Form einer Erpressung auftauchen. Bei einer Erpressung ist für das Opfer immer offensichtlich, daß der Täter die Macht hat, die angedrohte negative Folge herbeizuführen.

**Explizites Verbot mit Verharmlosung:** Der Täter begründet das Schweigegebot damit, daß die Handlungen es nicht wert wären, darüber zu sprechen. Sie seien nichts Besonderes, sondern ganz normal.

**Explizites Verbot mit Hinweis auf negative Konsequenzen für den Täter:** Der Täter verdeutlicht dem Kind, daß er selbst im Falle des Bekanntwerdens mit negativen Konsequenzen rechnen muß. Dem Kind wird eingeredet, daß es die Verantwortung für das weitere Wohlergehen des Täters hat und hofft darauf, daß das Kind aus Vermeidung des schlechten Gewissens, für die negativen Konsequenzen der Auslöser gewesen zu sein, schweigt.

**Explizites Verbot mit Hinweis auf negative Konsequenzen für eine nahestehende dritte Person:** Der Täter weist auf negative Konsequenzen für eine nahestehende dritte Person, die aus dem Bekanntwerden des Mißbrauchs resultieren würden, hin. Er hofft dabei darauf, daß das Kind, um diese Konsequenzen zu verhindern, über die sexuellen Handlungen schweigt.

**Hinweis, dem Kind würde nicht geglaubt werden:** Der Täter versucht, das Bekanntwerden der Vorfälle dadurch zu verhindern, daß er dem Kind ankündigt, ihm würde niemand glauben. In einigen Fällen kündigt der Täter dem Kind nicht nur an, daß diesem nicht geglaubt wird, sondern malt die Folgen dessen dramatisch aus, beispielsweise entwirft er ein Bild, in dem das Kind wegen Verleumdung ein Leben lang von einem Richter verfolgt würde.

**Das Kind hätte Mitschuld:** Der Täter behauptet eine Mitschuld bzw. Schuld des Opfers. Folgt man dieser Annahme, so verrät das Kind mit der Bekanntgabe des Mißbrauchs nicht nur etwas über das unerlaubte Verhalten des Täters, sondern steht selbst in einem negativen Licht. Der Täter kann folglich darauf hoffen, daß das Kind das Schweigegebot einhält, um sein eigenes Ansehen zu wahren.

**Erzeugen einer Angst vor Ausgrenzung:** Der Täter verdeutlicht dem Kind, daß es im Falle einer Mitteilung des Vorgefallenen damit rechnen muß, von anderen ausgegrenzt zu werden. Diese Vorstellung ist dem Motiv nach sozialem Anschluß entgegengesetzt, welches besagt, daß das Handeln des Menschen auf Annahme und Zuwendung durch soziale Partner gerichtet ist und Zurückweisung versucht wird zu vermeiden.

**Bestechung:** Der Täter versucht, das Kind durch Geschenke, beispielsweise Geld oder Kleidung, zur Verschwiegenheit zu verpflichten.

**Durch das Bekanntwerden des Mißbrauchs würde auch eine Regelüberschreitung des Kindes preisgegeben werden:** Der Täter weist das Kind darauf hin, daß durch das Bekanntwerden des Mißbrauchs auch eine Regelüberschreitung des Kindes bekannt würde. Die Folgen des unerlaubten Handelns des Kindes werden überhöht dargestellt. Beispielsweise wird dem Kind vorgehalten, daß es ohne Erlaubnis der Eltern ins Schwimmbad ging, wo die sexuellen Handlungen stattfanden, und daß das Bekanntwerden des unerlaubten Schwimmbadbesuchs zu großem Ärger führen würde.

### 3.1.2 Selbstaufgelegte Schweigegebote

**Angst vor Ärger:** Das Kind verheimlicht die Vorfälle, weil es Angst hat, Ärger zu bekommen, zum Beispiel von der Mutter geschimpft zu werden.

**Selbstaufgelegt, um negative Konsequenzen für nahestehende Dritte zu verhindern:** Das Kind schweigt, weil es negative Konsequenzen von nahestehenden Dritten fernhalten will. Beispielsweise legt es sich ein Schweigegebot auf, um die Mutter zu schonen.

**Selbstaufgelegt wegen der Befürchtung negativer Konsequenzen für das Familienleben:** Das Kind legt sich selbst ein Schweigegebot auf, um ein Auseinanderbrechen familiärer Strukturen zu verhindern. Beispielsweise schweigt das Kind der guten Beziehung zwischen der Mutter und dem Täter wegen oder aus Angst, selbst die Familie verlassen zu müssen.

**Angst des Kindes, daß ihm nicht geglaubt wird:** Das Kind befürchtet, daß ihm nicht geglaubt wird und vermeidet es daher, mit jemandem über die sexuellen Handlungen zu sprechen.

**Sich mitschuldig fühlen:** Das Kind rechnet es sich selbst zu, daß es zu den Handlungen gekommen ist. Eine Offenbarung des Mißbrauchs würde in den Augen des Kindes ein nicht gerechtfertigtes Schuldzuweisen auf den Täter bedeuten. Das Kind überschätzt hierbei völlig die eigene Entscheidungs- und Handlungsfreiheit.

**Angst vor Ausgrenzung:** Das Kind legt sich selbst ein Schweigegebot auf, weil es fürchtet, im Falle des Bekanntwerdens von anderen ausgegrenzt bzw. abgewiesen zu werden.

**Durch Bekanntwerden des Mißbrauchs würde auch eine Regelüberschreitung des Kindes bekanntwerden:** Dem Kind ist bewußt, daß mit der Bekanntgabe des sexuellen Mißbrauchs auch ein unerlaubtes Verhalten von ihm selbst bekannt werden würde, was mit dem Mißbrauch im engeren Sinne nicht zu tun hat. Um die eigene Regelüberschreitung zu verheimlichen, sagt es dann auch nichts über das unerlaubte Verhalten des Täters.

**Kein geeigneter Ansprechpartner vorhanden:** Das Kind ist der Meinung, es gäbe keinen geeigneten Gesprächspartner, mit dem ein solches Thema besprochen werden könnte.

**Als angenehm empfundene Beziehung aufrechterhalten:** Dem Kind ist bewußt, daß eine sexuelle Beziehung zwischen einem Erwachsenen und einem Kind verboten ist. Es möchte jedoch diese Beziehung fortführen, da diese mit Annehmlichkeiten verbunden ist oder das Kind sie als Liebesbeziehung ansieht. Daher versucht es, ein Bekanntwerden zu verhindern.

**Verharmlosung:** Das Kind tut das Geschehnis als belanglos ab und spricht daher nicht darüber.

**Vergessen wollen:** Das Kind entscheidet sich dazu, nicht über die Vorfälle zu sprechen, um diese zu vergessen.

## 3.2 Statistische Analysen

### 3.2.1. Häufigkeit der einzelnen Schweigegebote

Tabelle 1 hat die Häufigkeit, mit der verschiedene Gründe für das Verschweigen berichtet wurden, zum Gegenstand. Die Spalte „% der Nennungen“ gibt an, welchen Anteil die einzelne Kategorie an allen Nennungen hat. Die Spalte „% der Fälle“ gibt an, wieviel Prozent der Kinder Aussagen machten, die sich in eine der Kategorien einordnen lassen.

Tabelle 1: Häufigkeit der einzelnen Schweigegebote

Kategorie	absolut	% der Nennungen	% der Fälle
<b>Durch den Täter ausgesprochene Schweigegebote</b>			
Explizites Verbot	20	22.2	31.3
Drohung: Ärger	4	4.4	5.3
Drohung: Gewalt	10	11.1	15.6
Drohung: Familienleben	3	3.3	4.7
Drohung: Erpressung	2	2.2	3.2
Explizit mit Verharmlosung	3	3.3	4.7
Neg. Konsequenzen für den Täter	2	2.2	3.1
Neg. Konsequenzen für Dritte	1	1.1	1.6
Dem Kind wird nicht geglaubt	5	5.6	7.8
Das Kind hätte Mitschuld	1	1.1	1.6
Angst vor Ausgrenzung erzeugen	3	3.3	4.7
Bestechung	1	1.1	1.6
Regelüberschreitung des Kindes	1	1.1	1.6
<b>Selbstaufgelegte Schweigegebote</b>			
Angst vor Ärger	5	5.6	7.8
Neg. Konsequenzen für Dritte	4	4.4	6.3
Neg. Konsequenzen f. d. Familienleben	2	2.2	3.1
Mir wird nicht geglaubt	5	5.6	7.8
Sich mitschuldig fühlen	1	1.1	1.6
Angst vor Ausgrenzung	2	2.2	3.1
Regelüberschreitung des Kindes	2	2.2	3.1
Kein Ansprechpartner vorhanden	2	2.2	3.1
Angenehme Beziehung	4	4.4	6.3
Verharmlosung	1	1.1	1.6
Vergessen wollen	1	1.1	1.6
Keine Angaben zum Schweigegebot	19	5.6	7.8
Summe	90	100.0	140.6

### 3.2.2 Anzahl genannter Schweigegebote

Abbildung 1 stellt dar, wie sich die Anzahl genannter Schweigegebotskategorien verteilt. Es zeigt sich, daß 67% aller Kinder von einer einzigen Art des Schweigegebots gesprochen haben. Die Schilderungen von 25% der Kinder enthielten mehrere Kategorien von Gründen für das Verschweigen. Nur 8% der Kinder berichteten nicht von Gründen für das Verschweigen. Hierbei handelte es

sich fast ausschließlich um Vorfälle, die vom Kind nur kurze Zeit nach der Tat berichtet wurden, so daß man nicht vom Verhaltensmuster Verschweigen bei diesen Kindern sprechen kann. Dem entsprechend lagen bei diesen Kindern auch keine Gründe für das Verschweigen vor.

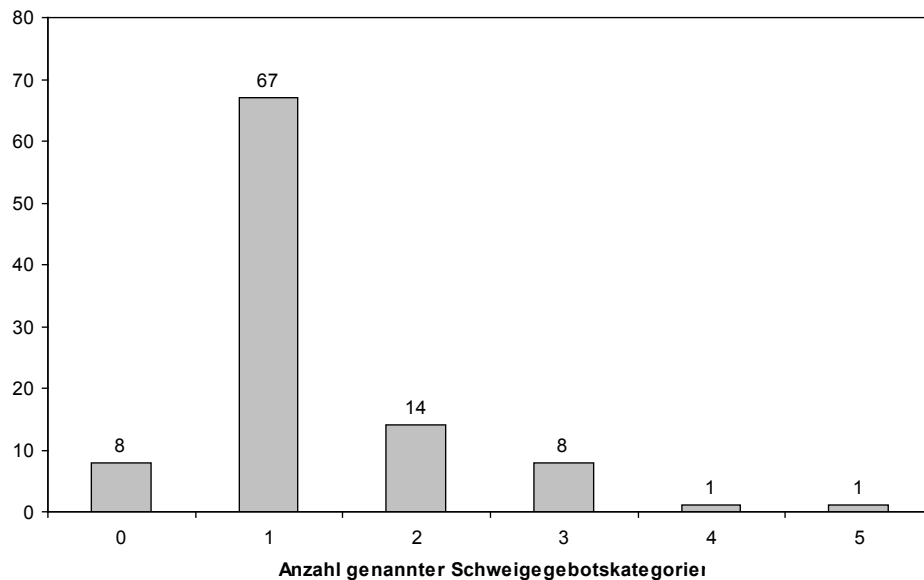


Abb. 1: Häufigkeit der Anzahl genannter Schweigegebotskategorien

Mit größerer Anzahl an Vorfällen wurde auch eine größere Anzahl an Gründen für das Verschweigen des sexuellen Mißbrauchs berichtet ( $r_s=0.365$ ,  $p=0.005$ ). Auch das Alter der Kinder und die Anzahl genannter Gründe für das Verschweigen korrelieren signifikant ( $r_s=0.476$ ,  $p=0.000$ ). Hinsichtlich der Täter-Opfer-Beziehung und der Anzahl an Schweigegebotskategorien gab es keine signifikanten Unterschiede ( $V=0.250$ ,  $p=0.521$ ). Auch zwischen Jungen und Mädchen lassen sich keine Unterschiede im Hinblick auf die Anzahl an Schweigegebotskategorien finden ( $V=0.224$ ,  $p=0.565$ ).

### 3.2.3 Vom Täter auferlegte und selbstauferlegte Schweigegebote

Die einzelnen Schweigegebote werden im folgenden zu größeren Gruppen zusammengefaßt. Es wird unterschieden, ob die Schweigegebote, die ein Kind mitteilte, allein vom Täter auferlegt wurden („nur Täter“), allein selbstauferlegt wurden („nur Kind“) oder ob ein Kind sowohl über ein vom Täter als auch über ein selbstauferlegtes Schweigegebot berichtete („Täter und Kind“).

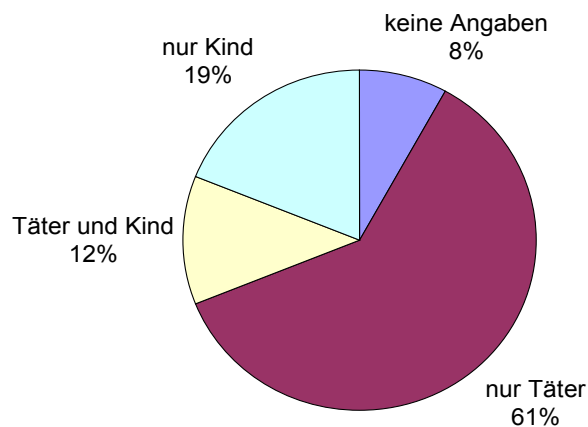


Abb. 2: Häufigkeit des vom Täter auferlegten und selbstaufgelegten Schweigegebots

Mit steigendem Alter des Kindes nimmt der Anteil der allein vom Täter ausgesprochenen Schweigegebote ab während gleichzeitig der Anteil der selbstaufgelegten Schweigegebote zunimmt ( $V=0.378$ ,  $p=0.009$ ). Hinsichtlich der Anzahl an Vorfällen zeigen sich keine Unterschiede zwischen den verschiedenen Schweigegebotsarten ( $V=0.081$ ,  $p=0.943$ ). Auch Jungen und Mädchen unterschieden sich nicht darin, ob sie ein vom Täter auferlegtes und/oder selbstaufgelegtes Schweigegebot berichteten ( $V=0.105$ ,  $p=0.724$ ). Über fremde Täter wurde immer ein durch den Täter ausgesprochenes Schweigegebot berichtet, wohingegen die Opfer der anderen Tätergruppen in 23% der Fälle nur selbstaufgelegte Schweigegebote mitteilten. Es ließ sich jedoch kein statistisch signifikanter Zusammenhang zwischen der Art des Schweigegebots und der Täter-Opfer-Beziehung (Fremde – alle anderen Tätergruppen) nachweisen ( $V=0.188$ ,  $p=0.353$ ).

## Literatur

- Arntzen, F. (1983). *Psychologie der Zeugenaussage. System der Glaubwürdigkeitsmerkmale*. München: Beck.
- Bullens, R. (1995). Der Grooming-Prozeß – oder das Planen des Mißbrauchs. In B. Marquardt-Mau (Hrsg.), *Schulische Prävention gegen sexuelle Kindesmißhandlung. Grundlagen, Rahmenbedingungen, Bausteine und Modelle* (S. 55-67). Weinheim: Juventa Verlag.
- Greuel, S., Offe, S., Fabian, A., Wetzels, P., Fabian, T., Offe, H. & Stadler, M. (1998). *Glaubhaftigkeit der Zeugenaussage*. Weinheim: Beltz Psychologie Verlags Union.
- Klaffer, K. & Undeutsch, U. (2000). *Täterstrategien in Zeugenaussagen zu sexuellem Mißbrauch*. Paper zum 1. Spanisch-deutschen Kongreß für Rechtspsychologie, Pamplona, Spanien, Juli 2000.
- Smith, D.W., Letourneau, E.J., Saunders, B.E., Kilpatrick, D.G., Resnick, H.S., Best, C.L. (2000). Delay in disclosure of childhood rape: results from a national survey. *Child Abuse & Neglect*, 24, 2, 273-287.
- Steller, M. & Köhnken, G. (1989). Criteria-based statement analysis. In C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 217-245). New York: Springer.

Trankell, A. (1971). *Der Realitätsgehalt von Zeugenaussagen. Methodik der Aussagepsychologie*. Göttingen: Vandenhoeck & Ruprecht.

Undeutsch, U. (1967). Beurteilung der Glaubhaftigkeit von Zeugenaussagen. In U. Undeutsch (Hrsg.), *Handbuch der Psychologie, Bd. 11: Forensische Psychologie*, S. 26-181. Göttingen: Hogrefe.



# Deliktmerkmale und Strafzumessung bei einem Vergewaltigungsdelikt: Vergleich einer Fragebogenstudie mit einer Conjoint-Analyse

*Michaela Brocke, Judith Bus, Heinz Holling & Wolfgang Bilsky*  
*Psychologisches Institut IV der Universität Münster*

Im vorliegenden Beitrag werden zwei Studien zum Einfluß von Merkmalen des Opfers sowie der Täter-Opfer-Beziehung auf die Strafzumessung bei einem Vergewaltigungsdelikt dargestellt. Der ersten Studie liegt ein 'between-subjects-Design' zugrunde, das zur Untersuchung von Fragestellungen wie der vorliegenden üblicherweise zum Einsatz kommt. In der zweiten Studie wurde die adaptive Conjoint-Analyse eingesetzt, die eine Untersuchung der Fragestellung mittels eines 'within-subjects-Designs' ermöglicht. Während bei der Untersuchung mit 'between-subjects-Design' der empirisch gesicherte Einfluß von zwei Deliktmerkmalen auf die Strafzumessung nicht nachgewiesen werden kann, bestätigen die Ergebnisse der Conjoint-Analyse die bereits vorliegenden Befunde. Durch das der Conjoint-Analyse zugrundeliegende 'within-subjects-Design' kann darüber hinaus gezeigt werden, daß jedem der drei Merkmale intraindividuell ein unterschiedlich hohes Gewicht bei der Strafzumessung zukommt und interindividuelle Unterschiede im Strafzumessungsverhalten bestehen.

## **1 Einleitung**

Zahlreiche Studien haben sich in der Vergangenheit mit der Verantwortungszuschreibung und der Strafzumessung nach Vergewaltigungen auseinandergesetzt (vgl. Krahe, 1985, 1991). Verschiedentlich wurde dabei die Wirkung potentiell relevanter Variablen auf die Einschätzung von Opfer und Täter analysiert. Berücksichtigt wurden in diesem Zusammenhang neben Täter- und Opfermerkmalen auch die Täter-Opfer-Beziehung sowie Beurteilermerkmale. Die Befunde erwiesen sich allerdings nur zum Teil als konsistent; dies gilt insbesondere für den Einfluß der Täter-Opfer-Beziehung auf die Strafzumessung und die Verantwortungszuschreibung (z.B. Bolt & Caswell, 1981; Bridges, 1993; Krulowitz, 1982; Smith, Keating, Hester & Mitchell, 1976; Szymanski, Devlin, Chrisler & Vyse, 1993). Unüberprüft blieb bisher, ob und in welche Richtung sich im Falle einer früheren Partnerschaft die Information, von welchem Partner die Trennung ausging, auf das Sanktionsverhalten auswirkt.

Zur Prüfung dieser und anderer offener Fragen wurde von Bus (2001) eine Fragebogenstudie mit juristischen Laien zum Einfluß von Täter-Opfer-Beziehung und Opfervariablen auf die Strafhärte bei einem Vergewaltigungsdelikt durchgeführt. In dieser Untersuchung hatten die Probanden die Aufgabe, aus mehreren Sanktionsalternativen diejenige auszuwählen, die aus ihrer Sicht einer im Rahmen eines 'between-subjects-Design' systematisch variierten Deliktbeschreibung (Vignette) angemessen erschien. In einer zweiten Studie wurde die Wirksamkeit der betreffenden Treatmentvariablen erneut, in diesem Fall jedoch im Rahmen eines 'within-subjects-Design' untersucht. Realisiert wurde diese zweite Untersuchung mittels einer Conjoint-Analyse (Green & Srinivasan, 1978). Ziel dieses explorativen Methodenvergleichs war es, Aufschluß über die Einsetzbarkeit conjoint-analytischer Verfahren im Rahmen des hier interessierenden Forschungskontextes zu gewinnen. Nachfolgend werden das Design und die Realisierung beider Studien erläutert und die Ergebnisse einander gegenübergestellt. Da der Einsatz von Conjoint-Analysen in diesem Forschungskontext unüblich ist, wird das methodische Vorgehen in der zweiten Studie entsprechend ausführlicher dargestellt.

## 2 Methode

Sowohl die Fragebogenstudie als auch die Conjoint-Analyse basieren auf einem 2×2×3-faktoriellen Design. Als *unabhängige Variablen* fungierten jeweils zwei Opfervariablen, 'Alkoholisierung' und 'Verhalten des Opfers' (beide zweigestuft), sowie die Variable 'Täter-Opfer-Beziehung' (dreigestuft). Durch Kombination der in Tabelle 1 zusammengefaßten Fallmerkmale ergaben sich insgesamt 12 verschiedene Fallkonstellationen, die den Probanden in Form von Vignetten vorgegeben wurden. Eine von ihnen ist nachfolgend in der in Studie 1 verwendeten Form zur Veranschaulichung wiedergegeben.

Frau T., eine 28-jährige Frau, hat auf einer Geburtstagsfeier ihrer Freundin zufällig ihren ehemaligen Lebensgefährten getroffen. Frau T. hat sich von diesem vor 4 Monaten getrennt. Da Frau T. sich ziemlich beschwipst fühlt, wird sie nach der Geburtstagsfeier von ihrem ehemaligen Lebensgefährten zu Fuß nach Hause gebracht. Vor ihrer Tür erbittet er Zutritt in ihre Wohnung, um sich ein Taxi zu rufen. In ihrem Wohnzimmer beginnt er, sie zu küssen. Als Frau T. versucht, dies zu verhindern, greift er ihren Arm, um sie wieder zu küssen. Frau T. schreit und versucht ihn von sich zu schubsen, aber er stößt sie auf die Couch und vergewaltigt sie.

Tabelle 1: Untersuchte Faktoren / Fallmerkmale und ihre jeweiligen Ausprägungen

Faktor / Merkmal	Ausprägungen
Verhalten des Opfers	sie bittet ihn in ihre Wohnung (provozierendes Verhalten)
	er erbittet sich unter einem Vorwand Zutritt zu ihrer Wohnung (kein provozierendes Verhalten)
Alkoholisierung	das Opfer ist beschwipst
	das Opfer hat keinen Alkohol getrunken
Täter-Opfer-Beziehung	Täter und Opfer sind flüchtig miteinander bekannt
	Täter und Opfer waren ein Paar, er hat sich von ihr getrennt
	Täter und Opfer waren ein Paar, sie hat sich von ihm getrennt

*Abhängige Variable* war die je individuelle Strafzumessung. Nach dem Lesen der jeweiligen Deliktbeschreibung bestand die Aufgabe des Probanden darin, in standardisierter Form auf die geschilderte Straftat zu reagieren.

Neben den genannten Variablen wurden in beiden Studien mehrere Beurteilermerkmale erfaßt, die im Rahmen der Fragebogenstudie (Studie 1) als *Kovariablen* zur Kontrolle der nicht auf die experimentellen Variablen zurückzuführenden Varianz zwischen den Probanden dienten. Hierbei handelte es sich um folgende Variablen:

- *allgemeine Strafhärtetendenz*, gemessen mit der Skala zur Erfassung der Strafhärteeinstellung von Bilsky, Mecklenburg und Wetzels (1993);
- *Akzeptanz von Vergewaltigungsmythen*, erfaßt mit einer Kurzform der Illinois Rape Myth Acceptance Scale (IRMA-SF) von Payne, Lonsway und Fitzgerald (1999; Übersetzung von Bus, 2001);
- *Opfer- und Täterempathie*, erfaßt mit einer gekürzten und modifizierten Form der Rape Empathy Scale (Deitz, Thiemann Blackwell, Daley & Bentley, 1982; vgl. Bus, 2001);
- *direkte oder indirekte Opfererfahrung* (ein Item).

Darüber hinaus wurde das Geschlecht der Probanden als weitere zu kontrollierende Variable berücksichtigt. Nachfolgend werden die beiden Studien im Hinblick auf die untersuchten Stichproben, das methodische Vorgehen sowie die Datenauswertung näher erläutert.

## 2.1 Fragebogenstudie (Studie 1)

### *Versuchspersonen*

An Studie 1 nahmen insgesamt 333 Studenten (200 Frauen und 133 Männer) im Alter von 19 bis 61 Jahren ( $x_{\text{med}}=24$ ) teil. Von den befragten Personen gaben 63% an, nie zuvor, 22% einmal und 15% mehrmals *selbst oder indirekt* über nahestehende Personen sexuelle Gewalt erfahren zu haben.

### *Methodisches Vorgehen*

Die Datenerhebung fand mittels eines schriftlich zu bearbeitenden dreiteiligen Fragebogens statt. Zunächst wurde jedem Probanden nach dem Zufallsprinzip eine der zwölf Vignetten (Fallkonstellationen) zur Beurteilung zugewiesen. Die vorzunehmende Beurteilung wurde in einer gleichzeitig vorgelegten Instruktion erläutert. Sie bestand in einer dichotomen Entscheidung über die Verhängung einer Freiheitsstrafe ("keine Freiheitsstrafe" versus "Freiheitsstrafe") und, im Falle einer Freiheitsstrafe, der Bestimmung eines als angemessen erachteten Strafmaßes anhand einer 16-stufigen Skala ('ein halbes Jahr', 'ein Jahr', ..., jeweils im Jahresabstand zunehmend). Diese Strafhärteskalierung war in Anlehnung an eine entsprechende Aktenanalyse von Gregor (1987) gewählt worden (vgl. Bus, 2001). Im Anschluß an die Deliktbeurteilung wurden im zweiten Teil der Befragung die zuvor genannten Beurteilermerkmale (Kovariablen), im dritten abschließend einige demographische Angaben zur Person erfaßt.

### *Datenauswertung*

Die Prüfung der Treatmenteffekte erfolgte in einem ersten Schritt durch eine *univariate Varianzanalyse* mit den Fallmerkmalen als unabhängige Variablen und der zugemessenen Strafe als abhängige Variable. In einem zweiten Schritt wurde darauf aufbauend analysiert, ob durch zusätzliche Berücksichtigung der Beurteilermerkmale im Rahmen einer *Kovarianzanalyse* eine bessere Absicherung der Treatmenteffekte möglich ist.

## 2.2 Conjoint-Analyse (Studie 2)

### *Versuchspersonen*

An Studie 2 nahmen 75 Studenten (62 Frauen und 13 Männer) im Alter von 19 bis 44 Jahren ( $x_{\text{med}}=20$ ) teil. Der Anteil derjenigen Personen, die sexuelle Gewalt zuvor direkt oder indirekt erfahren hatten, ist demjenigen in Studie 1 vergleichbar: 69% hatten nie, 18% einmal und 12 % mehrmals entsprechende Erfahrungen gemacht.

### *Methodisches Vorgehen*

Die Strafhärtemessung und die Erfassung der Beurteilermerkmale erfolgte in computergestützter Form. Da auf letztere an dieser Stelle nicht weiter eingegangen wird, beschränkt sich die folgende Darstellung auf zentrale Aspekte der mittels Conjoint-Analyse durchgeführten Strafhärtemessung.

Den Probanden wurde zunächst die Untersuchungsthematik sowie die Bearbeitung der im Rahmen der Conjoint-Analyse vorgegebenen Aufgaben (Deliktbeschreibungen) mündlich erläutert. Da in der Conjoint-Analyse im Vergleich zur Studie mit between-subjects-Design allein die Merkmalsausprägungen für die Delikte dargeboten werden, wurde den Probanden in der mündlichen Instruktion als Rahmen zur Einbettung der dargebotenen

Merkmalsausprägungen die Rahmenhandlung der ersten Studie genannt. Die Vorgabe der jeweiligen Fallinformationen und die Erfassung der auf sie bezogenen Strafe erfolgte anschließend computerunterstützt unter Verwendung des Programms *Alasca* von Holling, Jütting und Großmann (2000). Hierbei hatten die Probanden nach einer Rangordnungsaufgabe, die ausschließlich dem Vertrautwerden mit den untersuchten Merkmalsausprägungen diente, 25 Paarvergleichsaufgaben sowie acht *Objektsbeurteilungsaufgaben* zu bearbeiten. Vor der Bearbeitung der Aufgaben wurde den Probanden während der Datenerhebung ein weiteres Mal die zuvor in mündlicher Form gegebene Instruktion erneut in schriftlicher Form dargeboten.

Bei jedem *Paarvergleich* wurden den Probanden zwei Varianten des Vergewaltigungsdeliktes vorgegeben, die sich im Hinblick auf die in Tabelle 1 zusammengefaßten Merkmale unterschieden. Die Aufgabe der Probanden bestand darin, auf einer 7-stufigen Ratingskala anzugeben, welche der beiden Varianten härter zu bestrafen ist (s. Abbildung 1). Um aus den abgegebenen Urteilen der Personen auf den Einfluß der präsentierten Merkmale auf die Strafentscheidung schließen zu können, wurden die Probanden darauf hingewiesen, daß sich die präsentierten Delikte allein in den beschriebenen Merkmalsausprägungen unterscheiden.

Abbildung 1: Beispiel einer Paarvergleichsaufgabe in Alasca

Die Delikt-Paare waren zunächst anhand von zwei (10 Aufgaben), anschließend anhand von drei Merkmalen (15 Aufgaben) zu vergleichen. Die Vorgabe des nächsten Delikt-Paares erfolgte in Abhängigkeit von den zuvor abgegebenen Urteilen des Probanden. Bei dieser Form der *adaptiven Conjoint-Analyse* ist sichergestellt, daß der jeweils nächste Vergleich die größtmögliche Information zur Schätzung des Einflusses der Merkmalsausprägungen auf die Strafzumessung liefert. Konkret bedeutet dies, daß nach jeder bearbeiteten Aufgabe eine neue Schätzung der Ergebnisparameter vorgenommen wird.

Im letzten Teil der Conjoint-Analyse wurden den Probanden in adaptiver Sequenz acht *Objektbeurteilungsaufgaben* vorgelegt (s. Abbildung 2). Dabei wurde jeweils ein Delikt mittels der untersuchten Merkmalsausprägungen beschrieben. Die Probanden gaben auf einer 9-stufigen Skala an, wie hoch der Täter bestraft werden sollte. Die Pole der Skala wurden in Anlehnung an Studie 1 in den mündlichen Instruktionen mit einer unteren Grenze von einem halben Jahr und einer oberen Grenze von 15 Jahren Freiheitsstrafe verankert.

Abbildung 2: Beispiel einer Objektbeurteilungsaufgabe in Alasca

### *Datenauswertung*

Die Conjoint-Analyse ermittelt vorrangig zwei Arten von Ergebnisparametern, die den Einfluß der variierten unabhängigen Variablen auf die abhängige Variable beschreiben. Hierbei handelt es sich um sogenannte *'Teilnutzenwerte'* der untersuchten Merkmalsausprägungen (Faktorstufen) und um *'Wichtigkeiten'* der Merkmale (Faktoren). Die Teilnutzenwerte werden nachfolgend als Strafwerte interpretiert, da sie indizieren, ob und in welchem Ausmaß der jeweiligen Merkmalsausprägung eine strafverschärfende Wirkung beigemessen wird. Die Wichtigkeiten geben an, welchen Einfluß dem jeweiligen Merkmal insgesamt bei einer Strafzumessung zukommt. Beide werden wie folgt bestimmt.

Aus den Urteilen eines Probanden, d.h. den insgesamt 33 Antworten aus den Paarvergleichen und Objektbeurteilungen, werden zunächst mittels linearer Regression Beta-Gewichte der Merkmalsausprägungen der beurteilten Objekte (hier: der Deliktvarianten) ermittelt; diese entsprechen den *Teilnutzenwerten* (hier: den *Straf- bzw. Einflußwerten*), die den Merkmalsausprägungen zukommen. Das Verfahren erfordert, daß sich der *'Gesamtnutzenwert'* eines Objektes additiv aus den Teilnutzenwerten der einzelnen Merkmalsausprägungen zusammensetzt (vgl. Jungermann, Pfister & Fischer, 1998; s. aber Abelson & Levi, 1985) und sich die Merkmalsausprägungen nicht gegenseitig beeinflussen, d.h. unabhängig voneinander sind. Die *Wichtigkeiten* der Merkmale sind dann für jedes Merkmal durch die Spanne der Teilnutzenwerte seiner Ausprägungen zu ermitteln, d.h. der niedrigste Teilnutzenwert der Merkmalsausprägung wird vom höchsten subtrahiert. Teilnutzenwerte der Merkmalsausprägungen und Wichtigkeiten werden bei der Conjoint-Analyse auf individueller Ebene ermittelt. Zur anschließenden Aggregation der individuellen Beta-Gewichte wird eine Normierung der Teilnutzenwerte vorgenommen, die eine interindividuelle Vergleichbarkeit herstellt. Dabei wird der Merkmalsausprägung mit dem höchsten Beta-Gewicht ein Wert von 100 zugewiesen, die Ausprägungen jedes Merkmals mit dem geringsten Beta-Gewicht wird auf Null gesetzt (s. Backhaus, Erichson, Plinkel & Weiber, 2000). Alle übrigen Ausprägungen werden entsprechend normiert.

In der vorliegenden Untersuchung bezog sich die Auswertung der Daten somit zunächst auf die Ermittlung der *Strafwerte (Teilnutzenwerte)* der Merkmalsausprägungen und *Wichtigkeiten* der Merkmale (Alkoholisierung, Verhalten des Opfers und Täter-Opfer-Beziehung) auf Stichprobenebene. Beide wurden anhand der Antworten aus den

Paarvergleichen sowie Objektbeurteilungen ermittelt. Zuvor wurde dazu in einem ersten Schritt durch ein Regressionsmodell mit Interaktionstermen geprüft, ob die Annahme der Unabhängigkeit der Ausprägungen zutrifft.

Durch die Aggregation der Strafwerte können allerdings Unterschiede in den Relationen der Strafwerte zwischen Personen bzw. Personengruppen unentdeckt bleiben (s. Schubert, 1991). Mißt beispielsweise eine Probandengruppe der Merkmalsausprägung 'das Opfer ist alkoholisiert' einen höheren Strafwert zu als der Ausprägung 'das Opfer hat keinen Alkohol getrunken' und liegt bei einer zweiten Gruppe von Probanden kein Unterschied zwischen den Strafwerten dieser Merkmalsausprägungen vor, so würde das Strafwerte-Profil der zweiten Gruppe durch die Aggregation der Strafwerte 'verloren gehen'. Daher wurde in einem zweiten Schritt eine *hierarchische Clusteranalyse* (Ward-Verfahren; quadrierte euklidische Distanz) auf Basis der individuellen, normierten Strafwerte (Teilnutzenwerte) durchgeführt, um Personengruppen zu identifizieren, die durch unterschiedliche Gewichtungen der Merkmalsausprägungen der beurteilten Deliktvarianten gekennzeichnet sind (vgl. Backhaus et al., 2000).

### 3 Ergebnisse

#### 3.1 Ergebnisse der Fragebogenstudie (Studie 1)

Die im Rahmen der konventionellen Fragebogenstudie erhobenen Daten wurden in einem ersten Schritt einer dreifaktoriellen Varianzanalyse mit dem Strafmaß als abhängige Variable und den beiden Opfervariablen sowie der Täter-Opfer-Beziehung als unabhängigen Variablen unterzogen. Keines der drei Deliktmerkmale leistete einen signifikanten Beitrag zur Vorhersage des Strafmaßes (vgl. Tabelle 2). In einem erweiterten Modell mit Zwei- und Dreifach-Interaktionen der unabhängigen Variablen ließen sich ebenfalls keine signifikanten Haupt- und Interaktionseffekte nachweisen.

In einem zweiten Analyseschritt wurden die Beurteilermerkmale zusätzlich als Kovariablen berücksichtigt. Die Ergebnisse der entsprechenden mehrfaktoriellen Kovarianzanalyse sind in Tabelle 3 wiedergegeben; wie ersichtlich, konnte in diesem Fall ein signifikanter Einfluß der Täter-Opfer-Beziehung nachgewiesen werden ( $R^2 = 0.18$ , adjustiertes  $R^2 = 0.15$ ).

Tabelle 2: Ergebnisse der univariaten Varianzanalyse mit den Deliktmerkmalen

Quelle	Quadratsumme	df	Mittel der Quadrate	F	p
Korrigiertes Modell	85.99	4	21.49	1.28	0.31
Konstanter Term	11310.84	1	11310.84	635.60	0.00
Alkoholisierung	4.78	1	4.78	0.27	0.61
Täter-Opfer-Beziehung	80.54	2	40.27	2.26	0.11
Verhalten des Opfers	0.39	1	0.39	0.02	0.88
Fehler	5712.35	321	17.79		
Gesamt	17071.00	326			

Da die Täter-Opfer-Beziehung dreifach gestuft vorliegt, wurden als a-posteriori Vergleiche zwischen den Merkmalsausprägungen Helmert-Kontraste berechnet, um zu prüfen, ob sich die einzelnen Merkmalsausprägungen signifikant voneinander unterscheiden. Bei Helmert-

Kontrasten wird die Auswirkung einer jeden Merkmalsausprägung mit der mittleren Auswirkung aller folgenden Ausprägungen verglichen. Hier zeigte sich ein signifikanter ( $p=0.02$ ) Unterschied in der Strafzumessung zwischen der Alternative 'flüchtig bekannt' und den beiden anderen Merkmalsausprägungen, die auf eine frühere Partnerschaft von Täter und Opfer hinweisen. Die Information, von welcher Seite die Beendigung der Partnerschaft ausging, führte nach den Ergebnissen des Helmert-Kontrastes zu keinem signifikanten Unterschied in der Strafzumessung.

Tabelle 3: Ergebnisse der Kovarianzanalyse mit den Delikt- und Beurteilermerkmalen

Quelle	Quadratsumme	df	Mittel der Quadrate	F	p
Korrigiertes Modell	1047.26	11	95.21	6.29	0.00
Konstanter Term	36.52	1	36.52	2.41	0.12
Alkoholisierung	24.83	1	24.83	1.64	0.20
Täter-Opfer-Beziehung	95.68	2	47.84	3.16	0.04
Verhalten des Opfers	1.29	1	1.29	0.09	0.77
Geschlecht	44.39	1	44.39	2.93	0.09
Opfererfahrung	175.43	2	87.72	5.79	0.00
Opferempathie	52.56	1	52.56	3.47	0.06
Täterempathie	2.02	1	2.02	0.13	0.72
Mythenakzeptanz	328.23	1	328.23	21.69	0.00
Strafhärteeinstellung	194.91	1	194.91	12.88	0.00
Fehler	4751.07	314	15.13		
Gesamt	17071.00	326			

Insgesamt zeigte sich in der Fragebogenstudie somit allein ein signifikanter Unterschied in der Strafzumessung zwischen zwei Ausprägungen der Variable 'Täter-Opfer-Beziehung'. Die Variablen 'Alkoholisierung' und 'Verhalten des Opfers' wiesen entgegen bisherigen Forschungsbefunden (z.B. Best & Demmin, 1982; Schuller & Wall, 1998; Schult & Schneider, 1991; Workman & Freiburg, 1999) keinen Einfluß auf das von den Probanden präferierte Strafmaß auf.

### 3.2 Ergebnisse der Conjoint-Analyse (Studie 2)

Im ersten Auswertungsschritt der Conjoint-Studie wurde die Unabhängigkeit der Deliktmerkmale über eine lineare Regression der Urteile der Probanden geprüft. Während in diesem Gesamtmodell jede der unabhängigen Variablen einen signifikanten Beitrag zur Vorhersage der Antwort der Probanden leistete, ergab sich kein signifikanter Beitrag der Interaktionen zwischen den unabhängigen Variablen. Die Unabhängigkeit der Merkmale ist insofern gegeben (s. Tabelle 4).

Tabelle 4: Ergebnisse der linearen Regression (effektkodierte Deliktmerkmale)

unabhängige Variablen	B	SF	Beta	t	p
provokatives Verhalten (Verhalten des Opfers 1)	-0.66	0.02	-0.50	-28.72	0.00
Alkoholisierung (Alkoholisierung 1)	0.10	0.02	0.08	4.41	0.00
flüchtig bekannt (Täter-Opfer-Beziehung 1)	0.26	0.04	0.15	7.32	0.00
Paar, Täter hat sich getrennt (Täter-Opfer-Beziehung 2)	-0.22	0.04	-0.13	-6.04	0.00
Alkoholisierung 1 x Täter-Opfer-Beziehung 1	0.02	0.02	0.02	0.83	0.41
Alkoholisierung 1 x Täter-Opfer-Beziehung 2	-0.02	0.02	-0.01	-0.70	0.48
Verhalten des Opfers 1 x Alkoholisierung 1	0.01	0.01	0.02	0.94	0.35
Verhalten des Opfers 1 x Täter-Opfer-Beziehung 1	-0.01	0.02	-0.01	-0.33	0.74
Verhalten des Opfers 1 x Täter-Opfer-Beziehung 2	-0.03	0.02	-0.03	-1.49	0.14

Anmerkung: B=nicht standardisierte Koeffizienten, Beta=standardisierte Koeffizienten, SF=Standardfehler

Nach Normierung der durch Regressionen ermittelten individuellen Strafwerte ergaben sich die in Abbildung 3 wiedergegebenen *mittleren Strafwerte* für die untersuchten Merkmalsausprägungen. Die graphische Darstellung der mittleren Strafwerte zeigt, daß die einzelnen Deliktmerkmale einen unterschiedlichen Einfluß auf die Strafzumessung ausüben.

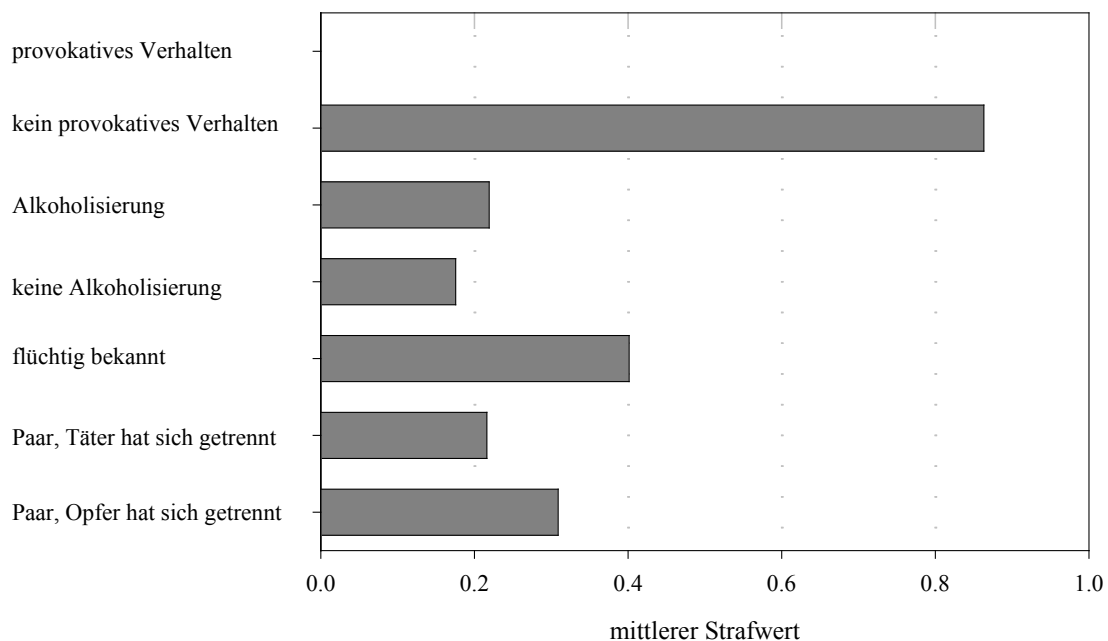


Abbildung 3: Einflussgewichte der Deliktmerkmalsausprägungen bei der Strafzumessung

Auf aggregierter Ebene kommt dabei dem Verhalten des Opfers der höchste Einfluß auf die Strafzumessung zu: Liegt 'kein provokatives Verhalten' des Opfers vor, so fällt das Strafmaß für den Täter höher aus. Der zweithöchste mittlere Strafwert wurde für die Merkmalsausprägung 'Täter und Opfer sind flüchtig bekannt' ermittelt; demzufolge wird in entsprechenden Fällen ein höheres Strafmaß gewählt als bei Delikten, bei denen von einem hohen Maß an Vertrautheit zwischen Täter und Opfer auszugehen ist. Der Unterschied



zwischen den entsprechenden Strafwerten des Merkmals 'Alkoholisierung' schließlich fällt relativ gering aus; hier weisen die Ergebnisse darauf hin, daß auf aggregierter Ebene bei Vorliegen einer Alkoholisierung des Opfers eine leicht höhere Strafe gewählt wird. In Abbildung 4 sind die *Wichtigkeiten* der drei Faktoren (Merkmale) dargestellt, die sich durch die Spanne ihrer Ausprägungen ergeben; bei dieser Darstellung wurde dem Merkmal mit dem höchsten Einflußgewicht der Wert 100 zugeordnet. Hier zeigt sich auf andere Weise, daß dem Merkmal Alkoholisierung die geringste, dem Verhalten des Opfers demgegenüber die größte Bedeutung bei der Strafzumessung zukommt.

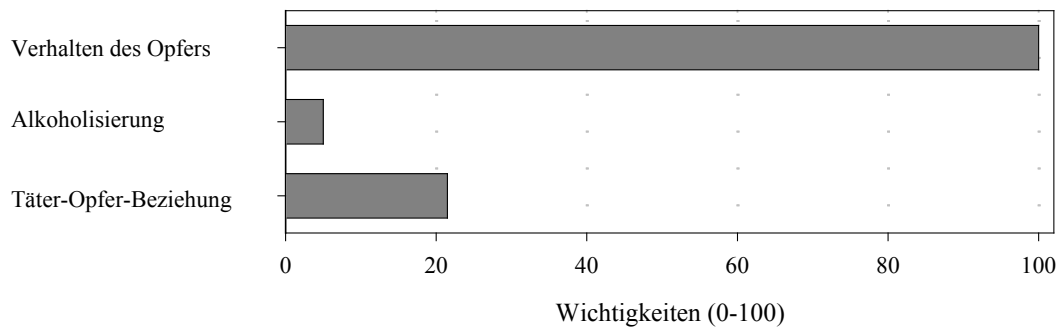


Abbildung 4: Wichtigkeiten der Deliktmerkmale bei der Strafzumessung

Die mit den individuellen normierten Strafwerten durchgeführte hierarchische Clusteranalyse legte eine Vier-Cluster-Lösung nahe. In Abbildung 5 sind für jedes der vier Cluster die mittleren Strafwerte der Merkmalsausprägungen als Profile darstellt. Die Profile charakterisieren Probandengruppen, die sich in der Gewichtung der Merkmalsausprägungen unterscheiden.

Allen Probandengruppen ist gemeinsam, daß sie im Hinblick auf das *Verhalten des Opfers* der Merkmalsausprägung 'kein provokatives Verhalten' einen höheren Strafwert, d.h. strafverschärfende Bedeutung beimessen.

Bezüglich des Merkmals *Alkoholisierung* lassen sich die vier Cluster in zwei Gruppen aufteilen. Die Probanden des Clusters 1 (n=15) und Clusters 3 (n=12) messen der Merkmalsausprägung 'das Opfer ist alkoholisiert' einen höheren Strafwert zu, die Cluster 2 (n=29) und Cluster 4 (n=19) demgegenüber der Merkmalsausprägung 'das Opfer hat keinen Alkohol getrunken'.

Beide Gruppen von Clustern unterscheiden sich intern schließlich in der Gewichtung der Merkmalsausprägungen der *Täter-Opfer-Beziehung*: Hier weist jedes der vier ermittelten Cluster eine unterschiedliche Konstellation der Strafwerte auf (s. Abbildung 5).

Insgesamt zeigen die Ergebnisse der Clusteranalyse, daß die einzelnen Deliktmerkmale bei der Zumessung einer Strafe von verschiedenen Personen unterschiedlich gewichtet werden. Die Analyse von interindividuellen Unterschieden bzw. die Aufdeckung von „typischen“ Profilen von Subgruppen wurde durch die mit der Conjoint-Analyse ermittelten *individuellen* Strafwerte ermöglicht.

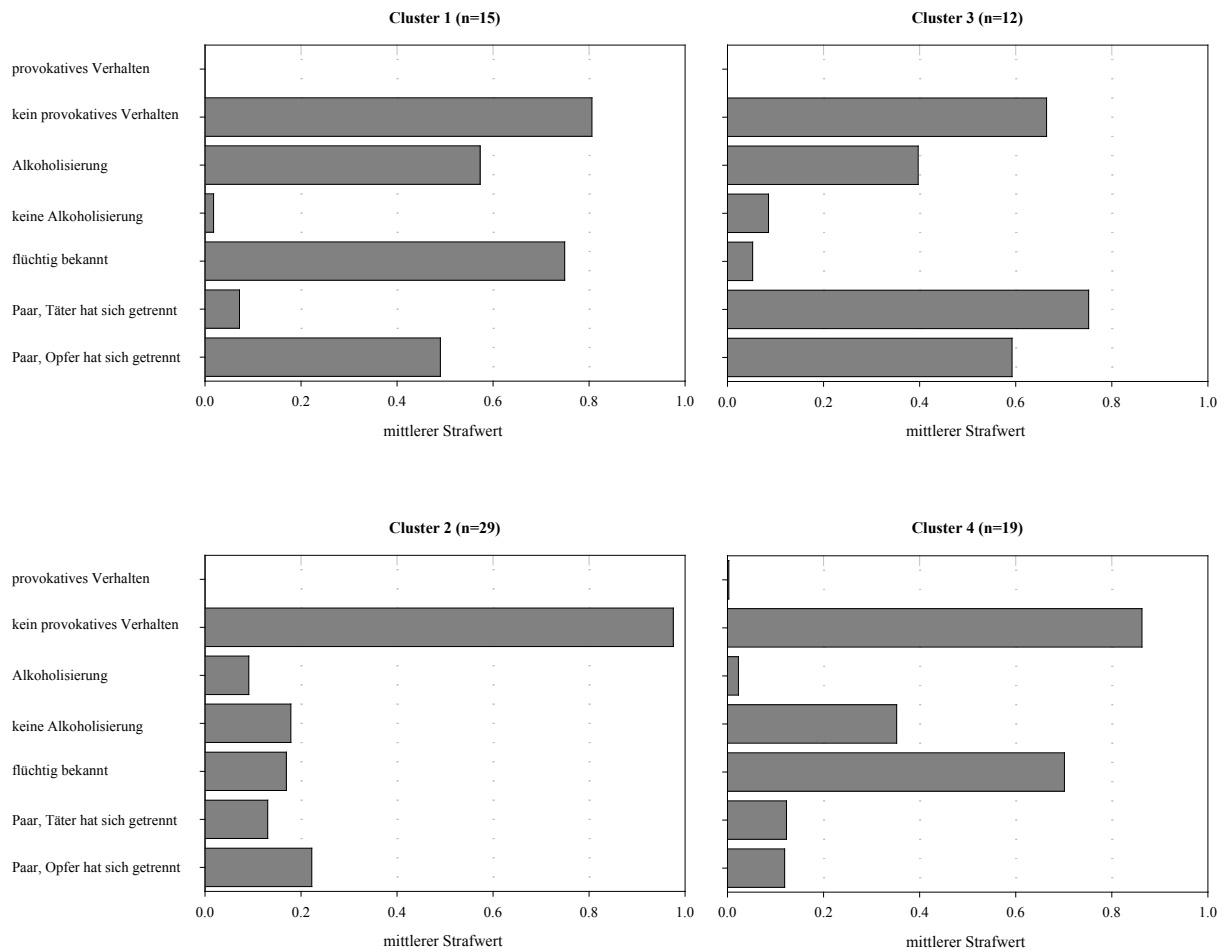


Abbildung 5: Profile der mittels Clusteranalyse identifizierten Probandengruppen

#### 4 Zusammenfassende Bewertung

Unser Methodenvergleich hat gezeigt, daß im Rahmen einer konventionellen, auf einem between-subjects-Design basierenden Fragebogenstudie ein unmittelbarer Nachweis des erwarteten Einflusses ausgewählter Fallvariablen auf die Sanktionierung eines Vergewaltigungsdelikts nicht möglich war. Erst nach statistischer Kontrolle individueller Unterschiede hinsichtlich mehrerer Beurteilermerkmale konnte ein signifikanter Einfluß eines der drei untersuchten Deliktmerkmale - der 'Täter-Opfer-Beziehung' - auf die für angemessen gehaltene Strafhöhe nachgewiesen werden.

Demgegenüber konnte mit der Conjoint-Analyse ermittelt werden, welche Gewichte den drei untersuchten Deliktmerkmalen bei der Strafzumessung zukommen. Auf aggregierter und auf individueller Ebene war dabei dem Verhalten des Opfers der mit Abstand höchste Einfluß auf die Strafzumessung beizumessen. Insbesondere für die Gewichtung der Merkmale 'Täter-Opfer-Beziehung' und 'Alkoholisierung des Opfers' ließen sich zudem mittels Clustering der conjointanalytischen Daten Subgruppen identifizieren, die durch unterschiedliche Gewichtungen der Merkmalsausprägungen charakterisiert sind.

Die Ergebnisse dieser Studie bestätigen die in einer weiteren Untersuchung zur Strafzumessung gemachten positiven Erfahrungen mit der Verwendung der adaptiven Conjoint-Analyse. Dort hatte Göldenitz (2001), unter Bezugnahme auf eine entsprechende Studie von Reichert (1999), ebenfalls mittels der Conjoint-Analyse den Einfluß von Alter des

Täters, Geständigkeit, Schwere der Tat und Vorstrafen auf die Strafzumessung bei einem Diebstahls- und einem Körperverletzungsdelikt bestimmen können.

Der Einsatz von Conjoint-Analysen innerhalb der Strafzumessungsforschung ist sicherlich ungewöhnlich. Dies spricht jedoch nicht gegen ihre grundsätzliche Verwendbarkeit in diesem Forschungsfeld sondern dürfte eher bereichsspezifische Arbeitsgewohnheiten charakterisieren. Tatsächlich gestattet es die adaptive Form der Conjoint-Analyse, die innerhalb der experimentellen Psychologie vielfach bewährte, jedoch zeitlich aufwendige Methode des Paarvergleichs aufgrund iterativer Schätzalgorithmen effektiv und ökonomisch für die Untersuchung von Strafzumessungsentscheidungen im Bereich einer grundlagenorientierten Forschung zu nutzen. Eine wichtige Voraussetzung ist dabei, den im Kontext der Conjoint-Analyse üblichen Begriff des 'Nutzen' in angemessener Weise auf die jeder Strafzumessung vorausgehende Entscheidungsfindung zu übertragen.

## Literatur

- Abelson, R.P. & Levi, A. (1985). Decision making and decision theory. In G. Lindzey & E. Aronson (Eds.), *Handbook of Social Psychology* (3<sup>rd</sup> ed., pp. 251-309). New York: Random House.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2000). *Multivariate Analysemethoden - Eine anwendungsorientierte Einführung*. Berlin: Springer.
- Best, J.B. & Demmin, H.S. (1982). Victim's provocativeness and victim's attractiveness as determinants of blame in rape. *Psychological Reports*, 51, 255-258.
- Bilsky, W., Mecklenburg, E. & Wetzels, P. (1993). *Persönliches Sicherheitsgefühl, Angst vor Kriminalität und Gewalt, Opfererfahrung älterer Menschen. Skalenanalyse und Skalenkonstruktion zur KFN-Opferbefragung 1992* (19-23). Kriminologisches Forschungsinstitut Niedersachsen, Hannover.
- Bolt, M. & Caswell, J. (1981). Attribution of responsibility to a rape victim. *The Journal of Social Psychology*, 114, 137-138.
- Bridges, J.S. (1993). Perceptions of date and stranger rape: A difference in sex role expectations and rape supportive beliefs. *Sex Roles*, 24, 291-307.
- Bus, J. (2001). *Einfluß der Täter-Opfer-Beziehung und Einfluß von Opfervariablen auf die Strafzumessung bei einem Vergewaltigungsdelikt*. Unveröffentlichte Diplomarbeit, Westfälische Wilhelms-Universität Münster.
- Deitz, S.R., Thiemann Blackwell, K., Daley, P.C. & Bentley, B.J. (1982). Measurement of empathy toward rape victims and rapists. *Journal of Personality and Social Psychology*, 43, 372-384.
- Göldenitz, C. (2001). *Analyse von Strafzumessungsurteilen mittels Conjoint-Analyse*. Unveröffentlichte Diplomarbeit, Westfälische Wilhelms-Universität Münster.
- Green, P.E. & Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research*, 5, 103-123.
- Gregor, R. (1987). Strafzumessung bei Vergewaltigung. *Monatsschrift für Kriminologie und Strafrechtsreform*, 70, 261-277.
- Holling, H., Jütting, A. & Großmann, H. (2000). *ALASCA - Computergestützte Entscheidungs- und Nutzenanalyse*. Göttingen: Hogrefe.
- Jungermann, H., Pfister, H.-R. & Fischer, K. (1998). *Die Psychologie der Entscheidung: Eine Einführung*. Heidelberg: Spektrum.
- Krahé, B. (1985). Die Zuschreibung von Verantwortlichkeit nach Vergewaltigungen: Opfer und Täter im Dickicht der attributionstheoretischen Forschung. *Psychologische Rundschau*, 36, 67-82.

- Krahé, B. (1991). Social psychological issues in the study of rape. *European Review of Social Psychology*, 2, 279-309
- Krulewitz, J.E. (1982). Reactions to rape victims: Effects of rape circumstances, victim's emotional response, and sex of helper. *Journal of Counselling Psychology*, 29, 645-654.
- Payne, D.L., Lonsway, K.A. & Fitzgerald, L.F. (1999). Rape myth acceptance: Exploration of its measurement using the Illinois Rape Myth Acceptance Scale. *Journal of Research in Personality*, 33, 27-68.
- Reichert, A. (1999). *Zusammenhang zwischen Ursachenzuschreibung für Kriminalität und Strafhärte: Eine empirische Untersuchung an einer Stichprobe juristischer Laien*. Unveröffentlichte Diplomarbeit, Westfälische Wilhelms-Universität Münster.
- Schubert, B. (1991). *Entwicklung von Konzepten für Produktmotivationen mittels Conjointanalyse*. Stuttgart: Poeschel.
- Schuller, R.A. & Wall, A.M. (1998). The effects of defendant and complainant intoxication on mock juror's judgement of sexual assault. *Psychology of Women Quarterly*, 22, 555-573.
- Schult, D.G. & Schneider, L.J. (1991). The role of sexual provocativeness, rape history, and observer gender in perceptions of blame in sexual assault. *Journal of Interpersonal Violence*, 6, 94-101.
- Smith, R.E., Keating, J.P., Hester, R.K. & Mitchell, H.E. (1976). Role and justice considerations in the attribution of responsibility to a rape victim. *Journal of research in personality*, 10, 346-357.
- Stormo, K.J. (1998). An investigation of factors affecting perceptions of sexual interest and attributions of blame in acquaintance rape. *Dissertation Abstracts International*, 58, 7 B.
- Szymanski, L.A., Devlin, A.S., Chrisler, J.C. & Vyse, S.A. (1993). Gender role and attitudes toward rape in male and female college students. *Sex Roles*, 29, 37-57.
- Workman, J.E. & Freeburg, E.W. (1999). An examination of date rape, victim dress, and perceiver variables within the context of attribution theory. *Sex Roles*, 41, 261-277.

# Psychologische Sachverständigengutachten im Entlassungsverfahren von geistig abnormen zurechnungsfähigen Rechtsbrechern in Österreich

## Eine qualitative und quantitative Analyse

*Christine Brugger*

*Psychologisches Institut der Universität Wien*

Der vorliegende Beitrag ist die Zusammenfassung einer Diplomarbeit über Sachverständigengutachten im Entlassungsverfahren aus der Maßnahme nach § 21 (2) StGB. In dieser Arbeit werden die Rolle und Aufgabe des psychologischen und zum Vergleich auch des psychiatrischen Sachverständigen im Entlassungsverfahren genauer beschrieben und dargestellt. Die Untersuchung beschränkt sich auf das Jahr 1998 und den Gerichtssprengel Wien mit den beiden Anstalten für Untergebrachte nach § 21 (2) StGB: die Sonderanstalt Mittersteig (inkl. Außenstelle Floridsdorf) in Wien und die Justizanstalt Stein/Donau. Insgesamt wurden 44 Entlassungsverfahren mit 63 psychologischen und psychiatrischen Gutachten analysiert. 14 Verfahren führten zu einer bedingten Entlassung, die restlichen 30 Verfahren hatten eine weitere Anhaltung nach § 21(2) StGB zur Folge.

Im Vergleich zum Deutschen Maßregelvollzug unterscheidet das österreichische Strafgesetzbuch zwischen zurechnungsfähigen und zurechnungsunfähigen geistig abnormen Rechtsbrechern: Eine Verurteilung nach § 21(1) StGB erfolgt dann, wenn jemand eine Tat, die mit einer ein Jahr übersteigenden Freiheitsstrafe bedroht ist, begeht und nur deshalb nicht bestraft werden kann, weil er sie unter dem Einfluß eines die Zurechnungsfähigkeit ausschließenden Zustandes begangen hat, der auf einer geistigen oder seelischen Abartigkeit von höherem Grad beruht. In einem solchen Fall hat ihn das Gericht in eine Anstalt für geistig abnorme Rechtsbrecher einzuweisen, wenn nach seiner Person, nach seinem Zustand und nach der Art der Tat zu befürchten ist, daß er sonst unter dem Einfluß seiner geistigen oder seelischen Abartigkeit eine mit Strafe bedrohte Handlung mit schweren Folgen begehen werde. Im Unterschied zu § 21(1) StGB werden nach § 21(2) StGB auch jene geistig abnormen Rechtsbrecher eingewiesen, die ohne zurechnungsunfähig zu sein, unter dem Einfluß einer geistigen oder seelischen Abartigkeit von höherem Grad eine Tat begangen haben, die mit einer ein Jahr übersteigenden Freiheitsstrafe bedroht ist. In einem solchen Fall ist die Unterbringung zugleich mit dem Ausspruch der Strafe anzuordnen. Der Unterschied zwischen geistig abnormen zurechnungsunfähigen und geistig abnormen zurechnungsfähigen Rechtsbrechern liegt im österreichischen Strafgesetzbuch darin, daß nach § 21(2) (geistig abnormer zurechnungsfähiger Rechtsbrecher) zusätzlich zur Unterbringung in eine Anstalt für geistig abnorme Rechtsbrecher auch eine Freiheitsstrafe ausgesprochen wird.

Grundsätzlich erfolgt die Einweisung in eine Anstalt für geistig abnorme Rechtsbrecher erst nach der Untersuchung des Betroffenen durch mindestens einen Sachverständigen aus dem Gebiet der Psychiatrie (vgl. §§ 429, 439 StPO). Die Dauer der Unterbringung in eine Anstalt für geistig abnorme Rechtsbrecher wird auf unbestimmte Zeit angeordnet und wird solange vollzogen, wie es ihr Zweck erfordert. Der Zustand des Untergebrachten soll soweit gebessert werden, daß die Begehung einer strafbaren Handlung nicht mehr zu erwarten ist, und weiters soll ihm zu einer rechtschaffenen und den Erfordernissen des Gemeinschaftslebens angepaßten Lebenseinstellung verholfen werden – vgl. § 164 StVG. Über die Aufhebung der Unterbringung entscheidet das Gericht mindestens einmal jährlich. Im Unterschied zur Einweisung, die gesetzlich die Beiziehung zumindest eines Sachverständigen vorschreibt, erfolgt die bedingte Entlassung laut § 17(1) StVG unter Einholung von Äußerungen des

Anstaltsleiters, des öffentlichen Anklägers sowie des Verurteilten. Nach § 17(2) StVG erfolgt die Beiziehung eines ärztlichen oder psychologischen Sachverständigen nur, wenn der Sachverhalt im Hinblick auf den Gesundheitszustand oder die Wesensart des Verurteilten nicht genügend geklärt erscheint. Es gibt also grundsätzlich keine gesetzliche Verpflichtung, einen externen Gutachter im Entlassungsverfahren aus der Maßnahme beizuziehen. Zur Beschreibung des Täters, der nach § 21(2) StGB in Österreich verurteilt wird, werden in den folgenden Tabellen neben der Hauptdiagnose im Einweisungsgutachten auch das Delikt und die Vorstrafen der untersuchten Stichprobe dargestellt:

Tabelle 1

Hauptdiagnose bei der Einweisung (n = 42 Untergebrachte)	abs.	%
Debilität	2	4,8 %
Persönlichkeitsstörung	22	52,4 %
Perversionen	1	2,4 %
Neurosen	3	7,1 %
Schizophrenie	1	2,4 %
Pädophilie	9	21,4 %
Pyromanie	3	7,1 %
Kleptomanie	1	2,4 %
gesamt	42	100 %

Tabelle 2

Delikt (n = 44 Untergebrachte)	abs.	%
Eigentum	1	2,3 %
Gewalt	13	29,5 %
Sexual	22	50 %
Brandstiftung	8	18,2 %
gesamt	44	100 %

Die Hälfte der Untergebrachten wurde aufgrund eines Sexualdeliktes verurteilt: 18 Untergebrachte aufgrund sexuellen Mißbrauchs an Minderjährigen und vier aufgrund von Vergewaltigung bzw. versuchter Vergewaltigung oder sexueller Nötigung. 13 Untergebrachte der Stichprobe sind aufgrund eines Gewaltdeliktes, v.a. wegen Körperverletzung aber auch wegen Mordes oder versuchten Mordes und Totschlags, nach § 21(2) StGB eingewiesen worden. Nur ein Untergebrachter wurde wegen Einbruchsdiebstahls in die Maßnahme eingewiesen.

Tabelle 3

Vorstrafen (n = 42 Untergebrachte)	abs.	%
keine	5	11,9 %
1	3	7,1 %
2 bis 5	16	38,1 %
6 bis 10	14	33,3 %
11 bis 20	2	4,8 %
über 20	2	4,8 %
gesamt	42	100 %

Tabelle 4

Einschlägige Vorstrafen (n = 42 Untergebrachte)	abs.	%
Keine	7	16,7 %
1	5	11,9 %
2 bis 5	23	54,7 %
6 bis 10	6	14,3 %
über 10	1	2,4 %
Gesamt	42	100 %

Bei der analysierten Zufallsstichprobe wurden von insgesamt 63 Gutachten 32 psychologische Gutachten erstellt. Die Zahl der Bestellungen eines psychologischen Gutachtens von Seiten des Gerichtes sind in ihrer Häufigkeit den psychiatrischen Gutachten gleichzusetzen: es wurden insgesamt 31 psychiatrische Gutachten in dieser Stichprobe erstellt. Die psychologischen Gutachten wurden von vier gerichtlich beeideten Sachverständigen im Gerichtssprengel Wien ausgeführt, die psychiatrischen Gutachten von insgesamt acht Sachverständigen. Auffällig ist die Verteilung der Gutachten, wie die folgenden Tabellen zeigen:

Tabelle 5

Verteilung der psychologischen Gutachten	abs.	%
SV A	19	59,4 %
SV B	10	31,3 %
SV C	2	6,2 %
SV D	1	3,1 %
Gesamt	32	100 %

Tabelle 6

Verteilung der psychiatrischen Gutachten	abs.	%
SV 1	10	32,3 %
SV 2	8	25,8 %
SV 3	7	22,6 %
SV 4	2	6,5 %
SV 5	1	3,2 %
SV 6	1	3,2 %
SV 7	1	3,2 %
SV 8	1	3,2 %
gesamt	31	100 %

Die Dominanz einiger Gutachter wird sehr deutlich. Es scheint in erster Linie bei der Bestellung des Sachverständigen durch das Gericht die Person, der Status oder auch der Bekanntheitsgrad des Gutachters ausschlaggebend zu sein. Dies zeigt sich sehr deutlich bei den psychologischen Gutachtern (vgl. Tabelle 5), aber auch bei den psychiatrischen Sachverständigen (vgl. Tabelle 6).

Im folgenden wird die gutachterliche psychologische Tätigkeit anhand der Gutachten der vier psychologischen Sachverständigen der Stichprobe im Entlassungsverfahren 1998 zusammengefaßt dargestellt. Die grundsätzliche Fragestellung des Gerichtes an die Gutachter ist folgende: Zum einen soll festgestellt werden, ob die einweisungsrelevante Gefährlichkeit des Unterbrachten weiterbesteht oder abgebaut wurde. Zum anderen soll eine Prognose über das voraussichtliche Verhalten nach einer allfälligen bedingten Entlassung abgegeben werden. In der Einleitung geben die psychologischen Sachverständigen im allgemeinen neben dem Namen des Auftraggebers, des Unterbrachten und dessen Alter auch den Zeitpunkt und Ort – nicht aber die Dauer - der Untersuchung sowie die genaue Fragestellung des Gerichtes wieder. Keiner der Gutachter gibt grundsätzliche Angaben zum „Untersucher“. (Ob mehrere Untersucher beteiligt waren, die nicht im Gutachten genannt werden, kann anhand des Produktes Gutachten nicht festgestellt werden. In vielen Fällen werden die testpsychologischen Untersuchungen von Assistenten durchgeführt.) Als zusätzliche Datenquelle werden den Gutachtern jeweils der gesamte Gerichtsakt (Einweisungsgutachten, eventuell schon erstellte Sachverständigengutachten in Entlassungsverfahren aus den Vorjahren, Urteile, Gerichtsprotokolle usw.) zur Gutachtenerstellung übergeben. Die Stellungnahmen des psychologischen, psychiatrischen und sozialen Dienstes der jeweiligen Anstalt sowie der Anstaltsleitung selber werden jährlich dem Gericht im Zuge des Entlassungsverfahrens übermittelt und liegen ebenfalls dem Akt bei. Sie enthalten hauptsächlich eine Beschreibung der therapeutischen Behandlung des Unterbrachten und dessen Einstellung dazu sowie die Einschätzung des jeweiligen Dienstes über Fortschritte und Entwicklung. Die einzelnen Dienste geben nicht immer eine dezidierte Empfehlung für oder gegen eine bedingte Entlassung in der Stellungnahme ab. Nicht alle psychologischen Gutachter verweisen in ihren Gutachten auf Vorinformationen. Es finden sich vor allem Auszüge aus den jährlichen Stellungnahmen der Fachdienste der Anstalten sowie Zitate und Wiedergaben aus den Einweisungsgutachten und bereits erstellten Gutachten in vergangenen Entlassungsverfahren. Aus dem Strafakt zitiert kein Gutachter explizit. Es finden sich aber in vielen Gutachten Angaben zum Delikt, der verhängten Freiheitsstrafe und dem Strafende. Auffällig ist, daß die eigentliche Befragung der Untersuchten in den Gutachten kaum bis gar nicht wiedergegeben wird. Die Befragung in Form einer Anamnese wird von keinem psychologischen Gutachter durchgeführt. Es wird im Großteil der Gutachten mit folgenden



Sätzen auf die unveränderte Lebensgeschichte verwiesen, die im Laufe der Unterbringung schon mehrmals erhoben wurde und deswegen nicht nochmals ausgeführt wird: „Da das Gutachten dem Akt angeschlossen wird, erübrigt sich ein detaillierter Aktenauszug, es werden lediglich diejenigen Sachverhalte erhoben, die für die unmittelbare Gutachtenerstellung notwendig sind“, „Der aktenkundige Inhalt, betreffend die forensische Vorgeschichte braucht nicht neuerlich wiedergegeben werden, da dieses Gutachten dem Akt anliegt“ usw.

Den Hauptteil in den Gutachten, und damit auch in der gutachterlichen psychologischen Tätigkeit, nimmt die psychologische Testung und Befunderstellung ein. Grundsätzlich kann die Testung in den Gutachten in drei verschiedene Teile unterschieden werden: Die Leistungsdiagnostik, in der vor allem Leistungs- und Intelligenztests zur Anwendung kommen, die neuropsychologische Diagnostik mittels Flimmerfrequenzanalyse sowie Verfahren zur Testung der Daueraufmerksamkeit und die Psychodiagnostik, in der vor allem Persönlichkeitstests und das Rorschach-Formdeute-Verfahren eingesetzt werden. Am meisten werden in einem Gutachten neun Testverfahren eingesetzt, am wenigsten werden in jeweils zwei Gutachten drei psychologisch diagnostische Verfahren zur Begutachtung angewandt. Der Rorschachtest findet in allen 32 psychologischen Gutachten Verwendung. In sämtlichen Gutachten fehlen die genaue Darstellung und Charakterisierung (Testversion, Alter, Verwendung usw.) der verwendeten Untersuchungsmethoden. Das Befremdendste bei der Befunderstellung ist, daß außer in den Gutachten von einem Sachverständigen, in allen anderen Gutachten die Interpretation ohne Angabe der numerischen Testresultate erfolgt. Auch Angaben zu den Testgütekriterien (Validität, Reliabilität, Normpopulationen, Standardschätzfehler usw.) fehlen. Somit ist ein Nachvollziehen der Testbefunde und Testinterpretationen nicht möglich. Ebenso der Vergleich mit früheren Testergebnissen aus bereits erstellten Gutachten wird nur in sehr wenigen Fällen bei der Interpretation explizit durchgeführt. Interessant wäre in diesem Zusammenhang ein direkter Vergleich der Testergebnisse aus früheren Gutachten mit dem aktuell durchgeführten, um einen möglichen Lerneffekt bzw. eine Verschlechterung und eventuell eine „sozial erwünschte Antworttendenz“ darzustellen. Neben der Testung der Leistungsfähigkeit des Probanden mittels eigener Verfahren werden auch in einigen Gutachten Angaben über den „neuropsychologischen Status“ und das Verhalten des Untergebrachten im Laufe der Untersuchung wiedergegeben. Einen „neuropsychologischen“ Befund erhebt nur ein Sachverständiger in seinen Gutachten. Auffallend ist in diesem Zusammenhang, daß die „Organismusvariable“ (Alkoholabhängigkeit, hormonelle Behandlung, organisches Psychosyndrom, medikamentöse Einstellung usw.) kaum bis gar nicht in die eigentliche Stellungnahme der Gutachten einfließt. Wie die Fragestellung des Gerichtes bezüglich des Abbaus der einweisungsrelevanten Gefährlichkeit in der abschließenden Stellungnahme beantwortet wird, hängt in erster Linie von der Definition des Begriffes „Gefährlichkeit“ ab. Einige psychologische Gutachter definieren einweisungsrelevante Gefährlichkeit dadurch, daß angegeben wird, aufgrund welchen abnormen Verhaltens verbunden mit einer psychischen Störung oder Abartigkeit der Untersuchte in den Maßnahmenvollzug – laut Einweisungsgutachten – eingewiesen wurde. Die Gefährlichkeit ist dann laut Stellungnahme nicht mehr gegeben, wenn das abnorme Verhalten bzw. die psychische Störung durch die unterschiedlichen Datenquelle (in erster Linie anhand der Testinterpretationen) nicht mehr feststellbar ist. In anderen psychologischen Gutachten wiederum wird „Gefährlichkeit“ dadurch bestimmt, daß angegeben wird, welche Bereiche – und hier sind vor allem auch die Testinterpretationen entscheidend – sich abweichend zur Norm darstellen. D.h. es besteht die „Einweisungsrelevante Gefährlichkeit“ beim Untergebrachten dann weiterhin, wenn dieser von der Normalität abweichendes Verhalten oder eine abweichende Einstellung zeigt. Ein weiterer Weg zur psychologischen Stellungnahme ist jener, daß aufgezeigt wird, in welchen Bereichen der Untergebrachte noch „Defizite“ aufweist. Es wird jemand dann noch als

gefährlich eingestuft, wenn er innerhalb seiner Persönlichkeit, seines Verhaltens noch soweit Defizite aufweist, aufgrund derer er in bestimmten Situationen nicht fähig ist adäquat, d.h. entsprechend der Norm, zu reagieren. Auch in diesem Fall stellen die Testinterpretationen die wesentlichste Datenquelle dar.

Die Frage nach der Zukunft wird nicht in allen Gutachten beantwortet. Grundsätzlich ist eine positive Zukunftsprognose für die Sachverständigen dann gegeben, wenn Wohn- und Arbeitsplatz nach der bedingten Entlassung vorhanden ist und wenn den Empfehlungen zu gerichtlichen Weisungen, die in den Gutachten ausgesprochen werden, Folge geleistet wird. Auch realistische Zukunftsperspektiven und Pläne nach einer bedingten Entlassung von Seiten des Untergebrachten, die er in der Befragung äußert, werden zu Beurteilung der Zukunftsprognose herangezogen. Zum Großteil wird nur in jenen Gutachten, die eine bedingte Entlassung empfehlen, die Frage nach einer positiven Zukunftsprognose überhaupt beantwortet. In Gutachten, in denen ein Weiterbestehen der Gefährlichkeit festgestellt wird, wird sehr oft auf den zweiten Teil der Fragestellung des Gerichtes nicht mehr eingegangen bzw. wird oft vermerkt, daß noch keine wesentlichen Schritte von Seiten der Anstalt unternommen wurden, daß eine Zukunftsprognose beurteilt werden kann. Dementsprechend kann die Zukunftsprognose nur dann abgegeben werden, wenn sich der Untergebrachte bereits am Freigang oder auf Vollzugslockerungen bewährt hat und Pläne getroffen wurden, wo der Untergebrachte nach einer etwaigen bedingten Entlassung arbeiten und wohnen wird. Eine weitere Möglichkeit zur Beurteilung der Zukunftsprognose findet sich in der Beurteilung jenes sozialen Netzes, welches von Seiten der Anstalt und diversen Diensten für den Untergebrachten nach einer bedingten Entlassung aufgebaut wurde. Eine positive Zukunftsprognose liegt also kaum im Einflußbereich des Untergebrachten selber, sondern obliegt vielmehr der Anstalt und deren Sonderdiensten. Auch wenn die Testwerte des Untergebrachten innerhalb der Norm liegen, aber noch keine Schritte im Zuge des Entlassungsvollzuges eingeleitet wurden, wird die Zukunftsprognose nicht beurteilt, d.h. aus den Testwerten wird keine Zukunftsprognose abgeleitet.

Im Anschluß an die Beantwortung der gerichtlichen Fragestellung werden im Großteil der psychologischen Gutachten noch Empfehlungen für gezielte therapeutische Interventionen bzw. Empfehlungen für gerichtliche Weisungen nach einer etwaigen bedingten Entlassung von Seiten der psychologischen Sachverständigen abgegeben. Empfehlungen für Weisungen werden zur Risikominimierung ausgesprochen. Es werden vor allem eine sichere bzw. betreute Wohn- und Arbeitsmöglichkeit empfohlen, therapeutische Nachbetreuung, Alkoholikerberatung usw., so daß ein kontrollierendes Netz auch nach einer bedingten Entlassung gegeben ist um damit in erster Linie das Risiko eines möglichen Rückfalls in Delinquenz zu minimieren.

Die Fragestellung des Gerichtes an die psychiatrischen Sachverständigen ist dieselbe wie an die psychologischen Gutachter. Ebenso wie bereits bei der gutachterlichen Tätigkeit der Psychiater beschrieben, erhalten auch die psychiatrischen Sachverständigen zur Begutachten den vollständigen Gerichtsakt. Die Ausgangslage für die gutachterliche Tätigkeit ist also bei beiden Gutachtern – sowohl psychiatrisch als auch psychologisch – dieselbe. Entsprechend der unterschiedlichen Wissenschaft gibt es aber grundsätzliche Unterschiede. Das Hauptuntersuchungsinstrument des Psychiaters ist die Erhebung der Anamnese und des psychopathologischen Befundes durch das Interview/Exploration. Bei 5 psychiatrischen Gutachtern nimmt die Anamnese/Exploration/Befragung (je nachdem wie der Sachverständige den Punkt „Befragung“ benennt) tatsächlich den Großteil des Gutachtens ein. Bei zwei Gutachtern nimmt jeweils die Wiedergabe aus Vorgutachten und Stellungnahmen mehr Platz im Gutachten ein. Bei einem psychiatrischen Gutachter ist dagegen der psychologische Befund (Testinterpretation), übernommen aus einem vorher erstellten psychologischen Gutachten, der Hauptteil des Gutachtens. Generelle Unterschiede zwischen

den Sachverständigen zeigen sich beim Inhalt und der Wiedergabe der Befragung in den Gutachten. Ebenso wie bei den psychologischen Gutachten aufgezeigt, gibt es auch bei den psychiatrischen Gutachtern Unterschiede betreffend Genauigkeit und Vollständigkeit in der Wiedergabe der Untersuchungsergebnisse. Einige psychiatrische Sachverständige geben das vollständige Tonbandprotokoll bzw. das vollständige Gespräch mit dem Untergebrachten im Gutachten wieder. Dadurch wird auch die Schlußfolgerung und Beantwortung der Fragestellung verständlicher und nachvollziehbarer. Dagegen finden sich bei einigen psychiatrischen Sachverständigen nur sehr ungenaue und oberflächliche Darstellungen der eigenen Untersuchung, vor allem des Gespräches mit dem Untergebrachten, das eigentlich die Hauptdatenquelle im psychiatrischen Gutachterprozeß darstellt. Dies erklärt auch die sehr hohe Heterogenität in der Länge der psychiatrischen Gutachten. Neben der Befragung bildet auch die neurologische Untersuchung einen wesentlichen Teil des psychiatrischen Sachverständigengutachten. Sie nimmt, im Gegensatz zu den psychologischen Gutachten, grundsätzlich sehr wohl Einfluß in die eigentliche Stellungnahme der psychiatrisch/neurologischen Gutachter. Jeder der acht Sachverständigen gibt in seinen Gutachten ein neurologisches/physiologisches Zustandsbild des Untergebrachten wieder. Der physiologische Befund dient auch zur Erklärung des psychischen Zustandsbildes durch organische Auffälligkeiten. Die psychodiagnostische Testung, die Haller (1996) als sehr wesentlich für eine differenzierte Persönlichkeitsdiagnostik erachtet, findet sich sehr unterschiedlich bei den einzelnen Gutachtern wieder. Ein psychiatrischer Gutachter verwendet ein psychodiagnostisches Verfahren (nach Rorschach) zur Testung der Persönlichkeit. Das psychologische Testverfahren wird vom psychiatrischen Sachverständigen selber angewandt. Zwei Gutachter setzen Instrumente zur Gefährlichkeitseinschätzung ein. Bei den übrigen Gutachtern finden sich testpsychologische Befunde aus Vorgutachten, die teilweise in die Gutachten aufgenommen werden, aber kaum in die Stellungnahme miteinbezogen werden. Nur ein psychiatrischer Gutachter verwendet zur Beantwortung der gerichtlichen Fragestellung in allen Gutachten vorwiegend die Testbefunde des psychologischen Vorgutachters. Die eigentliche Schlußfolgerung und Beantwortung der Fragestellung der anderen Gutachter erfolgt überwiegend anhand der durchgeführten Befragung. Auch der neurologische Zustandsbefund wird, wie bereits erwähnt, in den psychiatrischen Gutachten sehr oft in die Schlußfolgerung miteinbezogen. Die Definition des Begriffes „Gefährlichkeit“ wird überwiegend anhand des Krankheitsbegriffes durchgeführt, d.h. die Gefahr ist dann nicht mehr gegeben, wenn die seelische und geistige Abartigkeit höheren Grades nicht mehr vorhanden ist bzw. nachgewiesen werden kann. Die Zukunft wird, ähnlich den psychologischen Gutachten, in erster Linie in jenen Gutachten beantwortet, in denen eine bedingte Entlassung empfohlen wird. Empfehlungen für Weisungen, die in jenen Gutachten ausgesprochen werden, die eine bedingte Entlassung empfehlen, werden entsprechend der Risikominimierung eines möglichen Rückfalls abgegeben.

Nach der Darstellung der Gutachten ist in erster Linie festzuhalten, daß diese – aus nachstehenden Gründen – qualitativ nicht stellvertretend für psychologische und psychiatrische Sachverständigengutachten in Österreich stehen. Die Interessen der Auftraggebers, in diesem Fall des Richters, liegen in der Klärung des Gesundheitszustandes oder der Wesensart des Verurteilten, in der Erreichung einer zusätzlichen Absicherung der grundsätzlichen Entscheidung und dadurch vielleicht auch in der Abgabe der alleinigen Verantwortung. Es liegt aber in keiner Weise in seiner Aufgabe und offensichtlich auch nicht in seinem Interesse die eigentliche Arbeit des Sachverständigen zu überprüfen oder zu kritisieren. Rechtlich gesehen hätte der Richter grundsätzlich dazu sehr wohl die Möglichkeit. Der Klient – in diesem Fall der Untergebrachte – hat rechtlich gesehen die Möglichkeit der Ablehnung eines bestimmten Sachverständigen und die Bestellung eines Privatgutachtens.

Dies ist in der Praxis aufgrund des Fehlens der notwendigen finanziellen Mitteln und vor allem auch der Kontakte zu anderen qualifizierten Gutachtern im Vollzug kaum möglich. Die einzige rechtliche Beschwerdemöglichkeit, die auch in der Praxis genutzt wird, ist der Einspruch gegen die richterliche Entscheidung beim Oberlandesgericht. Die grundsätzliche Problematik liegt also darin, daß die Gutachter, vor allem jene, die bereits einen gewissen „Expertenstatus“ haben, im Entlassungsverfahren keinen qualitativen Mindeststandards entsprechen müssen, weil es von keiner Seite – weder vom Auftraggeber, vom Klient noch von der Gesellschaft – gefordert wird und dadurch ein „nicht nachvollziehbares“ oder „oberflächliches“ Gutachten für den Sachverständigen keine rechtlichen oder berufspolitischen Konsequenzen hat. Standards zur Gutachtenerstellung, wie sie die Arbeiterkammer in Zusammenarbeit mit dem Psychologenbeirat in Österreich zu schaffen versucht, werden aller Voraussicht nach nur in jenen Fällen eingehalten, in denen die Möglichkeit, das Interesse und die Zeit zur Überprüfung derselben gegeben sind. Auffälligkeiten dieser Stichprobe, wie fehlende numerische Testwerte oder genauere Angaben zu den Testverfahren usw., sind sicher in dem oben beschriebenen Kontext zu sehen. Die Dominanz der Testung in den psychologischen Gutachten ist sicher nicht nur in Sachverständigengutachten im Entlassungsverfahren aus der Maßnahme zu finden. Man gewinnt den Eindruck, daß der psychologische Gutachter seine Schlußfolgerung und Stellungnahme mittels psychometrischer Verfahren untermauern möchte. Die Befragung des Untersuchten nimmt, trotz Entwicklung von strukturierten Interviewverfahren und neuerer Risikodiagnostikinstrumente, kaum Einfluß auf die gutachterliche psychologische Stellungnahme. Der psychiatrische Sachverständige verläßt sich – ganz im Gegenteil – in erster Linie auf das explorative Gespräch zur Befundaufnahme. Aber auch in der psychiatrischen Begutachtung nimmt die psychologische Testung eine immer bedeutendere Rolle ein. Interessant ist aber, daß – zumindest in der dargestellten Arbeit von einem psychiatrischen Sachverständigen – die psychologische Testung von Psychiatern auch selber durchgeführt wird.

Obwohl in der aufgezeigten Stichprobe das Verhältnis an Gutachten zwischen Psychologen und Psychiatern recht ausgewogen ist, gibt es standespolitisch doch offensichtliche Unterschiede. Diese zeigen sich nicht nur in zahlreichen Gesetzen (vgl. StGB, StVG, StPO, usw.), sondern vor allem auch beim Gebührenanspruch der einzelnen Sachverständigen. Die dargestellte Arbeit setzt sich mit einem Bereich in der psychologischen und psychiatrischen Begutachtung auseinander, der allgemein ein sehr geringes gesellschaftliches und mediales Echo verursacht. Verschiedenste Ursachen sind verantwortlich für Mängel in diesen Gutachten. Es liegt aber nicht nur beim Gesetzgeber, sondern sollte vor allem im Sinne der österreichischen Berufsgruppen von Psychologen und Psychiatern sein, qualitative Standards und Richtlinien auch in diesem Bereich durchzusetzen.

## **Literatur**

- Haller, R. (1996). *Das psychiatrische Gutachten*. Wien: Manzsche Verlags- und Universitätsbuchhandlung.  
Holzbauer, A. & Brugger, S. (1996). *Strafvollzugsgesetz (StVG)*. Wien: Verlag Österreich.

## **Abkürzungsverzeichnis**

%	Prozentwert
abs.	absolute Zahlen

StGB	Strafgesetzbuch
StPO	Strafprozessordnung
StVG	Strafvollzugsgesetz
SV 1 – 8	psychiatrische Sachverständige 1 bis 8
SV A – D	psychologische Sachverständige A bis D



# Die Generierung empirischer Täterprofile: Eine post-hoc Klassifikation am Beispiel der Tötung des Intimpartners<sup>1</sup>

*Thomas P. Busch & O. Berndt Scholz*  
*Psychologisches Institut der Universität Bonn*

Seit mehr als 30 Jahren werden in Europa, in Kanada und den USA von der Polizei Täterprofile zur Unterstützung der praktischen Ermittlungsarbeit herangezogen. Lag in den 70er Jahren des letzten Jahrhunderts der Schwerpunkt der Täterprofilierung noch in den USA und Kanada, so findet seit einigen Jahren die Erstellung von psychologischen Täterprofilen auch im deutschsprachigen Raum ein zunehmendes Interesse.

Ein Täterprofil kann als eine Art psychologisches „Phantombild“ eines unbekanntes Täters verstanden werden. In einem solchen psychologischen Profil werden die charakteristischen Erlebens- und Verhaltensmuster des Täters mit dem Ziel beschrieben, die polizeilichen Ermittlungen vor Ort zu unterstützen (z.B. Åsgård, 1998; Jackson und Bekerian, 1999). Dies kann beispielsweise in Fällen notwendig werden, bei denen der Kreis potentieller Verdächtiger sehr groß ist und nicht mit anderen kriminalistischen Methoden eingeschränkt werden kann. Das gilt ebenso bei Delikten, die auf einen Wiederholungstäter schließen lassen oder wenn die Tat als Ausdruck einer psychischen Störung des Täters angesehen werden kann.

Grundsätzlich können zwei Arten von Täterprofilen unterschieden werden, (a) die *praktisch-deduktiv erstellten Täterprofile* und (b) die *wissenschaftlich-induktiv erstellten Täterprofile*.

a) Bei einem *praktisch-deduktiv orientierten Vorgehen* erfolgt die Generierung eines Täterprofils „bottom up“. Der Ausgangspunkt der Täterprofilierung ist immer ein konkreter - bis zu diesem Zeitpunkt nicht aufgeklärter - Einzelfall. Die Qualität dieser Profile ist maßgeblich von dem persönlichen Erfahrungshorizont des einzelnen Profilers mit der betreffenden Deliktart abhängig, sie werden weitgehend auf der Basis von Urteilsheuristiken erstellt. Praktisch-deduktiv erstellte Profile sind somit überwiegend impressionistisch.

b) Bei der *wissenschaftlich-induktiven Vorgehensweise* erfolgt die Täterprofilierung „top down“. Der Generierung eines Täterprofils geht die statistische Analyse einer größeren Anzahl gleichartiger, in der Vergangenheit aufgeklärter Delikte voraus. Aufgrund der empirischen Befunde kann der „ideal-typische“ Täter durch quantifizierbare und empirisch fundierte Merkmale in seinem Erleben und Verhalten beschrieben werden.

Eine Literaturanalyse zu den verwendeten Methoden der Täterprofilierung kam zu dem Ergebnis, dass die Verwendung von Begriffen wie „criminal profiling“, „offender profiling“ oder „psychological profiling“ uneinheitlich erfolgt und die Begriffsdefinitionen unscharf sind. Es kann keine einheitliche Beschreibung gefunden werden nach welchen methodischen Regeln Täterprofile zu erstellen sind. Die mangelnde Transparenz bei der Konstruktion von Täterprofilen mag vielfältige Gründe haben. Eine der wichtigen Konsequenzen besteht aber darin, dass sie sich wegen der fehlenden Transparenz einer wissenschaftlichen Beurteilung im Hinblick auf ihre methodische Genauigkeit und praktische Leistungsfähigkeit weitgehend entziehen.

Insgesamt gesehen ist für die Erstellung von Täterprofilen ebenso wie beim Einsatz anderer, damit vergleichbarer Methoden ein regelgeleitetes Vorgehen zu fordern, welches sowohl bezüglich der verwendeten Informationen als auch bezüglich der angewandten Methoden

nachvollziehbar und transparent ist. Nachfolgend wird ein regelgeleitetes Vorgehen zur Generierung wissenschaftlich-induktiver Täterprofile vorgestellt und am Beispiel der Tötung des Intimpartners praktisch demonstriert.

## Die regelgeleitete Generierung von Täterprofilen

Die wissenschaftlich-induktiv orientierte Erstellung von psychologischen Täterprofilen kann als ein Versuch verstanden werden, auffälliges Verhalten anhand von psychologischen Merkmalen zu klassifizieren. Ziel einer solchen post-hoc Klassifikation ist es, festzustellen, wie gut sich die einzelnen Täter einer Deliktart auf der Basis psychologischer Merkmale zu spezifischen Tätergruppen zusammenfassen lassen. Unter dem Gebot der Nachvollziehbarkeit und Transparenz sind die einzelnen Arbeitsschritte und Regeln zu definieren, nach denen die Analyse erfolgen soll. In Abbildung 1 wird das hier vorgestellte regelgeleitete Vorgehen zur Erstellung der Täterprofile schematisch dargestellt. Die einzelnen Schritte sollen am Beispiel der Tötung des Intimpartners praktisch demonstriert werden.

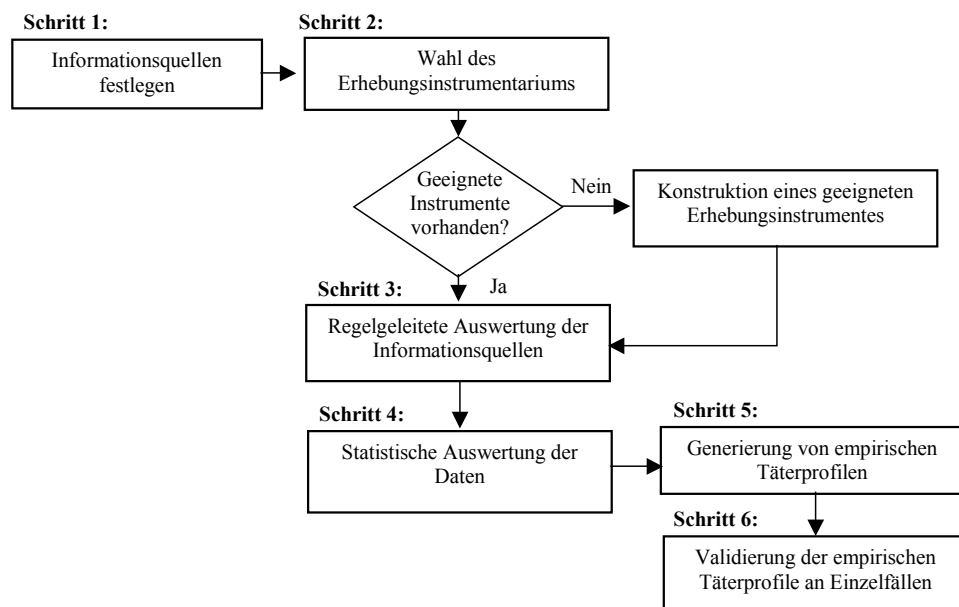


Abbildung 1: Regelgeleitetes Vorgehen zur Erstellung von Täterprofilen

### Schritt 1: Auswahl der Informationsquellen

Die Grundlage zur Erstellung von psychologischen Täterprofilen sind Informationen über Erlebens- und Verhaltensmuster von Tätern einer Deliktart (z. B. Tötungsdelikte). Bei der wissenschaftlich-induktiven Erstellung von Täterprofilen kann prinzipiell auf eine Vielfalt von geeigneten Informationsquellen zurückgegriffen werden, da der Täter ja bereits ermittelt worden ist. Allgemein kann zwischen Primärquellen (Direktes Interview des Täters, psychometrische Untersuchung des Täters, etc.) und Sekundärquellen (Aussageprotokolle des Täters, Gerichtsakten, Gutachten, etc.) unterschieden werden. Um möglichst unverfälschte Informationen zu erhalten, sollte bevorzugt auf Informationsquellen zurückgegriffen werden, die direkt vom Täter generiert werden. Bei den Primärquellen kann dies generell angenommen werden. Es finden sich allerdings auch bei den Sekundärquellen Dokumente, welche als weitgehend tätergeneriert angesehen werden können: die polizeilichen



Erstvernehmungsprotokolle.

*Auswahl der Informationsquellen am Beispiel der hier vorgestellten Studie:* Die Gesamtstichprobe bestand aus 294 Gerichtsakten von im Jahre 1993 in Bayern und Nordrhein-Westfalen zur Aburteilung gekommenen Tötungsdelikten. Aus dem umfangreichen Gesamtmaterial wurden zur weiteren Analyse randomisiert 47 polizeiliche Erstvernehmungsprotokolle gezogen. Aufgrund der Art des verübten Tötungsdeliktes wurden die Einzelfälle einer von zwei Untersuchungsgruppen zugeordnet. Die erste, deliktspezifisch sehr homogene Untersuchungsgruppe bilden die Tötungsdelikte am Intimpartner (24 Fälle). In der zweiten, deliktspezifisch eher heterogen zu nennenden Untersuchungsgruppe werden alle Tötungsdelikte zusammengefasst, bei denen nicht der Intimpartner des Täters das Opfer war (23 Fälle).

### Schritt 2: Wahl des Erhebungsinstrumentariums

Abhängig von der Art der ausgewählten bzw. zur Verfügung stehenden Informationsquelle(n) ist ein geeignetes Erhebungsinstrument zu wählen. Zur Analyse von Primärquellen liegen eine große Zahl an bereits bewährten Instrumenten vor (z. B. psychometrische Fragebögen zur Erfassung dispositioneller Verhaltensweisen, standardisierte Interviewleitfäden, etc.). Zur Analyse von Sekundärquellen ist ein Katalog von psychologischen Merkmalen erforderlich, anhand dessen die Akten analysiert werden. Um der Willkürlichkeit der Informationserhebung vorzubeugen, sollte die Konstruktion solcher Erhebungsinstrumente regelgeleitet durchgeführt werden. Folgende allgemeine Regeln sind dabei zu beachten: (1) Sichtung der relevanten empirischen und theoretischen Befunde und (2) theoriegeleitete Konstruktion des Kategoriensystems.

*Beispiel der Konstruktion eines geeigneten Erhebungsinstrumentes:* Zur regelgeleiteten und systematischen Analyse der polizeilichen Erstaussagen wurde ein spezifisches Erhebungsinstrument entwickelt, mittels dessen all jene Informationen erfasst werden können, welche das Verhalten des Täters bei der Tatausführung beschreiben und sich auf sein Erleben und Verhalten im Alltag beziehen. Von Interesse bei Tötungsdelikten sind dabei insbesondere Befunde zu folgenden psychologischen Merkmalen der Täter: Aggressivität in der Vorgeschichte (vgl. Burgheim, 1993; Gierowski, 1992; Steck, 1990; etc.), Verhalten in Konfliktsituationen (vgl. Simons, 1988; Steck, Matthes, Wenger de Chávez & Sauter, 1997), interpersonelle Interaktionen (vgl. Burgheim, 1994; Wulf, 1979; etc.), soziale Beeinträchtigungen (vgl. Glatzel, 1987; Oberlies, 1997; Rasch, 1964; etc.) und das Planungs- und Kontrollniveau der Tat (vgl. Gierowski, 1992; Ditlev & Beckmann, 1998; etc.).

Aus den theoretischen und empirischen Befundquellen zur Psychologie von Tötungsdelikten wurden 45 psychologische Merkmale abgeleitet und nach inhaltlichen Gesichtspunkten folgenden sieben Kategorien zugeordnet: (1) Aggressivität in der Vorgeschichte, (2) Verhalten in Konfliktsituationen, (3) interpersonale Interaktion, (4) soziale Beeinträchtigung, (5) delinquente Vorgeschichte, (6) Planungsniveau der Tat und (7) Kontrolle bei der Tatausführung. Das vollständige Kategoriensystem zur inhaltsanalytischen Erfassung von psychologischen Tätermerkmalen bei Tötungsdelikten ist in Tabelle 1 dargestellt.

### Schritt 3: Regelgeleitete Auswertung der Informationsquellen

Unter Zuhilfenahme des ausgewählten bzw. gemäß Schritt 2 konstruierten Erhebungsinstrumentariums werden die vorliegenden Informationsquellen regelgeleitet

analysiert. Dabei werden die Regeln maßgeblich von der Wahl des Instrumentes bestimmt. Werden beispielsweise psychometrische Fragebögen oder standardisierte Interviewleitfäden verwendet, sind die zu befolgenden Regeln dem entsprechenden Fragebogen- bzw. Interviewmanual zu entnehmen. Bei der Konstruktion eines spezifisches Erhebungsinstrumentes – wie es in der Regel bei Aktenanalysen notwendig ist – sind die zu befolgenden Anwendungsvorschriften vor der Analyse des Untersuchungsmaterials explizit zu formulieren. Wichtig ist in beiden Fällen, dass die entsprechenden Regeln dem Täterprofil beigelegt werden. Nur so kann der Empfänger des Profils nachvollziehen und beurteilen, mit welchen Methoden die Informationen über den Täter gewonnen worden sind und auf welcher wissenschaftlichen Basis die Analyse durchgeführt wurde.

*Regelgeleitete Auswertung der polizeilichen Erstaussagen:* Die 45 psychologischen Merkmale wurden in den polizeilichen Erstvernehmungsprotokollen gesucht. Für jede Einzelaussage wurde das Vorliegen bzw. Nicht-Vorliegen der einzelnen Merkmale erfasst und mittels Zahlenwerten kodiert. Um eine regelgerechte Auswertung des Untersuchungsmaterials zu gewährleisten, wurde ein Kodiermanual erstellt. In diesem Manual ist definiert, nach welchen spezifischen Regeln ein Merkmal zu kodieren ist. Zur weiteren statistischen Auswertung werden durch Addition der zugeordneten Merkmalswerte für jeden Einzelfall individuelle Kategorienwerte ermittelt. Die so berechneten sieben Kategorienwerte bilden das Kategorienprofil jedes Einzelfalles ab.

Tabelle 1: Kategoriensystem zur inhaltsanalytischen Erfassung von psychologischen Tätermerkmalen bei Tötungsdelikten

<b>Kategorie</b>	<b>Ausprägung</b>	<b>Merkmale</b>	
<b>Aggressivität in der Vorgeschichte</b>	<i>autoaggressiv</i>	Suizidversuch, Selbsterniedrigung	
	<i>fremdaggressiv</i>	körperliche Auseinandersetzung, beleidigen oder demütigen anderer Personen	
	<i>unauffällig</i>	keines der oben aufgeführten Merkmale kodierbar	
<b>Verhalten in Konfliktsituationen</b>	<i>funktional</i>	<i>Vorgeschichte</i>	Reflexion über Situation, aggressionsfreier Bewältigungsversuch
		<i>Nachtat</i>	Milderung der Folgen für das Opfer, stellt sich der Polizei
	<i>disfunktional</i>	<i>Vorgeschichte</i>	Alkohol- oder Drogenabusus, suizidale Drohung, homicidale Drohung, sonstige Drohung
		<i>Nachtat</i>	Schutzbehauptung, Schuldverschiebung
<b>interpersonale Interaktion</b>		Selbstwertbedrohende Attacke, Erfahrung von Gewalt durch das Opfer Unauffällig	
<b>soziale Beeinträchtigung</b>	<i>Schule:</i>	kein qualifizierter Schulabschluss	
	<i>Beruf:</i>	keine Berufsausbildung, Arbeitslosigkeit	
	<i>Abhängigkeit:</i>	mehr als 10000,-DM Schulden, keine eigene Wohnung	
	<i>soz. Isolation:</i>	kein Freundes- & Bekanntenkreis, keine Akzeptanz in Bezugsfamilie	
<b>delinquente Vorgeschichte</b>	<i>unauffällig</i>	keines der oben aufgeführten Merkmale kodierbar Vorstrafe in der Vorgeschichte Unauffällig	
<b>Planungsniveau der Tat</b>	<i>Zielorientiert</i>	Antizipation der Tat, Waffe gezielt mitgebracht, ausgeformtes Motiv	
	<i>ziellos</i>	hill-climbing, Waffe zufällig am Tatort, Waffe vom Opfer mitgebracht, nicht ausgeformtes Motiv	
<b>Kontrolle bei der Tatausführung</b>	<i>Ausgeprägt</i>	Schutzhandlung, Verdeckungshandlung, mehrphasiges Tatgeschehen, Alternative erkannt, Primärmotiv, Sekundärmotiv	
	<i>Schwach</i>	Übertönen, Motivverlagerung	

#### Schritt 4: Statistische Auswertung der Daten

Die so gewonnenen Daten werden einer geeigneten statistischen Auswertung unterzogen. Hier kann keinem statistischen Verfahren generell der Vorzug gegeben werden. Häufig zur Anwendung gelangen Cluster- und Diskriminanzanalysen. Lassen sich bei der Analyse zwischen den unterschiedlichen Clustern bezüglich der Merkmalsausprägungen statistisch bedeutsame Unterschiede finden, kann mit hoher Wahrscheinlichkeit davon ausgegangen werden, dass ein für die entsprechende Gruppe „typischer“ Täter die ermittelten Merkmalsausprägungen aufweist.

##### *Statistische Auswertung der Daten am Beispiel der Tötung des Intimpartners:*

Beziehungstäter bilden zwei Cluster (jeweils 11 Fälle) und eine Restgruppe (2 Fälle). Betrachtet man die Kategorienprofile der Restgruppe, so ist festzustellen, dass bei beiden Tätern während der polizeilichen Erstaussagen keine Informationen zu „Aggression in der Vorgeschichte“ erhoben wurden. Da dieser Kategorie eine diskriminatorische Bedeutung zukommt (s.u.), wird auf eine weiterführende inhaltliche Diskussion dieser Gruppe verzichtet. Zur Prüfung der Frage, ob beide Täterprofile sich statistisch bedeutsam voneinander unterscheiden und welche der sieben Kategorien zur Unterscheidung zwischen ihnen geeignet sind, wurde eine Diskriminanzanalyse durchgeführt. Dabei zeigte sich, dass die Varianz zwischen den Täterprofilen mehr als das 10fache der Streuung innerhalb der Gruppen beträgt (Eigenwert der 1. Diskriminanzfunktion: 10.161), sich die Gruppen statistisch bedeutsam voneinander unterscheiden (residuale Wilks' Lambda = .031;  $\chi^2 = 62.800$ ;  $p \leq .000$ ). Die Verortung der einzelnen Fälle auf der Diskriminanzebene ist in Abbildung 2 dargestellt.

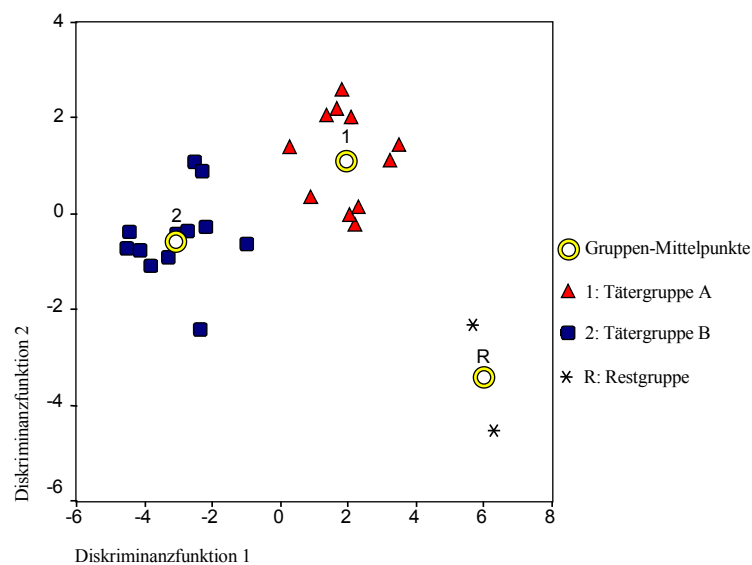


Abbildung 2: Verortung der Einzelfälle auf der Diskriminanzebene

Bei dem Streudiagramm sind auf der Abszisse die Werte der ersten Diskriminanzfunktion und auf der Ordinate die der zweiten Diskriminanzfunktion für jeden Einzelfall abgetragen. Aus der Anordnung der Punkte im Streudiagramm lässt sich ablesen, daß einerseits die Mitglieder der drei Gruppen eng um ihren jeweiligen Gruppenmittelpunkt streuen, andererseits die Gruppen deutlich von einander getrennt sind.

Die Bedeutung der diskriminatorisch relevanten Kategorien lässt sich anhand des mittleren Diskriminationskoeffizienten (mDk) beurteilen. Bei diesem Koeffizienten bedeuten höhere Werte einen höheren diskriminatorischen Einfluss der jeweiligen Kategorie. Außerdem wurde durch eine einfaktorielle ANOVA mittels multipler Vergleichstests überprüft, bei

welchen der diskriminatorisch bedeutsamen Kategorien signifikante Unterschiede ( $p \leq .05$ ) zwischen den Täterprofilen bestehen. Der Kategorie „Planungsniveau der Tat“ kann die größte diskriminatorische Bedeutung zugeschrieben werden ( $mDk = 1.123$ ;  $p \leq .001$ ). Sie wird gefolgt vom „Verhalten in Konfliktsituationen“ ( $mDk = .827$ ;  $p \leq .000$ ), der „Kontrolle bei der Tatausführung“ ( $mDk = .484$ ;  $p \leq .001$ ) und der „Aggressivität in der Vorgeschichte“ ( $mDk = .346$ ;  $p \leq .012$ ). Beide Täterprofile unterscheiden sich signifikant in den diskriminatorisch relevanten Kategorien. Bei der statistischen Überprüfung, welche der sieben Kategorien zur Differenzierung zwischen den Täterprofilen geeignet sind, zeigt sich, dass die Täterprofile bezüglich drei Kategorien gemeinsame Merkmalsausprägungen aufweisen. Die Kategorien „interpersonale Interaktion“, „soziale Beeinträchtigung“ und „delinquente Vorgeschichte“ vermögen nicht signifikant zwischen den Gruppen der Beziehungstäter zu trennen.

Schritt 5: Generierung von Täterprofilen unter Verwendung empirisch fundierter, psychologischer Merkmale

Zur Generierung von Täterprofilen werden die Ergebnisse der statistischen Analyse ausgewertet und zu einem psychologischen Profil des „typischen“ Täters integriert. Dabei sind nur diejenigen Merkmale zu berücksichtigen, welche in der Lage sind, statistisch bedeutsam zwischen den einzelnen Gruppen zur differenzieren.

*Generierung von Täterprofilen unter Verwendung empirisch fundierter, psychologischer Merkmale:* In Abbildung 3 werden die gruppentypischen Kategorienprofile der beiden Hauptgruppen graphisch dargestellt.<sup>2</sup>

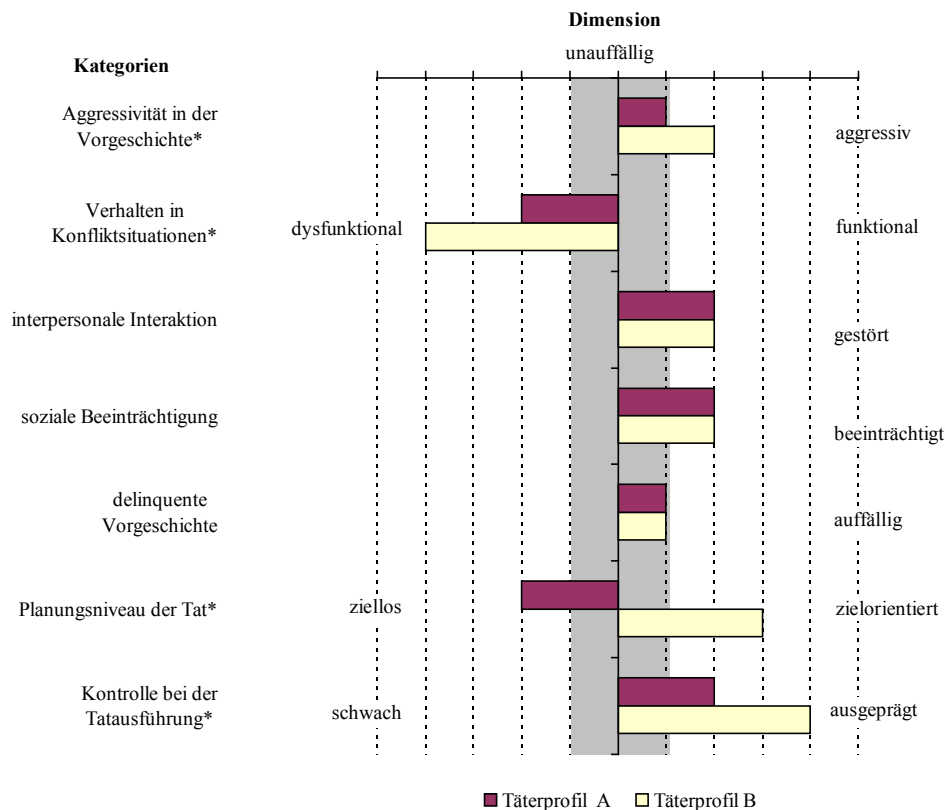


Abbildung 3: Gruppentypische Kategorienprofile der beiden Täterprofile  
 Auf der Ordinate sind die erhobenen Kategorien und auf der Abszisse die zugehörigen Dimensionen der Kategorien abgetragen. Der grau unterlegte Bereich markiert eine unauffällige Ausprägung der Kategorien. Die mit \* gekennzeichneten Kategorien trennen die beiden Täterprofile statistisch bedeutsam voneinander.

### Schritt 6: Validierung der Täterprofile an Einzelfällen

Die Validierung der Täterprofile erfolgt an Einzelfällen, die nicht in die statistische Auswertung einbezogen wurden (Validierungsgruppe). Diese Einzelfälle sollten dabei zwar aus der gleichen Deliktklasse (z.B. Tötungsdelikt) stammen, jedoch einer anderen (Unter-) Gruppe zugehören (z.B. Tötung des Intimpartners vs. sonstige Tötungsdelikte). Für jeden Einzelfall dieser Validierungsgruppe werden die zuvor als statistisch bedeutsam ermittelten psychologischen Merkmale erhoben und zu einem individuellen Profil des einzelnen Täters integriert. Dieses Einzelprofil wird mit dem statistisch ermittelten Profil verglichen. Sind die beiden Profile verschieden, d.h. zeigen die psychologischen Kategorien unterschiedliche Ausprägungen, kann davon ausgegangen werden, dass ein Profil deliktspezifisch ist. Sind jedoch das Einzelprofil und das wissenschaftlich-induktiv erstellte Profil gleich, muss dieses überprüft und weiter ausdifferenziert werden, ggf. muss es auch als zu allgemein verworfen werden.

*Validierung der Täterprofile A und B an Einzelfällen:* Um die Deliktspezifität der vorgestellten Täterprofile zu überprüfen, wurden die Profile mit den Fällen der Gruppe „sonstige Tötungsdelikte“ verglichen. Ein Vergleich der einzelnen Fälle auf kategorialer Ebene zeigt in keinem Fall eine Übereinstimmung zwischen den statistischen Profilen und

Tätern, welche ein Tötungsdelikt begingen, bei dem nicht der Intimpartner das Opfer war. Die vorgestellten gruppentypischen Täterprofile können somit für die hier untersuchte Stichprobe als deliktspezifisch für Beziehungstäter angesehen werden.

### **Abschließende Bewertung und Ausblick**

Die Ergebnisse der vorliegenden Studie machen augenfällig, dass ein gemäß den eingangs definierten Regeln geleitetes Vorgehen geeignet ist, Täterprofile zu generieren. Es konnten zwei typische Täterprofile auf der Basis quantifizierbarer und empirisch fundierter Merkmale beschrieben werden. Jedes dieser Merkmale ist erlebens- und verhaltensbezogen und weist einen mehr oder weniger direkten Bezug zur Tathandlung auf. Für die Untersuchungsgruppe „Beziehungstäter“ ließ sich zeigen, dass es bei der Tötung des Intimpartners nicht einen einzigen, umschriebenen Tätertyp gibt, sondern mindestens zwei Gruppen von Tätern, die sich deutlich im Hinblick auf Aggressivität während der Vorgeschichte, im Hinblick auf ihr Verhalten in Konfliktsituation des Alltags, das Planungsniveau der Tat und die Kontrolle bei der Tatausführung unterscheiden.

Bei der inhaltsanalytischen Bearbeitung der polizeilichen Erstvernehmungsprotokolle fällt auf, dass sie hinsichtlich zweier relevanter psychologischer Einflussfaktoren informatorische Schwächen aufweisen.

a) Die einzelnen Erstaussagen divergieren erheblich in ihrem Informationsgehalt zur interpersonalen Interaktion des Täters in der Vorgeschichte. Bei der Erstvernehmung des möglichen Täters eines Tötungsdeliktes sollte solchen Informationen substantielle Bedeutung beigemessen werden.

b) Bei einigen der hier bearbeiteten Erstvernehmungen wurden nur wenig Tatdetails angesprochen. Dies hatte zur Folge, dass zum Erleben und Verhalten des Täters bei der Tatdurchführung in diesen Fällen nur sehr wenig detaillierte Informationen vorlagen.

In zukünftigen Studien mit analoger Fragestellung sollte neben den polizeilichen Erstaussageprotokollen weiteres Datenmaterial ausgewertet werden, wie zum Beispiel die vollständigen Ermittlungs- und Gerichtsakten. Insbesondere Zeugenaussagen und forensische Berichte liefern wertvolle Informationen über das Verhalten des Täters bei der Tatdurchführung.

Im Gegensatz zu Turvey (1999) wird von den Autoren jedoch die Ansicht vertreten, dass die praktisch-deduktive und wissenschaftlich-induktive Generierung von Täterprofilen nicht als disjunkte Entitäten zu verstehen sind. Sie schließen einander nicht aus, sondern stehen in einem sich gegenseitig ergänzenden Verhältnis zueinander. So können wissenschaftlich-induktiv erstellte Täterprofile einen erheblichen Beitrag zur Differenzierung und Genauigkeit der praktisch-deduktiv erstellten Profile leisten, indem sie dem in der Praxis tätigen Profiler empirisch überprüfte Merkmalssets zugänglich machen. Der Praktiker kann damit den „Suchraum“ eines Einzelfalles eingrenzen, systematisch auswerten und daraus eine Ermittlungsstrategie spezifizieren.

### **Literatur**

Åsgård, U. (1998). Die schwedischen Erfahrungen bei der Täterprofilierung und die Bewertung einiger Aspekte eines Mord- und Entführungsfalles in Deutschland. In Bundeskriminalamt (Ed.), *Methoden der Fallanalyse: Ein internationales Symposium*. (S. 161-166). Wiesbaden: Bundeskriminalamt - Kriminalistisches Institut.

- Burgheim, J. (1993). *Psychologische Bedingungen bei Entstehung und Verlauf von Tötungsdelikten in der Situation der Partnertrennung*. Konstanz: Hartung-Gorre Verlag.
- Burgheim, J. (1994). Zur Dynamik von Tötungsverbrechen am Beispiel der sogenannten Trennungstaten. *Zeitschrift für Strafvollzug und Straffälligenhilfe*, 43, 277 – 285.
- Busch, T. P. & Scholz, O. B. (in Press). Die regelgeleitete Generierung psychologischer Täterprofile: Eine post-hoc Klassifikation am Beispiel der Tötung des Intimpartners. *Kriminalistik*.
- Ditlev, G., & Beckmann, J. (1998). Der Fall "Meschede": Fragen, Fakten und Erklärungen. In Bundeskriminalamt (Ed.), *Methoden der Fallanalyse: Ein internationales Symposium*. (S. 107-124). Wiesbaden: Bundeskriminalamt - Kriminalistisches Institut.
- Gierowski, J. K. (1992). Untersuchungen zur Motivation bei Tötungsdelikten. In T. R. Payk (Hrsg.), *Dissozialität. Psychiatrische und forensische Aspekte* (S. 65-75). Stuttgart: Schattauer.
- Glatzel, J. (1987). *Mord und Totschlag: Tötungshandlungen als Beziehungsdelikte. Eine Auswertung psychiatrischer Gutachten*. Heidelberg: Kriminalistik-Verlag.
- Jackson, J. L., & Bekerian, D. A. (1999). *Offender Profiling: Theory, Research and Practice*. Chichester: John Wiley & Sons.
- Oberlies, D. (1997). Tötungsdelikte zwischen Männern und Frauen - Eine Untersuchung geschlechtsspezifischer Unterschiede anhand von 174 Gerichtsurteilen. *Monatsschrift für Kriminologie und Strafrechtsreform*, 80, 133 – 147.
- Rasch, W. (1964). *Tötung des Intimpartners*. Stuttgart: Enke.
- Simons, D. (1988). *Tötungsdelikte als Folge misslungener Problemlösungen*. Stuttgart: Verlag für Angewandte Psychologie.
- Steck, P. (1990). Merkmalscluster bei Mordhandlungen Ergebnisse einer clusteranalytischen Studie. *Monatsschrift für Kriminologie und Strafrechtsreform*, 73(6), 384-398.
- Steck, P., Matthes, B., Wenger de Chávez, C., & Sauter, K. (1997). Tödlich endende Partnerkonflikte: Versuch einer Replikation und Erweiterung der Befunde J. Burgheims (1993, 1994). *Monatsschrift für Kriminologie und Strafrechtsreform*, 80(6), 404-418.
- Turvey, B. (1999). *Criminal Profiling - An introduction to behavioral evidence analysis*. San Diego: Academic Press.
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997b). *HCR-20: Assessing the Risk for Violence* (Version 2). Vancouver: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Wulf, B. R. (1979). *Kriminelle Karrieren von "Lebenslänglichen". Eine empirische Analyse ihrer Verlaufsformen und Strukturen anhand von 141 Straf- und Vollzugsakten*. München: Minerva.

<sup>1</sup> Vorliegende Studie entstand auf der Grundlage eines vom Zweitautor geleiteten Forschungsprojektes, das vom Bundesministerium der Justiz [Az Z B 1 – 5106 I (691) – Z2 1058/95] und des Ministeriums für Wissenschaft und Forschung des Landes NRW [Az IV A 7 – 302 00 96] im Jahre 1996 vergeben und finanziell gefördert worden war. Es wurden durch die Landesjustizministerien von Bayern und NRW Gerichtsakten von 294 im Jahre 1993 zur Aburteilung gekommener Tötungsdelikte zur Verfügung gestellt. Die Autoren danken beiden Ministerien, den Justizbehörden der Bundesländer Bayern und NRW sowie den Staatsanwaltschaften für Ihre Unterstützung.

<sup>2</sup> Eine ausführlichere Darstellung dieser Täterprofile ist zu finden bei Busch & Scholz (in Press).





# Vergleich von kriminellen Handlungen mit und ohne Gewaltanwendung hinsichtlich Tatplanung, Tathergang und Nachtatverhalten

*Matthias Fischer*

*Fachbereich Psychologie der Universität Konstanz*

## **1 Einleitung**

### 1.1 Theoretischer Hintergrund

Aufgrund ihrer meist gravierenden Folgen und ihrer Seltenheit fordern Gewalttaten zur psychologischen Erklärung heraus. Die neuere kriminalpsychologische Forschung bemüht sich dabei, das kriminelle Verhalten als Ergebnis eines Interaktionsprozesses von individuellen Dispositionen (Täterpersönlichkeit) und aktuellen Situationsgegebenheiten (Tatsituation) aufzufassen. Von verschiedenen Autoren wurden handlungstheoretische und motivationspsychologische Modelle und Überlegungen herangezogen, um neue Beschreibungs- und Erklärungsansätze für kriminelle Handlungen zu finden (z.B. Kornadt, 1974; Simons, 1988; Werbik, 1978). Gemeinsam ist diesen Ansätzen, Straftaten als Handlungen aufzufassen, die der Erreichung eines Zieles bzw. der Lösung eines Problems dienen. Verschiedene Untersuchungen folgten dieser veränderten Sichtweise und befassten sich mit der Wirkung kognitiver Faktoren auf den Verlauf der kriminellen Handlung. Als ein solcher, die Taddynamik beeinflussender Faktor wurde die Planung der Tat untersucht: bereits Lempp (1977) zeigte, dass jugendliche Mörder keine sorgfältige und längere Planung durchführen, sondern sich vielmehr kurzfristig zur Tat entschließen. Dietz (1983), Servay und Rehm (1986) sowie Volbert (1990) machten deutlich, dass die Planung von Raubdelikten häufig undetailliert und kurzfristig ist, wobei beim Täter dennoch meist die Ansicht vorhanden sei, die Tat weitgehend geplant zu haben. Den Zusammenhang zwischen mangelhafter Planung und der Anwendung von Gewalt stellt Simons (1988) her, indem er die Eskalation, die man häufig im Rahmen von Gewalttaten beobachten kann, durch eine mangelnde Planung bei der Tatvorbereitung, die wiederum durch eine aufgrund von erhöhtem Handlungsdruck entstandene Zielfixierung bedingt ist, erklärte. Diese und andere Untersuchungen (wie etwa Amir, 1971; Burgheim, 1993; Erez, 1980; Feeney, 1986; Steck et al., 1997) verweisen auf die Notwendigkeit der Einbeziehung handlungstheoretischer Konzepte, um kriminelle Gewalttaten verständlicher zu machen.

### 1.2 Eigene Fragestellung

Ziel dieser Studie ist nun, durch direkte Befragung des Täters kognitive Prozesse wie Idee und Zielsetzung, Abwägen und Planen der Tat bei verschiedenen Gewalthandlungen differenziert zu untersuchen und in Beziehung zu setzen mit Ereignissen und dem Verhalten vor, während und nach der Tat. Charakteristische Merkmale krimineller Handlungen mit Gewaltanwendung sollten durch den Vergleich mit kriminellen Handlungen ohne Gewaltanwendung (Eigentumsdelinquenz) gefunden werden. Als Rahmenmodell wird das Rubikonmodell der Handlungsphasen (Heckhausen & Gollwitzer, 1987) auf den speziellen Fall der kriminellen Handlung adaptiert. Dieses Modell sieht zur Erreichung eines Handlungsziels vier voneinander abgrenzbare Phasen, nämlich Abwägen, Planen, Handeln und Bewerten mit jeweils spezifischen Bewusstseinslagen vor. Ergänzend wird das Konzept der „Vorszene“ von Kahlert und Lamparter (1979) zur Binnenstrukturierung der Handlungsphase herangezogen.

Die Operationalisierung ist angelehnt an das Vorgehen Burgheims (1993) und Stecks (1997). Demgemäß wurden behaviorale, motivationale und volitionale Daten für die Tatanlaufzeit (Zeitraum der letzten vier Wochen vor der Tat), die Vorszene (die letzten Stunden vor der Tat), den Tathergang und das Nachtatverhalten des Täters erfasst. Der handlungspsychologische Ansatz wird dabei um eine persönlichkeitspsychologische Perspektive ergänzt, die über das Konzept der kriminellen Karriere in der Auslegung Göppingers (1975) operationalisiert wird.

## 2 Methode

### 2.1 Untersuchungsgruppen

Für den Vergleich der kriminellen Handlungen wurden 175 männliche Straftäter zu einer begangenen Straftat und zu ihrer Biographie befragt. Die Stichprobe wurde in folgende vier Tatgruppen gegliedert: Raub (N = 53), Vergewaltigung (N = 20) und Aggressionstat (berücksichtigt sind hier Gewalttaten ohne Verbindung zu Raub oder sexueller Gewalt, auch ohne Beziehungstaten) (N = 37) als Repräsentanten krimineller Handlungen mit Gewaltanwendung (Gewalt mit tödlichem Ausgang ist in der Stichprobe vertreten); Einbruchsdiebstahl als kriminelle Handlung ohne Gewaltanwendung.

Für den täterspezifischen Vergleich hinsichtlich der kriminellen Karriere wurden folgende sechs Tätergruppen gebildet: Räuber (N = 39), Vergewaltiger (N = 20), Aggressionstäter (N = 28) und Intensivtäter (N = 30) als Typen von Gewalttätern, Betrüger und Einbrecher als Gruppen von „nicht-Gewalttätern“. Die Täter wurden für diesen Vergleich nicht nach den berichteten Straftaten eingeteilt, sondern nach Sichtung der Vorstrafenregister. Intensivtäter sind hier Täter, die mindestens drei der untersuchten Deliktarten in jeweils mindestens zehn Fällen verübt hatten.

### 2.2 Datenmaterial- und Verarbeitung

Die der Analyse zugrundeliegenden Daten wurden in der Zeit vom Mai 1999 bis zum Mai 2000 durch Mitarbeiter und Mitarbeiterinnen der Arbeitsgruppe Rechtspsychologie der Universität Konstanz erhoben. Mithilfe eines halbstandardisierten Interviews wurden insgesamt 175 Straftäter in verschiedenen Justizvollzugsanstalten Süddeutschlands zu einer begangenen Straftat und ihrer Biographie befragt. Zusätzlich wurden vor der Befragung die Gerichtsurteile eingesehen. Um eine statistische Analyse zu ermöglichen, wurden die Antworten auf die Fragen nach festgelegten inhaltlichen Kriterien kodiert und dichotomisiert. Die so gebildeten Variablen wurden analog dem Vorgehen Burgheims (1993) nach thematischen Gesichtspunkten zu 12 Quasi-Skalen zusammengefasst (*siehe Tab. 1*). Zusätzlich wurde noch die Variable „Unerwartetes Ereignis im Tatverlauf“ gebildet. Dieses Merkmal sollte dann zutreffen, wenn der Proband ein Ereignis während der Interaktion mit dem Opfer berichtet (z.B. Provokationen, Gewaltanwendung, Schreien oder Drohungen seitens des Opfers), das zu einer Verunsicherung seitens des Täters führte.

Die Auswahl der als relevant erachteten Variablen basiert insbesondere auf den Arbeiten von Burgheim (1993), Göppinger (1975), Simons (1988), Steck, Schrader und Aucher (1999) sowie Volbert (1990). Hinzugefügt wurden Fragen, die am Rubikonmodell der Handlungsphasen von Heckhausen und Gollwitzer (1987) orientiert sind. Der statistische Vergleich der Täter- bzw. Tatgruppen hinsichtlich der 12 Quasi-Skalen erfolgte mit einem Rangtest für ordinalskalierte abhängige Variablen ( $\chi^2$  – Test nach van-der-Waerden) auf

dem 5 %-Signifikanzniveau. Eine Alpha-Fehler-Adjustierung wurde nicht vorgenommen, zum einen, um aufgrund des explorativen Charakters der Studie den Fehler zweiter Art möglichst gering zu halten, zum anderen, weil die Hypothesen nicht als völlig unabhängig voneinander betrachtet werden können.

### 3 Ergebnisse

#### 3.1 Vergleich der sechs Tätergruppen hinsichtlich biographischer Belastungsfaktoren

Der Rangvergleich zwischen den sechs Gruppen hinsichtlich der Häufung biographischer Belastungsmerkmale führte in allen vier untersuchten Bereichen zu einem signifikanten Ergebnis (*siehe Tab. 2*).

Tabelle 1: Die 13 Quasi-Skalen und ihre Themenbereiche

Quasi-Skalen	Themen:
<i>Biographischer Bereich (Kriminelle Karriere im Sinne Göppingers)</i>	
1. Familiäre Belastungsfaktoren	Strukturelle Unvollständigkeit durch Scheidung oder Tod eines Elternteils, Heimaufenthalte, Misshandlung, Sucht in der Kernfamilie
2. Psychosoziale Auffälligkeiten	Schulische Auffälligkeiten, Ausbildungsabbrüche, Suchtprobleme, psychische Auffälligkeiten
3. Kriminelle Karriere	Jugenddelinquenz, Rückfälligkeit, Gruppenkriminalität
4. Gewaltbereitschaft	Gewalttätigkeiten gegenüber Geschwistern, Mitschülern, Partner, „Schlägereien“
<i>Tatvorfeld</i>	
5. Ökonomische Deprivation	Arbeitslosigkeit, geringes Einkommen, Schulden
6. Verhaltensveränderungen	ungewöhnliche Empfindungen und Verhaltensweisen wie Schlafstörungen, Gereiztheit, Fehltage bei der Arbeit
7. Alkohol- und Drogenkonsum	erheblicher Alkoholkonsum, Konsum weicher und harter Drogen
8. Kosten- Nutzen- Erwägungen	Abwägen, Beschäftigung mit alternativen Problemlöseversuchen und dem Tatrisko
9 Intensität der Tatantizipation	Antizipation von Schwierigkeiten, Tatabbruch, Gewaltanwendung und Nachtatverhalten
<i>Vorszene</i>	
10. Konfliktverschärfende Ereignisse	Besonderes Ereignis, Alkohol- und Drogenkonsum, negative Stimmung
<i>Tat</i>	
11. Merkmale der Bedrohtheit und Verunsicherung	Täter fühlte sich durch Opfer provoziert, körperlich attackiert, bedroht; unerwartetes Verhalten seitens des Opfers
<i>Nachtatverhalten</i>	
12. Ungeordnetheit des Nachtatverhaltens	Angst, Verwirrung, Selbstanzeige, Suizidversuche im Anschluss der Tat

Tabelle 2: Mediane der Anzahl der zutreffenden Variablen bzgl. der vier Skalen zu den biographischen Merkmalen der Täter

Quasi-Skala (Anzahl der Variablen)	Betrüger N = 18	Einbrecher N = 26	Räuber N = 39	Vergewaltiger N = 20	Aggressions-täter N = 28	Intensiv-täter N = 30	Chi <sup>2</sup> (df = 5) Signifikanz
Familiäre Belastungsfaktoren (11)	1	3	1	2	1.5	3	20.763 p = .00
Psychosoziale Auffälligkeiten (12)	3	3	3	5	4	6	31.458 p = .00
Kriminelle Karriere (6)	2	4	3	2	3	5	51.8791 p = .00
Gewaltbereitschaft (5)	0	1	1	1	2	2	39.1578 p = .00

### 3.2 Vergleich der vier Tatgruppen hinsichtlich konfliktverschärfender Verhaltensmuster und Ereignisse im Tatvorfeld (Zeitraum der letzten vier Wochen vor der Tat bis zum Tattag)

Der prüfstatistische Vergleich der vier Tatgruppen hinsichtlich konfliktverschärfender Ereignisse und Verhaltensmuster im Tatvorfeld führte nur für die Skala „ökonomische Deprivation im Tatvorfeld“ zu einem signifikanten Ergebnis (siehe Tab.3). Hinsichtlich „Verhaltensänderungen im Tatvorfeld“ sowie „Alkohol- und Drogenkonsum im Tatvorfeld“ ergaben sich keine signifikanten Ergebnisse.

Tabelle 3: Mediane der Anzahl der zutreffenden Variablen bzgl. der Skala „Ökonomische Deprivation im Tatvorfeld“

Quasi-Skala (Anzahl der Variablen)	Einbruch N = 26	Raub N = 53	Vergewaltigung N = 20	Aggressionstat N = 37	Chi <sup>2</sup> (df = 3) sign.
Ökonomische Deprivation im Tatvorfeld (6)	3	3	1	1	20.2421 p = .00

### 3.3 Vergleich der vier Tatgruppen hinsichtlich kognitiver Merkmale des Abwägens und Planens im Tatvorfeld (Zeitraum der letzten vier Wochen vor der Tat bis zum Tattag)

Die Vergleiche der vier Tatgruppen bezüglich der Skalen „Kosten-Nutzen-Erwägungen“ sowie „Intensität der Tatantizipation“ erbrachten jeweils signifikante Unterschiede zwischen den Tatgruppen (siehe Tab. 4).

Tabelle 4: Mediane der Anzahl der zutreffenden Variablen bzgl. der Skalen zum Abwägen und Planen der Tat

Quasi-Skala (Anzahl der Variablen)	Einbruch N = 26	Raub N = 53	Vergewaltigung N = 20	Aggressionstat N = 37	Chi <sup>2</sup> (df = 3) sign.
Kosten-Nutzen-Erwägungen (5)	2	2	1	0	26.0322 p = .00
Intensität der Tatantizipation (5)	2.5	2	0	0	35.2118 p = .00

### 3.4 Vergleich der vier Tatgruppen hinsichtlich der Tatmerkmale

Der Vergleich der vier Tatgruppen hinsichtlich konfliktverschärfender Ereignisse (Skala 10) in der Vorszene führte zu keinem signifikanten Ergebnis. Hinsichtlich der Skala 11 „Merkmale der Bedrohtheit und Verunsicherung“ unterschieden sich die Gruppen jedoch signifikant (siehe Tab. 5).

Tabelle 5: Mediane der Anzahl der zutreffenden Variablen bzgl. der Skala 11 „Merkmale der Bedrohtheit und Verunsicherung“

Quasi-Skala (Anzahl der Variablen)	Einbruch N = 26	Raub N = 52	Vergewaltigung N = 19	Aggressionstat N = 37	Chi <sup>2</sup> (df = 3) sign
Merkmale der Bedrohtheit und Verunsicherung (6)	0	0	2	3	35.9599 p = .00

Der Vergleich der vier Tatgruppen hinsichtlich der Merkmale der Skala „Ungeordnetheit des Nachtatverhaltens“ führte zu einem signifikanten Ergebnis (siehe Tab. 6).

Tabelle 6: Mediane der Anzahl der zutreffenden Variablen bzgl. der Skala 12 „Ungeordnetheit des Nachtatverhaltens“

Quasi-Skala (Anzahl der Variablen)	Einbruch N = 26	Raub N = 52	Vergewaltigung N = 19	Aggressionstat N = 37	Chi <sup>2</sup> (df = 3) sign
Ungeordnetheit des Nachtatverhaltens	0	1	2	1	23.4023 p = .00

### 3.5 Korrelationsstatistische Analyse

Wie aus Tabelle 7 ersichtlich bestehen signifikante negative Zusammenhänge zwischen dem Merkmal „Unerwartetes Ereignis im Tatverlauf“ und den Skalen „Gedankliche Hemmungen gegenüber der Tat“ sowie „Intensität der Tatantizipation“. Zwischen dem Merkmal „Unerwartetes Ereignis im Tatverlauf“ und der Skala „Ungeordnetheit des Nachtatverhaltens“ fand sich dagegen ein positiver Zusammenhang.

Tabelle 7: Korrelationsmatrix

Merkmal	Skala „Kosten-Nutzen-Erwägungen“	Skala „Intensität der Tatantizipation“	Skala „Ungeordnetheit des Nachtatverhaltens“
<b>Unerwartetes Ereignis im Tatverlauf</b>	-0,3314* (N = 133)	-0,3022* (N = 133)	0,4032* (N = 129)

## 4 Diskussion

Bei Betrachtung der Rangvergleiche hinsichtlich der Häufung biographischer Belastungsfaktoren deuten die vier signifikanten Ergebnisse darauf hin, dass die sechs Tätergruppen quantitativ unterschiedlich stark ungünstigen Sozialisationsbedingungen ausgesetzt waren. Dieses Ergebnis erhärtet die von Sannemüller (1999) geäußerte Vermutung, dass sich die Sozialisationsbedingungen bei verschiedenen Tätergruppen unterscheiden und sich deshalb eine differenzierte Betrachtung lohne. Zwei Tätergruppen profilieren sich dabei besonders: Für die Gruppe der Intensivtäter fanden sich in allen vier untersuchten biographischen Bereichen die höchsten Belastungen. Die Gruppe der Betrüger weist dagegen in allen Bereichen die geringste Häufung an biographischen Belastungsfaktoren auf. Die Vermutung, dass Gewalttäter verstärkt ungünstige biographische Besonderheiten aufweisen als Nicht-Gewalttäter lässt sich allerdings nicht bestätigen, da auch die Gruppe der Einbrecher insbesondere beim Vergleich der familiären Belastungsfaktoren sowie hinsichtlich der Merkmale der kriminellen Karriere hohe Werte aufweist. Dieses Ergebnis deutet darauf hin, dass ungünstige Sozialisationsbedingungen zwar mit der Häufigkeit und Breite krimineller Aktivität in positivem Zusammenhang steht, nicht notwendig aber mit Gewaltkriminalität. Der Schlüssel zum Verständnis krimineller Handlungen mit Gewaltanwendung könnte aber in der Untersuchung der kriminellen Handlung selbst gefunden werden.

Die Vergleiche hinsichtlich konfliktverschärfender Ereignisse und Verhaltensweisen im Tatvorfeld (Zeitraum der letzten vier Wochen vor der Tat) führte nur bei einer von drei untersuchten Skalen zu einem signifikanten Unterschied zwischen den vier Tatgruppen, nämlich hinsichtlich der Häufung der Merkmale ökonomischer Deprivation. Erwartungsgemäß weisen in diesem Bereich die beiden Gruppen der Vermögenstaten eine stärkere Belastung auf als die beiden Gruppen der Gewalttaten ohne das Ziel der Bereicherung. Raub und Einbruchsdiebstahl unterscheiden sich allerdings nicht, es ist hier also von ähnlichen finanziellen Problemsituationen auszugehen.

Hinsichtlich der Skalen zum Abwägen („Kosten-Nutzen-Erwägungen“) und Planen der Tat („Intensität der Tatantizipation“) zeichnet sich in den Ergebnissen eine Zweiteilung der vier Tatgruppen ab: die Eigentumstaten Einbruchsdiebstahl und Raub zeichnen sich durch eine elaboriertere gedankliche Beschäftigung in den Tagen und Wochen vor der Tat aus als Vergewaltigung und Aggressionstat. Vermutlich durch die finanzielle Problemsituation angeregt kommt es bei den beiden erstgenannten Gruppen also eher zum Abwägen verschiedener Alternativen, Bedenken des Tatriskos und des Strafmaßes und zu einer intensiveren Antizipation des Tatverlaufs wie dem Bedenken eventuell auftretender Probleme und Hindernisse im Tatverlauf als bei den beiden letztgenannten Gruppen. Diese scheinen kaum antizipiert und geplant zu sein und tragen mehr den Charakter einer spontanen „Impulshandlung“, ein Befund, der die Beobachtungen von Amir (1971), Sexualstraftaten seien mehrheitlich geplant, nicht unterstützt. Im Tatverlauf selbst profilieren sich diese beiden Tatgruppen dann als Handlungen, die vergleichsweise häufiger mit Verunsicherungen seitens des Täters aufgrund unerwartetem, aggressiv empfundenem Opferverhalten (Drohungen, Beleidigungen, Schreien, Gewaltanwendung) gekennzeichnet sind. Einbruch und Raub weisen, vermutlich aufgrund der intensiveren Antizipation des Tatverlaufes, diese Merkmale selten auf. Dass es bedeutsame Beziehungen zwischen der kognitiven Auseinandersetzung mit der Tat im Tatvorfeld, dem Auftreten von unerwarteten Ereignissen im Tatverlauf sowie der Geordnetheit des Nachtatverhaltens gibt, darauf deuten die signifikanten Korrelationen zwischen diesen Variablen hin. Je weniger der Tatverlauf antizipiert wird, je weniger Kosten-Nutzen-Erwägungen stattfinden, desto eher werden nach der Tat unerwartete, verunsichernde Ereignisse erinnert. Das Berichten unerwarteter Ereignisse korreliert wiederum positiv mit Merkmalen eines ungeordneten, desorganisierten Nachtatverhaltens. Handlungstheoretisch ließe sich dieser Befund dahingehend interpretieren, dass aufgrund von unerwarteten Ereignissen nicht antizipierte Anwendung von Gewalt erfolgt, diese Zielkonflikte hervorruft, die sich schließlich in einem ungeordneten, chaotischen Verhalten nach der Tat äußern. Wenn diese Interpretationen zutreffen, würde der Befund zum Gruppenvergleich hinsichtlich des Nachtatverhaltens die Tatgruppe Vergewaltigung als die Gruppe auszeichnen, die am stärksten durch Zielkonflikte im Tatverlauf charakterisiert ist, während Einbruch sich als relativ arm an Zielkonflikten darstellt. Insgesamt unterstreichen diese letztgenannten Befunde die Bedeutsamkeit der schon von Lempp (1977) geäußerten Vermutung, dass insbesondere Überraschungssituationen im Tatverlauf zu eskalierenden, kriminellen Gewalthandlungen beitragen.

## Literatur

- Amir, M. (1971). *Patterns in forcible rape*. Chicago: University of Chicago Press.
- Burgheim, J. (1993). *Psychologische Bedingungen bei Entstehung und Verlauf von Tötungsdelikten in der Situation der Partnertrennung*. Dissertation. Konstanz: Hartung-Gorre.
- Dietz, M.L. (1983). *Killing for Profit. The Social Organization of Felony Homicide*. Chicago: Nelson-Hall.

- Erez, E. (1980). Planning of crime and the criminal career: official and hidden offenses. *Journal of Criminal Law & Criminology*, 71, 73-77.
- Feeney, F. (1986). Robbers as Decision-Makers. In D.B. Cornish & R.V. Clarke (Eds.), *The Reasoning Criminal: rational choice perspectives on offending*, 53-71. New York: Springer.
- Göppinger, H. (1975). Homicide and Criminal Career. A First Provisional Report of the Investigations of Murderers at Tübingen. *Rassegna di Criminologia*, 6, 39-45.
- Heckhausen, H. & Gollwitzer, P.M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion*, 11, 101-120.
- Kahlert, T.H. & Lamparter, U. (1979). Tötungsdelikte bei Jugendlichen und Heranwachsenden. *Monatsschrift für Kriminologie und Strafrechtsreform*, 62, 206-217.
- Kornadt, H.J. (1974). Toward a Motivational Theory of Aggression and Aggression Inhibition: some Considerations about an Aggression Motive and their Application to TAT and Catharsis. In J. De Wit and W.W. Hartup (Eds.), *Determinants and Origins of Aggressive Behavior*. Den Haag: Mouton.
- Lempp, R. (1977). *Jugendliche Mörder. Eine Darstellung an 80 vollendeten und versuchten Tötungsdelikten von Jugendlichen und Heranwachsenden*. Bern: Huber.
- Sannemüller, U., Ullrich, S., Pillmann, F., Draba, S. & Marneros, A. (1999). Tötungsdelikte - soziodemographischer Hintergrund der Täter und tatspezifische Merkmale. *Archiv für Kriminologie*, 65-74.
- Servay, W. & Rehm, J. (1986). *Bankraub aus Sicht der Täter : täterleitende Faktoren bei Raubüberfällen auf Geldinstitute*. Wiesbaden: Mittelbayerische Druckerei- und Verlagsgesellschaft.
- Simons, D. (1988). *Tötungsdelikte als Folge mißlungener Problemlösungen*. Stuttgart: Verlag für Angewandte Psychologie.
- Smettan, J.R. (1992). *Kriminelle Bereicherung in Abhängigkeit von Gewinnen, Risiken, Strafen und Moral*. Freiburg: Eigenverlag Max-Planck-Institut.
- Steck, P., Matthes, B., Wenger de Chavez, C. & Sauter, K. (1997). Tödlich endende Partnerkonflikte. *Monatsschrift für Kriminologie und Strafrechtsreform*, 80, 404-417.
- Steck, P., Schrader, S. & Auchter, U. (1999). *Konfliktverlauf und Verhaltensmuster bei Tötungsdelikten*. Unveröffentlichter Forschungsbericht. Konstanz: Fachgruppe Psychologie der Universität Konstanz.
- Volbert, R. (1990). *Tötungsdelikte im Rahmen von Bereicherungstaten*. Dissertation. Berlin: Technische Universität.
- Werbik, H. (1978). *Handlungstheorien*. Stuttgart: Kohlhammer.





# Experimentelle Testung der Neutralisationstheorie abweichenden Verhaltens

*Immo Fritsche*

*Institut für Psychologie der Universität Magdeburg*

Die Neutralisationstheorie – erstmals formuliert von den amerikanischen Soziologen Sykes & Matza (1957) – besitzt eine ungebrochene Popularität bei der Erklärung normwidrigen Verhaltens. Die Grundannahme, Personen brächen gesellschaftliche Normen nicht aus einem Mangel an Normakzeptanz, sondern weil sie in der Lage sind, diese von ihnen selbst geteilten Normen mithilfe von Rechtfertigungen und Entschuldigungen situativ zu neutralisieren, wurde auch über den strafrechtlich relevanten Bereich hinaus auf andere Normverletzungen übertragen. Trotz des hohen Verbreitungsgrades der Theorie war es der psychologischen und soziologischen Forschung bislang nicht möglich, die kausale Wirkung von Neutralisationen auf Verhalten angemessen zu belegen, da zumeist lediglich Korrelationsstudien mit Selbstberichtsdaten vorliegen. In der vorgestellten Studie wurde erstmals ein laborexperimenteller Zugang zum Nachweis der Wirkung von Neutralisationen auf beobachtetes Verhalten gewählt. Die experimentelle Manipulation bestand in der Hemmung bzw. der Förderung der Neutralisierungsfähigkeit einer Versuchsperson, das anschließende normwidrige Verhalten (Benutzen von Einweg-Getränkedosen) wurde im Labor beobachtet. Die Erhebung zahlreicher Kontrollvariablen, wie z.B. die Normbindung der Vpn, ermöglicht die Untersuchung der Bedingungen der Neutralisations-Verhaltens-Wirkung. Es werden die Ergebnisse zweier unabhängiger Experimente (N = 80, N = 54) vorgestellt, von denen lediglich eines die Kernannahme der Theorie bestätigt. Die Voraussetzungen des Neutralisationseffekts werden diskutiert.

## Die Neutralisationstheorie abweichenden Verhaltens

Im Jahr 1957 stellten G. M. Sykes und D. Matza im American Sociological Review eine Theorie zur Erklärung delinquenten Verhaltens Jugendlicher vor, die sich bis zum heutigen Tag im Erklärungsrepertoire der Devianzforschung wiederfindet. Trotz vereinzelter Rekonzeptualisierungsbemühungen blieb die Kernidee der Theorie weitgehend unbeschadet: Demnach verletzen Individuen gesellschaftliche Normen nicht, weil sie diese nicht teilen, sondern weil sie in der Lage sind, individuell geteilte Normen situativ zu neutralisieren. Dies geschieht durch den Einsatz von Rechenschaftskognitionen, den sogenannten Neutralisierungstechniken. Sykes & Matza (1957) unterscheiden zwei funktionale Typen von Rechenschaften: Neutralisierungen und Rationalisierungen (siehe Abb. 1). Rationalisierungen treten auf, nachdem ein normwidriges Verhalten ausgeführt wurde und schützen das Individuum vor Selbst- und Fremdvorwürfen. Neutralisierungen hingegen sollte eine Person im Vorfeld eines Fehlverhaltens entwickeln und den normwidrigen Akt dadurch erst ermöglichen.

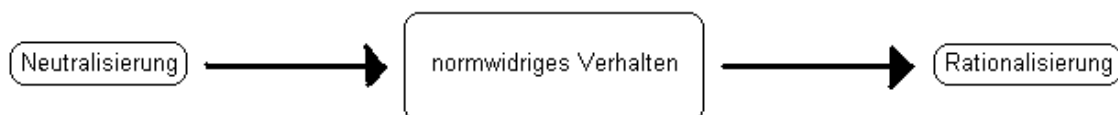


Abbildung 1: Zeitliche Verknüpfung der Variablen der Neutralisationstheorie

Gemessen an ihrer theoretisch-inhaltlichen Reichweite wurde die These, Neutralisierungen seien Prädiktoren individuellen Verhaltens, in der Folgezeit erstaunlich selten und methodisch unzulänglich überprüft. In der psychologischen und soziologischen Literatur finden sich insgesamt 14 Studien, die einen Zusammenhang zwischen Neutralisation und Verhalten

belegen (Ball, 1966; Egg & Sponzel, 1978; Norris & Dodder, 1979; Thurman, 1984; Amelang, Schahn & Kohlmann, 1988; Amelang, Zahn & Schahn, 1988; Hollinger, 1991; Dodder & Huges, 1993; Schahn, Dinger & Bohner, 1995; Winkel, 1997; Minor, 1981, 1984; Shields & Whitehall, 1994; Schwarz & Beyer, 1989). Mehrheitlich handelt es sich bei diesen Untersuchungen um Fragebogenstudien mit Einmalerhebung. Die erfragte Akzeptanz von rechtfertigenden und/oder entschuldigenden Aussagen wird hierbei in der Regel mit selbstberichteten Verhaltensscores korreliert. Diese Studien weisen z.T. gravierende methodische Unzulänglichkeiten auf.

a) Zehn der insgesamt 14 einschlägigen Studien liegen Daten aus Einmalerhebungen zugrunde. Diese Untersuchungen können daher nichts darüber aussagen, ob Neutralisationen normwidrigem Verhalten tatsächlich ursächlich vorangehen. Der in allen Studien gefundene Zusammenhang zwischen Neutralisationsakzeptanz und angegebenem normwidrigem Verhalten könnte ebensogut als Beleg für Rationalisierungsbedarf im Anschluss an Fehlverhalten gewertet werden.

Die drei vorliegenden Längsschnittuntersuchungen (Minor, 1981, 1984; Shields & Whitehall, 1994) können diesen methodischen Einwand nur bedingt ausräumen, da eine Konfundierung der erhobenen Neutralisationstendenz mit anderen Wirkvariablen nicht auszuschließen ist. Hinsichtlich der Kausalitätsfrage eindeutige Aussagen können lediglich experimentelle Versuchspläne liefern, in denen Neutralisation als echte unabhängige Variable randomisiert induziert wird, sich möglicherweise konfundierte Wirkgrößen also zufällig auf die untersuchten Personengruppen verteilen. Somit wird eine Kontrolle dieser unbekannt potentiellen Einflussgrößen möglich. Erstaunlicherweise berichten lediglich Schwarz & Beyer (1989) von der Umsetzung eines solchen Versuchsplans zur Testung der Neutralisations-Verhaltens-Hypothese. Leider ist die Induktion der Verfügbarkeit von Neutralisationen in dieser Studie unbeabsichtigt mit einer Induktion der kognitiven Verfügbarkeit eigenen normwidrigen Verhaltens verbunden, so dass auch diese Untersuchung hinsichtlich unserer Fragestellung nicht eindeutig interpretierbar ist.

b) Darüber hinaus arbeiten alle Studien zumindest teilweise mit *Selbstberichtsfragebögen*. Neutralisierung und Verhalten werden hierbei nicht beobachtet, sondern retrospektiv oder prospektiv erfragt. Eine Ausnahme bildet die Studie von Shields & Whitehall (1994), in der zumindest die Verhaltensvariable durch den objektiven Indikator polizeilicher Kriminalitätsakten erhoben wird. Auf ein ähnliches Außenkriterium für die Bestimmung abweichenden Verhaltens greifen auch jene Studien zurück, in denen eine Stichprobe von verurteilten Straftätern mit diesbezüglich unauffälligen Personen verglichen wird (Ball, 1966; Amelang, Schahn & Kohlmann, 1988; Shields & Whitehall, 1994).

c) Auch die Erhebungsweise der Neutralisationen unterscheidet sich von Studie zu Studie. Allerdings übernehmen einige AutorInnen die von Ball (1966) erstmals eingeführte Szenariotechnik zur Erhebung der Neutralisationsakzeptanz (Minor, 1981; Amelang, Schahn & Kohlmann, 1988; Amelang, Zahn & Schahn, 1988; Schwarz & Beyer, 1989; Shields & Whitehall, 1994; Schahn, Dinger & Bohner, 1995) und wählen die erhobenen Neutralisationen gemäß dem von Sykes & Matza (1957) sowie Minor (1981) vorgeschlagenen Neutralisationskatalog. Dieser Zusammenstellung von Neutralisationstechniken mangelt es zwar an einer systematisch-theoretischen Fundierung (Fritsche & Mayrhofer, 2001), doch zugunsten einer erhöhten Vergleichbarkeit der Untersuchungen sollte in Ermangelung von Alternativen zunächst auf diese klassische Liste zurückgegriffen werden.

d) Mit Ausnahme der Studie von Schwarz & Beyer (1989) liegt allen vorgestellten Untersuchungen die konzeptuelle Unzulänglichkeit zugrunde, die Neutralisations-Verhaltens-Beziehung nicht auf situationaler Ebene zu untersuchen. Grund ist das Fehlen eines entsprechenden Verhaltensmodells in der – überwiegend soziologisch geprägten – Neutralisationsforschung. Die Entwicklung eines situationsbezogenen Wirkungsmodells von Neutralisationen aber könnte angesichts der uneinheitlichen Ergebnisse der Längsschnittforschung (Minor, 1981, 1984; Shields & Whitehall, 1994) den Schlüssel zum Verständnis der Beziehung zwischen Neutralisation und Verhalten bergen.

Bemerkenswert ist die Erweiterung des Anwendungsbereichs der Neutralisationstheorie auf nicht-kriminelles Verhalten. Einerseits werden in der Mehrzahl der Neutralisations-Verhaltens-Studien neben anderen auch sehr leichte Formen abweichenden Verhaltens untersucht (z. B. Schummeln in der Prüfung und vorehelicher Geschlechtsverkehr bei Minor, 1981, Bummeln und zu spätes Erscheinen zur Arbeit bei Hollinger (1991) oder ungesetzlicher Alkoholgenuss bei Dodder & Huges, 1993). Andererseits bezieht sich die Untersuchung von Schahn, Dinger & Bohner (1995) auf den weitgehend strafrechtsfernen Bereich umweltschädigenden Verhaltens im Alltag. Eine Verallgemeinerung der Theorie auf jegliches normabweichende Verhalten, wie beispielsweise von Opp (1974, S. 105-110) vorgeschlagen, wird also auch durch die Praxis der Neutralisations-Verhaltens-Forschung gedeckt.

Trotz der weiter oben beschriebenen methodischen Unzulänglichkeiten bei der Identifizierung eines kausalen Zusammenhangs zwischen Neutralisation und Verhalten vermitteln die bisherigen Studien dennoch einen interessanten Eindruck von möglichen Moderatoren der Beziehung zwischen beiden Variablen. Unter der Annahme einer kausalen Wirkung des Neutralisationsmechanismus auf normwidriges Verhalten können diese Einblicke hypothesenbildenden Charakter haben. Im Folgenden sind die untersuchten potentiellen Moderatorvariablen aufgeführt.

*Norminternalisierung / Normakzeptanz.* Die Theorie der Neutralisation wurde von Sykes & Matza (1957) als Gegenentwurf zu einer Forschungstradition entwickelt, welche Norminternalisierung als den wichtigsten Prädiktor normkonformen bzw. –abweichenden Verhaltens betrachtet. Empirische Befunde zeigen jedoch einen gewichtigen Einfluss der Normakzeptanz auf Verhalten (z.B. Minor, 1984; Thurman, 1984; Amelang, Schahn & Kohlmann, 1988). Derlei Befunde können mithilfe von Minors Moderatorhypothese (1981) in die Neutralisationstheorie integriert werden. Hierbei sollte die Verhaltenswirksamkeit von Neutralisationen lediglich bei hoher Normakzeptanz bestehen. Die empirischen Befunde hierzu sind bislang uneindeutig, aber tendenziell entmutigend. Thurman (1984), Amelang, Schahn & Kohlmann (1988) sowie Dodder & Hughes (1993) finden Gegenteiliges, wonach die Neutralisationswirkung im Fall geringer Normakzeptanz stärker ist als bei hoher. Amelang, Zahn & Schahn (1988) sowie auch Schahn, Dinger & Bohner (1995) gelingt die Replikation eines solchen Effekts allerdings nicht.

*Schwere des Normbruchs.* Amelang, Schahn & Kohlmann (1988) berichten von einem höheren Neutralisations-Verhaltens-Zusammenhang bei vergleichsweise schweren als bei relativ leichten Vergehen. Diese beiden Deliktclassen wurden faktorenanalytisch gewonnen. Die Autoren erklären diesen Moderatoreffekt durch die unterschiedliche Stärke sozialer Missbilligung, die mit verschiedenen Verhaltensweisen verknüpft ist. „Die leichten Delikte werden von der Gesellschaft weniger stark mißbilligt, weshalb der Einsatz von Techniken der Neutralisation weniger notwendig ist“ (S.188). Da allerdings in keiner weiteren Studie ein solcher Effekt der Tatschwere berichtet wird, bedarf dieser Befund der Replikation.

*Häufigkeit vorangegangener Normverstöße.* Im Gegensatz zu den Ergebnissen seiner Gesamtstichprobe, war es Ball (1966) nicht möglich, innerhalb einer Teilstichprobe überführter Delinquenten anhand des Grades der Neutralisationsakzeptanz zwischen solchen mit vielen und solchen mit wenigen Straftaten zu unterscheiden. Dieses Ergebnis führte ihn zu der Vermutung, dass Neutralisationsakzeptanz eher im Anfangsstadium einer delinquenten „Karriere“ – also bei geringer Häufigkeit vorangegangener Normverstöße – als Verhaltensprädiktor von Bedeutung ist. Minor (1981) liefert mit seiner hardening-process-These eine zu diesem Ergebnis passende theoretische Erklärung. Leider gelingt es ihm selbst nicht, diesen Befund zu replizieren und da die Untersuchungen von Amelang, Schahn & Kohlmann (1988) sowie Hollinger (1991) sogar ein gegenteiliges Bild nahelegen, muss die Datenlage auch mit Bezug auf diese Moderatorhypothese als uneindeutig bewertet werden.

Die bisherige Forschung zur Verhaltensrelevanz von Neutralisationen lässt sich also hinsichtlich Umfang und Aussagekraft als recht eingeschränkt beurteilen. Zwar wird ein Zusammenhang zwischen potentiell neutralisierenden Kognitionen und Verhalten in allen Studien bestätigt, der Nachweis einer gerichteten Kausalität steht aber aus. Auch die Überlegungen und empirischen Untersuchungen zu möglichen Moderatoren dieses Zusammenhangs haben allenfalls hypothesenbildenden Charakter.

### **Der experimentelle Zugang**

Die einzige dem Autor bekannte experimentelle – und somit potentiell kausal interpretierbare - neutralisationstheoretische Studie ist eine von Norbert Schwarz & Andreas Bayer im Rahmen eines ZUMA<sup>1</sup>-Arbeitsberichts vorgelegte Arbeit aus dem Jahr 1989. Im Gegensatz zu allen übrigen oben erwähnten Untersuchungen zur Verhaltenswirksamkeit von Neutralisationen wurde die Neutralisationsvariable hier nicht in Form der Akzeptanz von neutralisierenden Aussagen operationalisiert, sondern als die kognitive Verfügbarkeit von Neutralisationen. Diese wurde insofern experimentell zufallsvariiert, als die Hälfte der N = 52 männlichen Studenten zunächst einen Teil des bereits bei Amelang, Schahn & Kohlmann (1988) eingesetzten Szenariofragebogens zur Neutralisationsakzeptanz bearbeitete (Szenarien zu Ladendiebstahl, Schwarzfahren, Wechselgeldmitnahme) und die Studenten anschließend nach der Wahrscheinlichkeit befragt wurden, mit der sie selbst die in den Szenarien beschriebenen Delikte begehen würden. Dem anderen Teil der Stichprobe wurde zuerst die Frage nach der eigenen zukünftigen Deliktbereitschaft gestellt und dann erst das Neutralisationsinventar zur Bearbeitung vorgelegt.

Diese Variation der Fragenreihenfolge innerhalb eines Fragebogens führte tatsächlich zu einer unterschiedlich hohen mittleren Verhaltensbereitschaft. Gemäß der Neutralisationstheorie gaben diejenigen Versuchsteilnehmer höhere Bereitschaftswerte an, die zunächst die Neutralisationsszenarios bearbeitet hatten (siehe Tab. 1).

Tabelle 1: Mittelwerte der berichteten Deliktbereitschaft als Funktion der Verfügbarkeit von Neutralisationstechniken; Anmerkung: 1 = sehr unwahrscheinlich; 10 sehr wahrscheinlich (aus Schwarz & Bayer, 1989, S. 23)

Delikt	Neutralisationstechniken	
	aktiviert	nicht aktiviert
Ladendiebstahl	1,80	1,26
Schwarzfahren	3,53	2,46
Wechselgeldmitnahme	5,34	2,96
Mittelwert	3,56	2,23

Methodisch wählen Schwarz & Beyer (1989) den Königsweg der Kausalitätsprüfung. Prädiktorvariablen werden hier nicht nur erhoben, sondern Personen werden zufällig einer bestimmten Ausprägung zugeordnet. Hierbei ist es zwar nicht möglich, „Aussagen über die natürliche Größe des untersuchten Phänomens zu machen“ (S. 15), jedoch kann auf diese Weise bestimmt werden, ob Neutralisationen bei entsprechender Verfügbarkeit einen Einfluss auf das Verhalten ausüben oder ob nicht.

Leider werden die Vorteile des experimentellen Designs hier von einigen Schwächen überschattet, die die Interpretierbarkeit der Ergebnisse möglicherweise einschränken.

a) Zunächst handelt es sich bei der experimentell manipulierten Variablen nicht um die Akzeptanz von Neutralisationen, wie in den vorher beschriebenen Untersuchungen, sondern lediglich um deren Verfügbarkeit. Dies stellt nicht unbedingt eine Einschränkung der Interpretierbarkeit der Ergebnisse dar, sondern erweitert deren Interpretation. Schwarz & Beyer (1989) spezifizieren – ohne dies allerdings kenntlich zu machen – neutralisationstheoretische Annahmen in einer theoretisch hoch relevanten Weise, welche explizit erwähnt werden muss. Demnach sollte es für die Wirksamkeit von Neutralisationen auf Verhalten nicht notwendig sein, eine Neutralisation inhaltlich zu bejahen, sondern es vielmehr ausreichen, diese im aktuellen kognitiven „Repertoire“ zu führen. Diese Spezifizierung der neutralisationstheoretischen Annahmen ist aus einer diskursanalytischen Perspektive sinnvoll, innerhalb derer Neutralisationsargumente als sozial repräsentierte „kognitive Ressource“ (Lauströer, 1996) verstanden werden können, die einem Akteur durch seine Kultur zur Verfügung gestellt wird. Da diese Argumente im Neutralisierungskontext eine strategische Funktion besitzen, kann sich eine neutralisierende Person „bedienen“, ohne die Neutralisation selbst inhaltlich vollständig akzeptieren zu müssen.

b) Eingeschränkt wird die Ergebnisinterpretation allerdings durch die Beschaffenheit der abhängigen Variablen. Bei dieser handelt es sich um die Einschätzung der Wahrscheinlichkeit eigenen künftigen Verhaltens. Dies ist allerdings weder gleichbedeutend mit tatsächlich ausgeführtem Verhalten noch mit einer Verhaltensintention. Es handelt sich vielmehr um eine rein kognitive und relativ verhaltensferne Variable, die tatsächlich folgendes Verhalten nur ungenügend widerspiegeln dürfte.

c) Der letzte Kritikpunkt betrifft die experimentelle Manipulation der Neutralisationsverfügbarkeit. Diese ist mit der Vorgabe eines normwidrigen Modellverhaltens konfundiert, da die der eigenen Verhaltensprognose vorangegangene Verfügbarmachung der Neutralisationen in einer der beiden Gruppen mit der Schilderung der normwidrigen Verhaltensweisen einherging, könnte sich diese Schilderung fördernd auf die Verfügbarkeit von Erinnerungen an zurückliegendes eigenes entsprechendes Fehlverhalten ausgewirkt haben. Eine derartige Referenzinformation sollte für sich genommen bereits zu einer erhöhten Prognose künftigen Fehlverhaltens führen.

## **Die Dosen-Studie**

Ziel der zu beschreibenden eigenen Untersuchung war es,

- die neutralisationstheoretisch angenommene Verhaltenswirksamkeit von Neutralisationen experimentell nachzuweisen sowie
- den Einfluss folgender potentieller Moderatoren auf diesen Effekt zu untersuchen:
  - Individuelle Normakzeptanz;

- Schwere des Normbruchs;
- Häufigkeit vorangegangener Normverstöße.

Methodisch hebt sich die Studie hierbei von der vorangegangenen Forschung neben der Untersuchungsmethode vor allem durch die der Neutralisation vorangehende Erhebung der individuellen Normakzeptanz sowie durch den Einsatz der Verhaltensbeobachtung ab.

Im Anschluss an die Forschung von Schahn et al. (1995) bezieht sich die Untersuchung auf den Bereich umweltschädigenden Verhaltens. Dieser Verhaltensbereich ist zwar in der Regel nicht gesetzlich sanktioniert, unterliegt jedoch einer starken gesellschaftlichen pro-Umwelt-Norm, welche sich in den Ergebnissen entsprechender Einstellungssurveys niederschlägt (z.B. Kuckartz, 2000).

Im Folgenden berichte ich über zwei ähnliche und aufeinander aufbauende Experimente.

## Experiment I

Im ersten Experiment sollte zunächst lediglich die Belegbarkeit des Neutralisations-Verhaltenseffekts mittels der eingesetzten Methode untersucht werden. Dies erschien aufgrund der Neuartigkeit des entwickelten Untersuchungsparadigmas für den Gegenstand notwendig.

In Tabelle 2 ist der formale Aufbau des Experiments dargestellt. Die Spalten beschreiben von links nach rechts den Ablauf des Experiments für die drei unabhängigen Versuchsgruppen.

Tabelle 2: Formale Struktur Experiment I

	FB Umwelt- einstellung (= Norm- akzeptanz)	Neutralisation vorgegeben <i>(UV<sub>alt 1</sub>)</i>	Neutralisation von Vpn selbst produziert <i>(UV<sub>alt 2</sub>)</i>	Normwidriges Verhalten <i>(AV)</i>	FB Gewohnheit
N=75					
Kontrolle <i>n<sub>K</sub></i> =26	FB I	nein	nein	Beobachtet	FB II
Treatment Verfügbarkeit <i>n<sub>TV</sub></i> =26	FB I	ja	nein	Beobachtet	FB II
Treatment Produktion <i>n<sub>TP</sub></i> =23	FB I	nein	ja	Beobachtet	FB II

N=75 per Aushang an der Universität Magdeburg geworbene Personen wurden zu einer „Untersuchung zur Kommunikation im Internet“ an das Institut eingeladen. Die zu 91,9 % aus StudentInnen der Universität sowie der Fachhochschule Magdeburg (48,0 % Studierende der Psychologie) bestehende Stichprobe weist mit 74,7 % (*n* = 56) einen klaren Frauenüberschuss auf (25,3 %, also 19 Männer). Das Durchschnittsalter lag mit 22 Jahren (min = 16, max = 33) erwartungsgemäß niedrig, das Bildungsniveau bei nur 4 Personen ohne Abitur (5,3 %) erwartungsgemäß hoch.

Vor Ort wurden die TeilnehmerInnen von der Versuchsleiterin / dem Versuchsleiter (VI) darüber unterrichtet, dass Ihre Aufgabe darin bestünde, mit einer anderen Versuchsperson, die bereits etwas früher eingetroffen sei und nun in einem Raum im anderen Stockwerk sitze, zu „chatten“, d.h., sich über zwei vernetzte Computer zu unterhalten, wobei die Untersuchung

mit der Beantwortung eines Fragebogens zu persönlichen Meinungen und Verhaltensweisen beginnen sollte. In diesen Fragebogen, der in weiten Teilen lediglich der Glaubhaftmachung der weiteren Deckgeschichte diente, war eine standardisierte Skala zur Messung der allgemeinen Umwelteinstellung (Schahn, Damian, Schurig & Fuchsle, 2000) integriert. Der resultierende Wert kann zwischen den Werten 1 (geringes Umweltbewusstsein) und 7 (hohes Umweltbewusstsein) variieren und gilt hier als ein Indikator der persönlichen Normakzeptanz. Anschließend an die Erhebung der allgemeinen Umwelteinstellung wurde zum Zweck der ergänzenden Induktion persönlicher Normakzeptanz das Item „Ich befürworte Umweltschutz.“ in eine Reihe andere Themen betreffender Aussagen gestellt. Dieses Item sollte auf einer vierstufigen Skala von 1 = „trifft überhaupt nicht zu“ bis 4 = „trifft völlig zu“ bewertet werden, wobei die beiden Skalenpunkte unterhalb des Mittelpunkts rot, die Punkte oberhalb des Skalenmittelpunkts grün unterlegt waren.

Anschließend erhielten die Versuchspersonen (Vpn) eine ausführliche Instruktion zum Chat mit der anderen (konföderierten!) Versuchsperson (Vp). Darin wurde als Ziel der Unterhaltung das Kennenlernen der anderen Person genannt, welches über das gegenseitige Befragen zu vorher ausgelosten Kombinationen von Themen aus dem öffentlichen Leben sowie „emotionalen Zuständen“ geschehen sollte. Diese Unterhaltung wurde strikten Regeln unterworfen, die beispielsweise vorschrieben, dass jede Frage erstens von einem realen Erlebnis handeln sollte und zweitens unbedingt zu beantworten sei. Beide Chat-PartnerInnen sollten sich abwechselnd in 15-Minuten-Blöcken zu ihren jeweiligen Themenkombinationen befragen. Da jede/r PartnerIn zwei solcher Themenkombinationen ziehen sollte, war der Gesamtzeitraum also auf eine Stunde angesetzt, die sich in vier Gesprächsabschnitte untergliedern sollte.

Der Instruktion folgte eine scheinbar zufällige Themenauswahl, bei der unter anderen die Themen „Umweltschutz“ und „schlechtes Gewissen“ zur Vergabe standen.<sup>2</sup> Diese waren ohne das Wissen der Vpn im Vorhinein der Auswahl entzogen, was die Ziehung dieser Themen durch die konföderierte Person plausibel machen sollte.

Zu Beginn des folgenden Chats bat der/die VI die Vp, mit der Befragung der anderen Person zu beginnen. Mit diesem ersten Durchgang war beabsichtigt, der Vp ein tieferes Verständnis der aufgestellten Chat-Regeln dadurch zu ermöglichen, dass sie diese auf ihre eigene Themenkombination anwenden musste. Nach maximal 20 Minuten Dauer mahnte die konföderierte Vp einen regelgemäßen „Themenwechsel“ an.

Die zweite Gesprächsrunde diente der Herstellung von Saliens des (antizipierten) möglichen eigenen Normbruchs sowie der experimentellen Induktion der Neutralisationsvariable. Die Grundidee dieser Induktion bestand darin, der Vp die Möglichkeit zu geben, ZeugIn einer „Neutralisationsepisode“ (vgl. Schönbachs (1990) Konzept der account episodes) auf Seiten der konföderierten Vp zu werden und je nach Versuchsbedingung diese auch beeinflussen zu können. Da sich diese simulierte Neutralisationsepisode auf eine Verhaltensmöglichkeit bezog, die sich auch der Vp selbst anschließend eröffnen würde, sollte die Episode – quasi als Projektion – die Funktion der eigenen Auseinandersetzung mit der Verhaltensoption erfüllen. Zur Simulation der Neutralisationsepisode besaß die konföderierte Vp eine vorgefertigte Auswahl an Textbausteinen, die nach dem in Tab. 3 dargestellten Versuchsplan eingesetzt wurden.

Tabelle 3: Verteilung von Textbausteinen und auf die Versuchsgruppen zur experimentellen Manipulation der Neutralisation (inkl. Instruktionen für die konföderierte Vp)

Kontrollgruppe	Treatment Verfügbarkeit	Treatment Produktion
also dann ... [die Sekretärin / der Untersuchungs-Chef] hat mir eben was zu trinken hingestellt :) Ein Tablett mit Pfandflaschen und Dosen. Und ich trinke jetzt Cola aus einer Dose. Stell Dir vor - ich haab tatsächlich ein schlechtes Gewissen, denn diese Dosen sind ja schlecht für die Umwelt.		
wann hattest Du das letzte Mal ein schlechtes Gewissen?	wann hattest Du das letzte Mal ein schlechtes Gewissen?	Ich suche nach Rechtfertigungen. wie würdest Du Dich in meiner Situation rechtfertigen?
<i>Antwort der Vp</i>	<i>Antwort der Vp</i>	<i>Antwort der Vp</i>
<p><i>[keine wertenden Aussagen zum schlechten Gewissen des anderen !]</i></p> <p>ich finde es ganz schön schwierig, zu der Themenkombi Fragen zu stellen!</p> <p><i>oder</i></p> <p>Hast Du schon von 3-Liter-Autos gehört? Meine Eltern wollen sich einen Lupo kaufen. Dabei habe ich gehört, dass nur ein bestimmtes Modell wirklich drei Liter verbraucht. Weißt Du, was die andern Autokonzerne machen?</p> <p><i>oder</i></p> <p>e: Als ich Kind war, habe ich mal einen Film über Walfang gesehen und hatte ein schlechtes Gewissen. Kannst Du Dich auch an solche Filme erinnern?</p>	<p>Aber um nochmal auf mein schlechtes Gewissen und die Cola-Dose zurückzukommen: Eigentlich hab ja nicht ich mich falsch verhalten, sondern die, die mir die Dose angeboten haben. Man ist mit seinem Umweltschutz nämlich immer abhängig davon, was andere machen. Kennst Du das auch, wenn Du eigentlich umweltfreundlich sein willst, aber durch andere keine Möglichkeit dazu hast?</p> <p><i>oder</i></p> <p>Übrigens ist das mit den Dosen bei mir eine Ausnahme. Mineralwasser z.B. kaufe ich fast immer in Pfandflaschen. Schließlich kommts drauf an, dass die "Gesamtlinie" stimmt. In welchen Situationen bist Du eher umweltfreundlich und in welchen nicht?</p> <p><i>oder</i></p> <p>Nochmal zu meiner Dose: Solange man die Dose nicht in den Wald wirft, sondern in den Mülltrenneimer, wird die Dose recycelt. So kommt auch die Umwelt zu ihrem Recht. Wohin würdest Du die Dose nachher werfen?</p>	<p>[wenn keine oder zu wenige (&lt;2) Rechtfertigungen genannt wurden, nochmalige Nachfrage:]</p> <p>Also, stell Dir vor, Du wärst ich und hättest diese Dose genommen, obwohl das nicht gerade Umweltschutz ist. Wie würdest Du Dein schlechtes Gewissen beruhigen?</p> <p>[notfalls nochmalige Nachfrage:]</p> <p>Also, ich will einfach nur Vorschläge, wie ich mein schlechtes Gewissen wegen der Dose wegbekomme. In den Regeln steht was von Antwortpflicht ... :(</p>
<i>Antwort der Vp</i>	<i>Antwort der Vp</i>	<i>Antwort der Vp</i>
Fortsetzung wie oben bis Gesprächszeitende	Fortsetzung bis mindestens zwei Neutralisationen vorgegeben wurden; dann bis Gesprächszeitende	Fortsetzung bis mindestens zwei Neutralisationen produziert wurden; dann bis Gesprächszeitende

Nach Beendigung des zweiten Gesprächsdurchgangs war die Vp an der Reihe, zu ihrer zweiten Themenkombination zu fragen. Bei Eintreffen ihrer ersten Frage zu der neuen



Themenkombination auf dem Rechner der konföderierten Vp, wurde ein/e für die Experimentalbedingung blinde/r GetränkebotIn (Gb) aktiviert, welche/r mit einer Entschuldigung für das verspätete Eintreffen („Entschuldigung, ich hab Dich ganz vergessen.“) das Untersuchungslabor betrat. Die/der Gb trug ein Tablett, auf welchem sich Cola, Orangenbrause und Mineralwasser befanden. Jede Getränkesorte gab es sowohl in mehreren Dosen als auch in teilweise angebrochenen 1,5 l-Pfandflaschen. Im Arbeitszimmer der Vp standen von Beginn an mehrere saubere Gläser. Seitens des/der Gb wurde die Vp gefragt, ob sie etwas zu trinken möge und ihr das gefüllte Tablett hingehalten. Für den Fall, dass die Vp grundsätzlich ablehnte, wies die/der Gb darauf hin, dass aber jeder Versuchsperson ein Getränk zustehe. Bei einem weiterhin abschlägigen Bescheid war die/der Gb instruiert, das Tablett unter dem Vorwand, die nächsten Vpn kämen eh erst am nächsten Tag, im Raum stehen zu lassen. Nach Verlassen des Raums notierte der/die Gb auf einem Fragebogen die Verhaltensentscheidung der Vp (Dose, Flasche, keines von beidem) und schätzte auf einer siebenstufigen Ratingskala ein, wie lange die Person bei der Auswahl des Getränks gezögert hatte (von 1 = „gar nicht gezögert“ bis 7 = „sehr lange gezögert“). Des Weiteren wurden von ihr registrierte Besonderheiten und Rückfragen seitens der Vp festgehalten.

Nachdem der dritte Gesprächsdurchgang beendet war, wurde der Chat vorgeblich aus Zeitgründen durch den VI abgebrochen und die Vp wurde über einen Standardtext sowie im persönlichen Gespräch mit dem VI über die wahren Hintergründe der Untersuchung aufgeklärt und anschließend gebeten, einen letzten Fragebogen auszufüllen. Innerhalb dieses Schlussfragebogens sollte mittels entsprechender Nachfragen die Wirksamkeit der Cover-Story sowie die Häufigkeit gewöhnlicher Dosennutzung erhoben werden. Die gewöhnliche Dosennutzung wird hierbei als eine Operationalisierung der Häufigkeit vergangener Normverstöße aufgefasst und wurde mittels eines Items erhoben: „Wie häufig trinken Sie normalerweise Getränke aus Dosen?“ (Antwortskala: nie (1), selten (2), manchmal (3), oft (4), immer wenn möglich (5). Des Weiteren wurden über zwei Items mit offenem Antwortformat potentielle Rationalisierungen der Dosennutzung im Nachhinein erfragt.

Am Ende der Untersuchung erhielten die TeilnehmerInnen ein im Aushang angekündigtes Honorar in Höhe von DM 15,-, welches von Studierenden der Psychologie wahlweise in eine Bescheinigung über Versuchspersonenstunden umgewandelt werden konnte.

Der Vorbereitung von Experiment I dienten zwei Vorstudien. Diese können auf Nachfrage dargestellt werden. Kern der beiden Untersuchungen war es, die Typizität der Verhaltensweise „Nutzung von Getränke-Einwegdosen“ als umweltschädigend zu zeigen sowie die Breite und Häufigkeit zu erwartender Neutralisierungen zu bestimmen. Beide Ziele wurden erreicht. Ein Ergebnis dieser Vorstudien sind die drei in der ersten Experimentalgruppe (Treatment Verfügbarkeit) der konföderierten Vp zur Verfügung stehenden Neutralisationen, welche den in einer offenen Szenarioaufgabe am häufigsten kodierten Neutralisationstechniken „Ablehnung der Verantwortung“, „Ablehnung des Unrechts“ sowie „Metapher des Hauptbuchs“ (zur Definition der Strategien siehe auch Fritsche & Mayrhofer, 2001) entsprechen.

## **Ergebnisse Experiment I**

Ziel des ersten Experiments war es, die kausale Wirkung von experimentell induzierter Neutralisation auf Verhalten empirisch zu zeigen. In der Gruppe verfügbarer oder aktiv produzierter Neutralisation sollte der Anteil der normwidrigen Dosennutzung jeweils höher

liegen als in der Kontrollgruppe. Darüber hinaus war es mithilfe der erhobenen Daten möglich, die Hypothese des moderierenden Einflusses der Normakzeptanz auf den postulierten Neutralisationseffekt zu testen.

Die Anzahl der Vpn reduzierte sich für die Auswertungen auf  $N_{\text{kor}} = 73$ , da zwei Personen aus den nachfolgenden Analysen ausgeschlossen wurden. Dies geschah im Fall einer Person, die weder eine Dose noch eine Flasche genommen hatte und bei einer weiteren Person, die von einem schwerwiegenden technischen Fehler der Computeranlage betroffen war, der zur Verzögerung und vorzeitigem Abbruch der Chat-Phase führte.

Abb. 2 zeigt den relativen Anteil der Dosennutzung innerhalb der drei Versuchsbedingungen. In offensichtlicher Weise unterscheiden sich die beiden Experimentalbedingungen hierbei erwartungswidrig nicht voneinander.

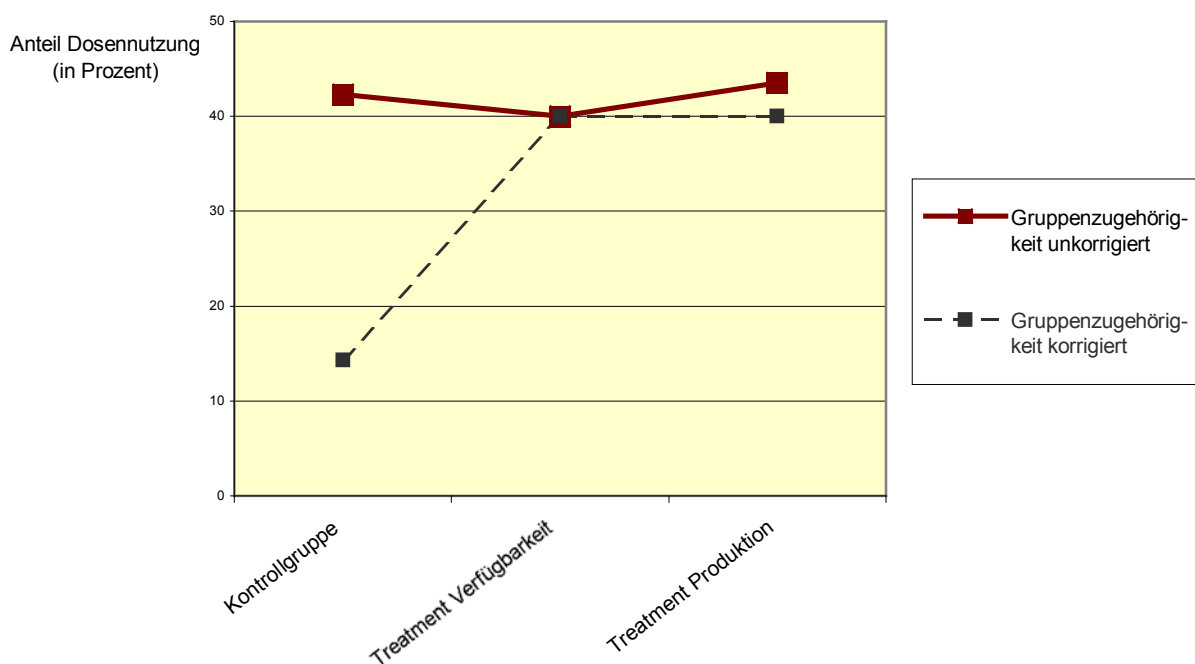


Abbildung 2: Relativer Anteil der Dosennutzung innerhalb der Versuchsbedingungen ( $N = 73$ )

Dieses ernüchternde Ergebnis führte zu post-hoc-Analysen, mit deren Hilfe Mutmaßungen über die Ursache des Nichtzustandekommen des Neutralisationseffekts angestellt wurden. Eine dieser nachträglichen Auswertungen bestand im Rekodieren der Gruppenzugehörigkeit. Hierbei wurden zunächst die Texte aller Chats einer quantitativen Inhaltsanalyse zum Vorkommen von Neutralisationsäußerungen auf Seiten der Vpn unterzogen. Anschließend fielen jene Vpn der Kontrollgruppe aus dieser heraus, die im Chat spontan Neutralisationen generiert hatten, sowie solche Personen aus der Experimentalgruppe „Produktion“, die trotz entsprechender Instruktion keine Neutralisationen genannt hatten. Die Probandenzahl für diese Experimentalgruppe reduzierte sich in Folge der Selektion auf  $n_{\text{Tpkor}} = 17$  und die der Kontrollgruppe gar auf  $n_{\text{Kkor}} = 7$ . Die relative Dosennutzung nach Korrektur der Gruppenzugehörigkeit ist ebenfalls in Abb. 2 dargestellt und zeigt – deskriptivstatistisch – eine Anteilsverteilung über die Gruppen, welche der Erwartung eines Neutralisationseffekts auf die Dosennutzung entspricht. Dieses aufgrund der geringen Größe der Kontrollgruppe und des faktischen Rückwurfs auf ein quasiexperimentelles Untersuchungsniveau natürlich mit äußerster Vorsicht zu interpretierende Ergebnis kann als Hinweis auf eine nicht ausreichend

starke experimentelle Manipulation aufgefasst werden. Schließlich hatten Mitglieder der Kontrollgruppe offensichtlich die Möglichkeit, spontan Neutralisationen zu generieren, und gleichwohl konnten sich jene Personen der zweiten Experimentalgruppe dem vorgesehenen Treatment entziehen.

Des Weiteren wurde die von der/dem GB unmittelbar nach dem Anbieten der Getränke eingeschätzte Dauer des Zögerns der Vp bei der Auswahl eines Getränks als weitere Abhängige Variable in für jede der beiden Experimentalgruppen getrennt durchgeführte Univariate Varianzanalysen (ANOVA) aufgenommen. Unabhängige Variablen waren die Neutralisationsbedingung sowie die künstlich dichotomisierten Variablen „Allgemeine Umwelteinstellung“ (Normakzeptanz, Mediansplit bei 5,17) und „Häufigkeit sonstiger Dosennutzung“ (Häufigkeit vorangegangener Normverstöße, Mediansplit bei 2). Neben einem Haupteffekt der letztgenannten personalen Variablen in beiden Analysen ( $F=5,93$ ,  $df=1$ ,  $p<.05$  bzw.  $F=6,01$ ,  $df=1$ ;  $p<.05$ ) lässt sich in der ersten ANOVA (Vergleich des mittleren Zögerns zwischen Kontrollgruppe und der Experimentalgruppe „Treatment Verfügbarkeit“) ein weiterer Haupteffekt der Normakzeptanz ( $F=10,22$ ,  $df=1$ ;  $p<.01$ ) sowie eine Interaktion der Häufigkeit vergangener Normverstöße mit der Versuchsgruppenzugehörigkeit ( $F=6,93$ ,  $df=1$ ;  $p<.05$ ) finden. Die zweite ANOVA (Vergleich des mittleren Zögerns zwischen Kontrollgruppe und der Experimentalgruppe „Treatment Produktion“) zeigt neben der bereits erwähnten Bedeutung der Häufigkeitsvariablen auch einen Haupteffekt der Zugehörigkeit zur Versuchsgruppe ( $F=6,03$ ,  $df=1$ ,  $p<.05$ ) in Richtung eines längeren Zögerns auf Seiten der Kontrollgruppe. Dieser Effekt scheint hier jedoch zusätzlich von den zwei personalen Variablen moderiert zu sein, da sowohl die Häufigkeit vergangener Normverstöße ( $F=10,19$ ,  $df=1$ ,  $p<.05$ ), wie tendenziell auch die Normakzeptanz ( $F=7,03$ ,  $df=1$ ,  $p=.06$ ) mit der Gruppenzugehörigkeit in Zweifach-Interaktion treten. Beide Interaktionseffekte sind in Abb. 3 und Abb. 4 dargestellt.

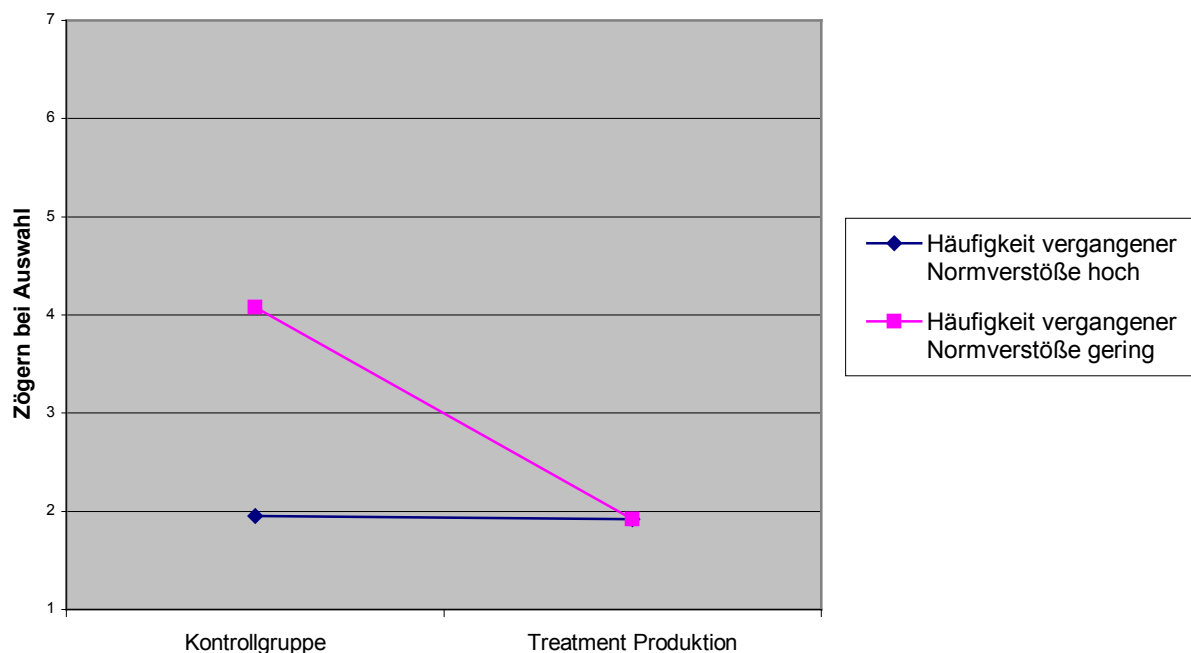


Abbildung 3: Vergleich der Dauer des Zögerns bei der Verpackungswahl – Interaktion „Häufigkeit vergangener Normverstöße“ x „Versuchsgruppenzugehörigkeit“

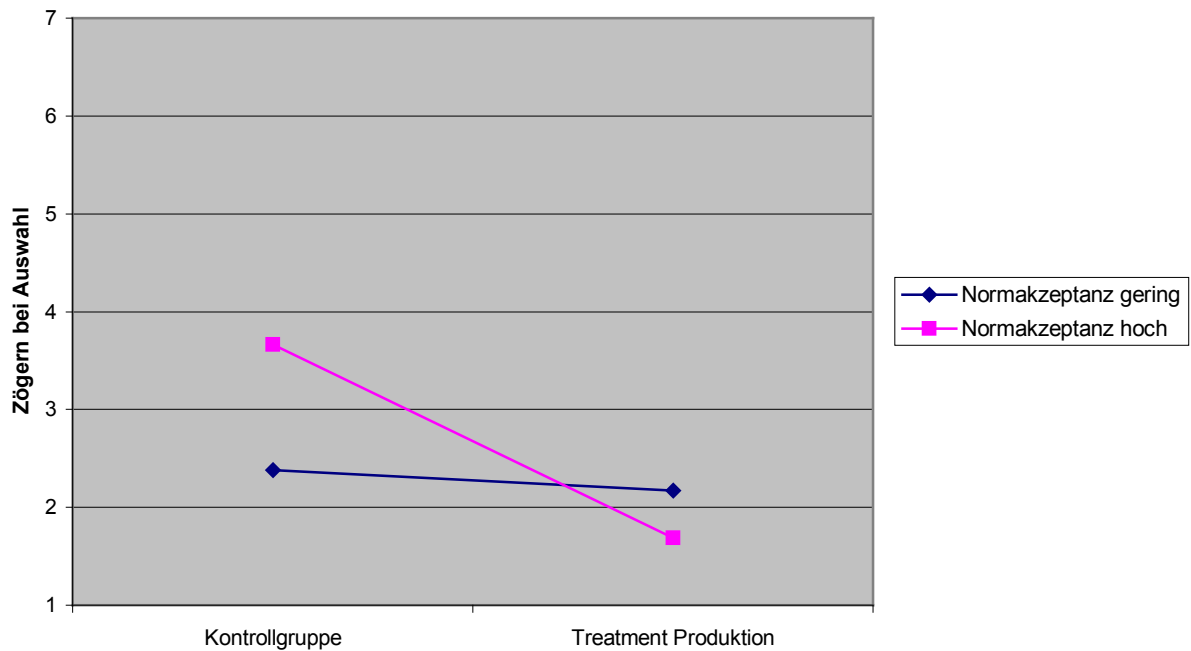


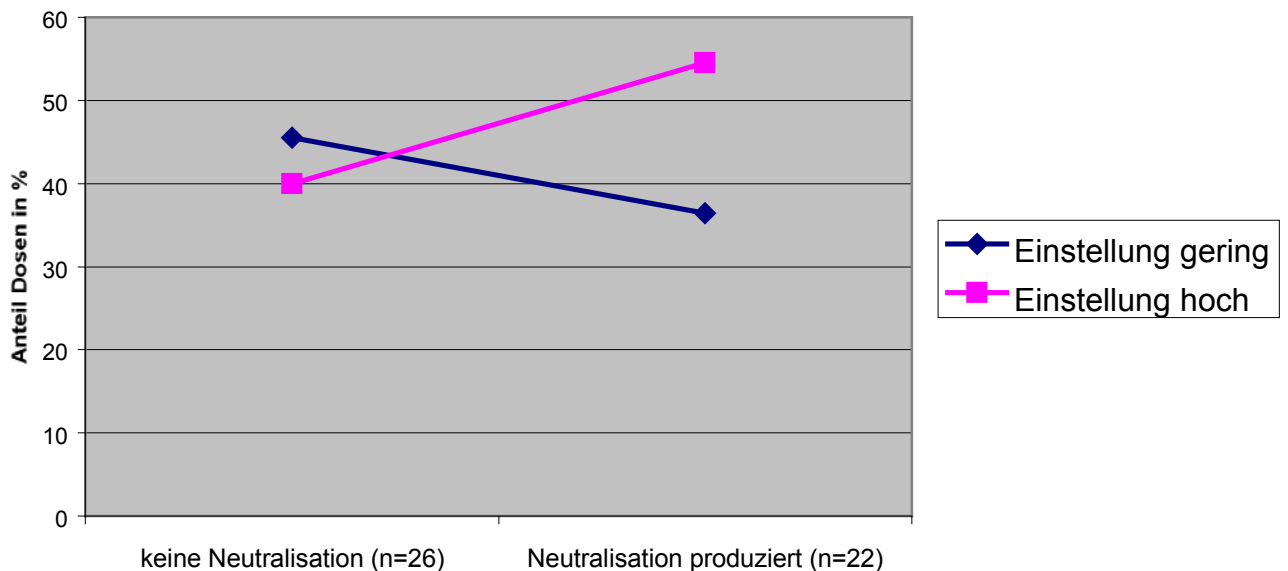
Abbildung 4: Vergleich der Dauer des Zögerns bei der Verpackungswahl – Interaktion „Normakzeptanz“ x „Versuchsgruppenzugehörigkeit“

Die Betrachtung der dargestellten Effekte der Versuchsgruppenzugehörigkeit auf eine indirekte Abhängige Variable wie die eingeschätzte Dauer des Zögerns bei der Verhaltensentscheidung führt zu dem Verdacht, dass die experimentelle Manipulation durchaus in der Lage war, Verhalten der Vpn in eine theoretisch plausible Richtung zu verändern. Die Ermöglichung der Neutralisation führt – zumindest auf Seiten von Personen, die die potentiell gebrochene Norm Umweltschutz stark akzeptieren – zu einer schnelleren Verhaltensentscheidung. Dieses Ergebnis ist auch konsistent mit der implizit neutralisationsbezogenen Annahme, die im Normaktivationsmodell von Shalom Schwartz (1977 sowie Schwartz & Howard, 1981) enthalten ist. Rechenschaftslegung wird dort nicht als Neutralisation, sondern als „Defense“ (Abwehr) konzeptualisiert und gilt als vereindeutigendes Element einer rationalen Entscheidung über sozial-normativ relevantes Verhalten, welches eine Verhaltensentscheidung bei unentschiedenen Abwägungen katalysiert. Dies geschieht durch die nachträgliche Umbewertung („Rekalibrierung“) vorangegangener Kognitionen. Bei Personen, denen diese Rekalibrierung bereits nahegelegt wurde, sollte also schon eine Bahnung zu einer eindeutigen Verhaltensentscheidung vorliegen. Dies kann die weniger zögerliche Verhaltensaufführung der Personen in der Experimentalgruppe „Treatment Produktion“ erklären.

Neben der Testung der Neutralisations-Verhaltens-Hypothese ging es auch darum, Moderatoren des postulierten Kausalzusammenhangs zu identifizieren. Von den oben angeführten möglichen Moderatorvariablen kann an dieser Stelle allerdings lediglich die individuelle Normakzeptanz untersucht werden, da die Moderatoren Schwere des Normbruchs sowie Häufigkeit vergangener Normverstöße erst innerhalb des zweiten Experiments angemessen erfasst wurden.

Alleiniger Indikator der individuellen Normakzeptanz war im ersten Experiment die allgemeine Umwelteinstellung, welche über eine aktuelle 6-Item-Skala zur Umwelteinstellung (Schahn et al., 2000) zu Beginn der Untersuchung erfasst wurde. Es wurden zwei für beide Experimentalgruppen separate logistische Regressionsanalysen (SPSS-Prozedur LOGISTIC

REGRESSION) über die Prädiktoren Kontroll- vs. Experimentalgruppe, Einstellung sowie die Interaktion beider Variablen bezogen auf das Kriterium der Dosennutzung durchgeführt. Ein statistisch bedeutsamer Interaktionseffekt – ein Hinweis auf die angenommene Moderatorfunktion der Normakzeptanz – trat jedoch in keiner der beiden Analysen auf ( $B = -.875$ ,  $p = .25$ , Nagelkerkes  $R^2 = .18$  bzw.  $B = .777$ ,  $p = .29$ , Nagelkerkes  $R^2 = .03$ ). Betrachtet man die Dosennutzung in den gekreuzten Gruppen allerdings auf deskriptiver Ebene, so fällt eine entgegengesetzte Richtung des Zusammenwirkens von Neutralisation und Normakzeptanz ins Auge (Abb. 5). Wurden Neutralisationen vorgegeben, so ist – vergleichbar mit den Befunden von Thurman (1984), Amelang, Schahn & Kohlmann (1988) sowie Dodder & Hughes (1993) – ein stärkerer Neutralisationseffekt für die Personen mit eher geringer Normakzeptanz zu erahnen, während bei Personen mit eher hoher Normakzeptanz die Verfügbarkeit von Neutralisationen sogar verhaltenshemmend zu wirken scheint. Dieses Verhältnis kehrt sich um, wenn die Personen der Kontrollgruppe mit denen verglichen werden, die aufgefordert worden waren, ihre eigenen Neutralisationen zu produzieren. Hier deutet sich ein Neutralisationseffekt für die Personen mit hoher und eine Hemmung auf Seiten der Personen mit geringer Normakzeptanz an.



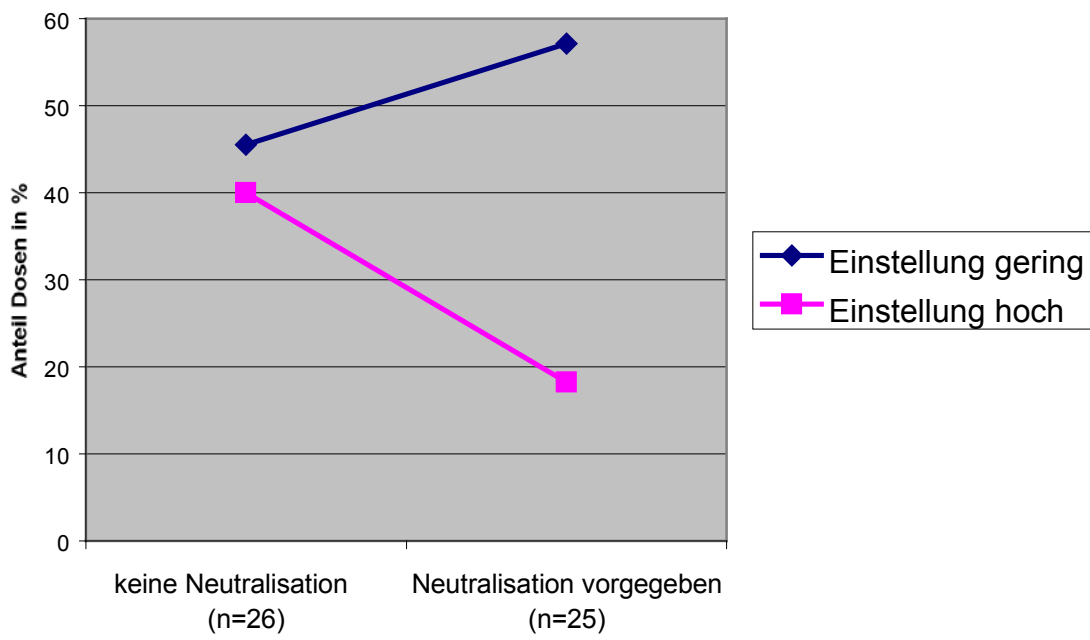


Abb. 5: Relativer Anteil der Dosennutzung in Abhängigkeit von Neutralisation und allgemeiner Umwelteinstellung; separate Vergleiche zwischen Kontrollgruppe und beiden Experimentalgruppen (Interaktionseffekte statistisch nicht signifikant!).

## Diskussion Experiment I

Die Ergebnisse des ersten Experiments bestätigen die Hauptannahme der Neutralisationstheorie – die Verhaltenswirksamkeit von Neutralisationen – nicht. Bevor dieses Resultat jedoch auf eine eventuelle Fehlerhaftigkeit der Theorie bezogen wird, sollte in Erwägung gezogen werden, dass die Operationalisierung der Hypothese im Experiment ihrerseits fehlerhaft gewesen sein kann. Allerdings macht die Neutralisationstheorie hinsichtlich der Antezedenzen, unter denen Neutralisation auf Verhalten wirken soll, nur wenige, ungenaue und empirisch ungesicherte Angaben und ist insofern tatsächlich mit Argwohn zu betrachten. Daher empfiehlt es sich, für die Fehlersuche zunächst ein eigenes Arbeitsmodell der Wirkung von Neutralisation und deren Voraussetzungen zu entwerfen, welches der im Experiment zu modellierenden beabsichtigten Situation entspricht. Dieses Modell ist in Abb. 6 dargestellt. Es bildet einen Leitfaden bei der Suche nach potentiellen Fehlern bei der Ausgestaltung des Experiments. Da sich die Ergebnisse dieser Fehlersuche in der Durchführung von Experiment II direkt niedergeschlagen haben, wird sie gemeinsam mit dem Leitfaden zu Beginn des folgenden Kapitels dargestellt.

Die Untersuchung eines Moderatoreffekts der Variable „Normakzeptanz“, operationalisiert über die allgemeine Umwelteinstellung, ergab kein statistisch bedeutsames Ergebnis. Der Umstand unterschiedlich gerichteter Interaktionsgewichte in beiden Experimentalgruppen lässt allerdings Vermutungen über die Abhängigkeit der Richtung des Effekts von den gegebenen Untersuchungsbedingungen zu. Möglicherweise erzeugt die Konfrontation von stark normativ gebundenen Personen mit vorgegebenen Rechenschaftsaussagen Mißtrauen gegenüber der Redlichkeit einer solchermaßen begründeten Verhaltensentscheidung, was wiederum zu einem reaktanten Verhalten führt. Im Fall eigener Neutralisationen sollte ein solches Mißtrauen weit geringer ausgeprägt sein. Im Gegensatz zu den Personen mit starker Normakzeptanz sollten Personen mit geringer Normakzeptanz zugänglich gemachte Neutralisationen weit weniger kritisch beurteilen, da einerseits die Verwendung tatermöglichender Neutralisationen für diese weit weniger dissonanzerzeugend ist und andererseits aufgrund des vermutlich in der Regel geringeren Neutralisierungsbedarfs dieser Personen kein individuelles Repertoire an Neutralisationen besteht.

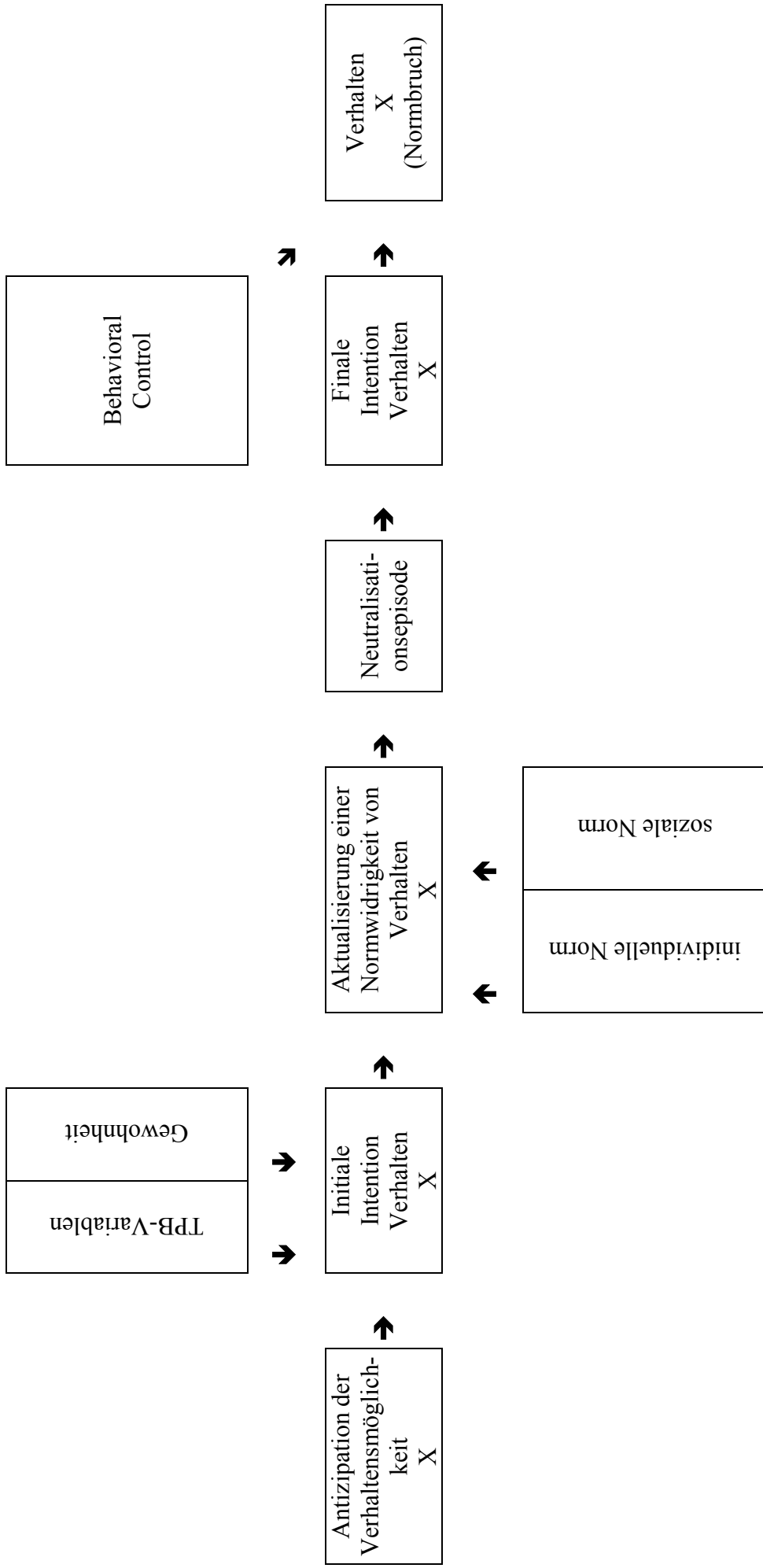


Abbildung 6: Arbeitsmodell zur situativen Wirkung von Neutralisation



## Experiment II

Im ersten Experiment war es nicht möglich, den Neutralisations-Verhaltens-Effekt nachzuweisen, was zu einer Revision des ursprünglichen Versuchsplans führte. Bei diesen Änderungen handelte es sich sowohl um die Induktion und/oder Kontrolle von möglichen Antezedenzvariablen des Neutralisationseffekts, als auch um die Verbesserung der technischen Umsetzung des Experiments (Vereindeutlichung der Treatments, Verbesserung des manipulation-check, etc.).

Induktion und Kontrolle von Antezedenzvariablen. Dem in Abb. 6 abgebildeten Prozessmodell ist als erste und distalste notwendige Bedingung für das Auftreten des Neutralisationseffekts die Antizipation der Verhaltensmöglichkeit zu entnehmen. Diese grundlegende Antezedenzannahme mag auf den ersten Blick banal wirken, ist jedoch in theoretischer wie auch untersuchungspraktischer Hinsicht durchaus von Belang. Theoretisch weist diese Annahme darauf hin, dass es im Vorfeld eines potentiell normwidrigen Verhaltens für eine Person Gelegenheit geben sollte, sich kognitiv mit einem Verhalten auseinanderzusetzen. Diese kognitive Auseinandersetzung setzt allerdings keine Bewusstheit voraus, sondern kann durchaus automatisiert ablaufen (vgl. Bargh, 1997). In Bezug auf die Optimierung der Untersuchungsdurchführung bedeutete diese Antezedenzannahme, dass gegenüber den Vpn stärker als im ersten Experiment geschehen auf ihre eigene spätere Möglichkeit der Dosennutzung hingewiesen werden sollte. Dies geschah durch die Aufnahme eines entsprechenden Hinweises in die Vorrede des Versuchsleiters zum Chat sowie durch die leichte Ergänzung des in allen Versuchsgruppen gleichen Einleitungsabschnitts im zweiten Chatabschnitt durch den Satz „Die/Der [GB] wollte auch noch zu Dir. War er/sie schon da?“.

Die Umsetzung der antizipierten Verhaltensmöglichkeit in Verhalten setzt in der Regel voraus, dass das Individuum eine entsprechende Verhaltensvornahme entwickelt hat. Diese initiale Intention sollte durch das Zusammenspiel von Erwartungs- und Wert-Überzeugungen entstehen, wie es in der Einstellungs-Verhaltens-Forschung beispielsweise von Ajzen & Madden (1986) in der Theory of Planned Behavior (TPB) beschrieben wird. Dem Arbeitsmodell dient diese Theorie als Vorlage, die jedoch erheblich erweitert wird. Neben der Hinzunahme der intentionsbeeinflussenden Variable „Gewohnheit“ (im zweiten Experiment mit einer präzise verankerten Skala erhoben) besteht die Erweiterung besonders in der Prozessualisierung der Annahmen zum Einstellungs-Verhaltens-Zusammenhang sowie der damit verbundenen Einführung einer Neutralisationsphase. Für die Suche nach bei der Ausgestaltung der Untersuchung vernachlässigten Bedingungen für den Neutralisationseffekt erwies sich das Konstrukt der Initialen Intention insofern als wertvoll, als die TeilnehmerInnen des ersten Experiments möglicherweise vor Beginn der simulierten Neutralisationsepisode noch nicht die Intention gefasst hatten, überhaupt etwas bzw. etwas aus einer Dose zu trinken. In einem solchen Fall bestünde für die Person überhaupt kein Anlass zu neutralisieren. Um dies im Nachfolgeexperiment weitgehend ausschließen zu können, wurde zu Beginn der Chat-Phase versucht, die initiale Intention zu induzieren. Hierzu wurden die Vpn vom Versuchsleiter nicht nur – wie bereits weiter oben berichtet – auf die ausstehende Getränkelieferung hingewiesen, sondern auch gefragt, ob sie „Cola, Fanta oder Wasser“ möchten. Die darauf folgende Bestellung sollte die Personen auf die Intention festlegen, ein Getränk zu nehmen. Zum Abschluss des Experiments wurde die Intentionsbildung rückwirkend innerhalb eines Kontrollfragebogens erfasst.

Im Anschluss an die Intentionsbildung kommt es im Arbeitsmodell zur Aktualisierung der Normwidrigkeit des intendierten Verhaltens, welche durch die Aktivierung entsprechender sozialer und/oder persönlicher Normen (i. S. von Schwartz, 1977) ausgelöst wird. Dieser

Schritt wurde der Übersichtlichkeit halber im Arbeitsmodell unabhängig von der Intentionbildung dargestellt. Alternativ zu dieser nachträglichen Aktualisierung der Normwidrigkeit ist die Aktualisierung auch im Verlauf der initialen Intentionbildung denkbar. Wichtig für das Arbeitsmodell ist es daher lediglich, dass nach der Bildung einer initialen Intention ein Normbruch ausreichender Stärke vom Individuum registriert wird. Dies ist entweder dann der Fall, wenn die Intentionbildung zugunsten der Verhaltensweise eine „knappe Sache“ war, also gegen die Ausführung des Verhaltens sprechende Überzeugungen von vergleichsweise hoher Wichtigkeit vorlagen und diese Überzeugungen zumindest teilweise normativer Natur waren. Andererseits können derartige normative „Gegenkognitionen“ auch nach Abschluss der Intentionbildung hinzukommen. Dieser Fall wird im Arbeitsmodell dargestellt. In Experiment I wurde die Aktualisierung der Normwidrigkeit durch einen entsprechenden Hinweis der konföderierten Vp induziert (s. o.). Zur Kontrolle, ob dies im Einzelfall gelungen ist, enthielt der Schlussfragebogen des zweiten Experiments mehrere entsprechende Nachfragen, z.B.: „Befanden sie sich während des letzten Fragenblocks in einem kleinen „Gewissenskonflikt“, ob Sie selbst die Dose nehmen sollten oder nicht?“. Zugleich wurden Kontrollfragen zur Ausprägung einer verhaltensspezifischen sozialen und persönlichen Norm formuliert („Nachdem Ihr Kommunikationspartner berichtet hatte, eine Getränkedose bekommen zu haben, *hatten Sie zu dem Zeitpunkt einmal den Gedanken, dass das Nehmen einer Dose gegen [gesellschaftliche / Ihre eigenen] Regeln verstößt?*“; Antworten von 1 = „ich hatte diesen Gedanken überhaupt nicht“ bis 6 = „ich hatte diesen Gedanken sehr stark“).

Wurde eine ausreichend starke Normwidrigkeit des intendierten Verhaltens festgestellt, die jedoch nicht zur Aufgabe dieser Intention führt, folgt der Versuch, diesen Normverstoß zu neutralisieren. Diese „Neutralisationsepisode“ kann man sich als innere Verhandlung eigener Schuld vorstellen. In den meisten Fällen wird es sich bei dieser Verhandlung um ein weitgehend automatisiertes Schnellverfahren handeln, in dem eine verfügbare und möglicherweise gewohnheitsmäßig eingesetzte Rechenschaftskognition nicht hinterfragt, sondern als hinreichend neutralisierend angenommen wird. Allerdings mag es andererseits auch zu einem ausführlichen Abwägen kommen, wenn beispielsweise der Normbruch als sehr gravierend angesehen wird oder ein ähnliches Verhalten vorher noch nie ausgeführt wurde. Müssen im Rahmen dieses Abwägens einzelne Neutralisationskognitionen als ungültig zurückgewiesen werden, so sollte die Person bemüht sein, gültige Neutralisationskognitionen zu finden. Ist dies nicht möglich, so wäre sie aus der hier dargestellten Perspektive nicht in der Lage, das ursprünglich intendierte Verhalten auszuführen. Dies sollte zur Revision der initialen Intention in Form einer finalen Intention sowie ebenfalls zur Nichtausführung des Verhaltens führen. Im zweiten Experiment wurde die Intention, das erwartete Getränk zu nehmen, vor und nach der simulierten Neutralisationsepisode rückwirkend erfragt. Die Neutralisationsepisode selbst wurde diesmal in weitaus intensiverer Weise manipuliert.

In Experiment I bestand die Manipulation der Neutralisationsvariable in einer Verfügbarkeits- sowie in einer Produktionsinstruktion. Die Ergebnisse einer nachträglichen Inhaltsanalyse der Chat-Äußerungen legen den Verdacht nahe, dass diese Manipulation aufgrund spontaner Neutralisation in der Kontrollgruppe zu schwach war. Aus diesem Grund sollte im zweiten Experiment das Auftreten spontaner Neutralisationen in der Gruppe der Nicht-NeutralisiererInnen verhindert bzw. aufgefangen werden. Gleichzeitig sollte die Kombination beider in Experiment I in getrennten Gruppen eingesetzten Maßnahmen zur Ermöglichung von Neutralisation zu einer Maximierung der Neutralisationsfähigkeit führen. So wurden unter Verzicht auf eine Kontrollgruppe zwei Experimentalgruppen mit gegensätzlichen Treatments eingerichtet. Den Vpn der ersten Gruppe sollte Neutralisation verunmöglicht, denen der zweiten Gruppe diese nahegelegt werden. Die Verunmöglichtung bestand zum einen

darin, dass Neutralisationskognitionen von der konföderierten Vp nicht angesprochen wurden, sondern die Betonung statt dessen auf ihrem schlechten Gewissen lag. Andererseits lag der konföderierten Vp bei spontaner Erwähnung von potentiellen Neutralisationen ein aus Vorstudien gewonnenes Manual mit Gegenargumentationen („Widerlegungen“) vor, aus dem bestimmten Neutralisationsstrategien zugeordnete Textbausteine in die eigenen Chatbeiträge einkopiert werden konnten. Zur Förderung von Neutralisation wurden die Treatments aus Experiment I insofern kombiniert, als die Vpn zur Äußerung von Rechtfertigungen aufgefordert und sie zugleich mit mindestens zwei der drei vorgegebenen Neutralisationskognitionen konfrontiert wurden. Tab. 4 gibt einen Überblick über die unterschiedlichen Instruktionen für die konföderierte Vp.

Tabelle 4: Instruktion konföderierte Vp in Experiment II

<b>Gruppe 1: Neutralisierung gehemmt</b>	<i>Gruppe 2: Neutralisierung gefördert</i>
<p><b>c:</b> also dann ... [die Sekretärin] ist eben mit nem Tablett voller Dosen rumgekommen. Und ich trinke jetzt Cola aus einer DOSE!! Hab deswegen aber ein ziemlich schlechtes Gewissen, denn diese Dosen sind ja schlecht für die Umwelt. Die wollte auch noch zu Dir ... war sie schon da?</p>	
<p>Meine Frage: kannst Du mein schlechtes Gewissen nachvollziehen?</p>	<p>Meine Frage: wie würdest Du Dich in meiner Situation vor Dir selbst rechtfertigen?</p>
<p><u>Reaktion der Vp</u> <u>meine Reaktion</u></p> <p>Aussage pro schlechtes Gewissen Bestätigung des schlechten Gewissens</p> <p>Neutralisation Spontane Widerlegung</p> <p>keines von beidem Bestätigung des schlechten Gewissens</p>	<p><u>Reaktion der Vp</u> <u>meine Reaktion (c)</u></p> <p>Nennung einer Neutralisation Bestätigung + Nennung anderer Neutralisation</p> <p>Ablehnen einer spezifischen Neutralisation kurze Gegenargumentation + Nennung anderer Neutralisation</p> <p>keine Neutralisation genannt Nennung einer Neutralisation</p>
<p><i>Wichtig: Nach der ersten Antwort wird die GetränkebotIn verständigt!</i></p>	

Die im Arbeitsmodell aufgeführte Verhaltenskontrolle als Voraussetzung von Verhalten kann als gegeben vorausgesetzt werden. Um sozialer Inhibition entgegenzuwirken, erhielt die/der GB im zweiten Experiment die Instruktion, das Tablett mit den Getränken eilig in den Raum zu tragen, es dort unter einem Vorwand stehenzulassen und es erst nach Ablauf der simulierten Neutralisationsepisode wieder abzuholen.

*Methodisch-technische Veränderungen.* Im zweiten Experiment kam es zu einigen Veränderungen der Versuchsbedingungen, welche zwar ebenfalls im Rahmen der Antezedenzvariablen erläutert werden könnten, die jedoch außerhalb des vorgestellten Arbeitsmodells liegen.

Den TeilnehmerInnen des ersten Experiments war vor der simulierten Neutralisation bekannt, dass sich auf dem Tablett neben als umweltschädigend deklarierten Dosen auch Pfandflaschen befanden. Es kann vermutet werden, dass ein nicht geringer Teil der Vpn den die Neutralisation überflüssig machenden Weg der Spezifikation der initialen Verhaltensintention gewählt hat. Statt der konfliktreichen Dosen ist es ein Leichtes, eine saubere Alternative zu wählen. Aus diesem Grund wurde die Verhaltensalternative erschwert. Dies geschah durch die Platzierung der Pfandflaschen im Vorflur des Versuchsraums, in dem sich jede Vp vor

Einlass ca. eine halbe Minute lang aufhielt. Auf dem Tablett fanden sich nunmehr nur noch Dosen.

Während in Experiment I die Entscheidung für oder gegen die Dose ca. 5 - 10 Minuten nach Beendigung der Neutralisationsepisode anstand, wurde dieser Zeitpunkt in Experiment II unmittelbar in die Neutralisationsepisode gelegt. Das Tablett wurde in den Raum der Vp gestellt, nachdem sie durch die konföderierte Vp eine Entgegnung auf Ihre erste Antwort erhalten hatte. Wieder abgeholt wurde es im Anschluss an einen von einer der beiden Parteien signalisierten Themenwechsel.

Dem in Tab. 4 dargestellten Instruktionsplan ist zu entnehmen, dass sich im zweiten Experiment im Gegensatz zum ersten alle Gruppen die gesamte Zeit über den Dosenkonflikt unterhielten. Diese Gleichsetzung schien im Nachhinein notwendig, um eine Erklärung des in Experiment I nicht vorhandenen Neutralisationseffekts durch eine unterschiedlich ausführliche Beschäftigung mit dem potentiellen Gewissenskonflikt ausschließen zu können.

Neben diesen ausführlich beschriebenen Änderungen und Ergänzungen der Versuchsbedingungen kam es auch zur Erhebung zahlreicher Kontrollvariablen, die sich unter anderem auf Merkmale der Versuchssituation und die Überprüfung des Erfolgs der experimentellen Manipulation bezogen. Diese Informationen wurden im Abschlussfragebogen erhoben, der im Gegensatz zu Experiment I bereits vor der Aufklärung der Vpn über den Hintergrund der Untersuchung ausgegeben wurde.

Die im Schlussfragebogen erhobene Variable „subjektive Schwere des Normbruchs“ sollte aufgrund ihrer Relevanz für die Testung einer der Moderatorhypothesen hier gesonderte Erwähnung finden. Bei der Erhebung dieser Variablen handelt es sich um den Versuch, die subjektive Repräsentation der Umweltschädlichkeit der Dosennutzung im Kontext anderer potentiell umweltschädigender Verhaltensweisen zu erfassen. Hierzu wurde zu insgesamt fünf Verhaltensweisen (Autonutzung auf Kurzstrecken, schwere Wasserverschmutzung, Fahrstuhlnutzung, Verwendung von Einweg-Getränkedosen, Nutzung von Kurzstreckenflügen, Verpackungsnutzung) gefragt, wie schädlich diese für die Umwelt seien (von 1 = „völlig unschädlich“ bis 5 = „sehr schädlich“). Die Formulierung des Items zur Dosennutzung lautete: „Ein Getränk aus einer Einweg-Dose trinken.“

Die Stichprobe des zweiten Experiments wurde wie in der Vorgängeruntersuchung per Aushang an der Universität und Ankündigung eines Honorars von DM 15,- oder zweier Versuchspersonenstunden geworben. Auch bei den TeilnehmerInnen des zweiten Experiments handelt es sich fast ausschließlich um Studierende der Magdeburger Universität und Fachhochschule (94,4 %), wobei der Anteil der Psychologiestudierenden (14,4 %) deutlich unter dem des ersten Experiments liegt. Ebenfalls fällt der Frauenüberschuss (58,2 %) in Experiment II geringer aus. Sehr ähnlich sind sich beide Stichproben in ihrem überdurchschnittlichen Bildungsgrad (niemand ohne Abitur) sowie in der Altersverteilung (mittleres Alter von 22 Jahren; min = 17, max = 35). Von insgesamt N = 55 TeilnehmerInnen wurde eine Person aufgrund technisch bedingten vorzeitigen Abbruchs ausgeschlossen.

## Ergebnisse Experiment II

Das wichtigste Ergebnis des zweiten Experiments ist das Eintreten des vorhergesagten Effekts der Neutralisationsmanipulation auf die Dosennutzung ( $\chi^2=6,13$ ;  $p=.01$ , Abb. 7). Während 24 (92,3 %) der insgesamt 26 Vpn in der Bedingung „Neutralisation gefördert“ eine Dose nahmen, so waren es in der Bedingung „Neutralisation gehemmt“ lediglich 18 (64,3 %) von 28 Personen.

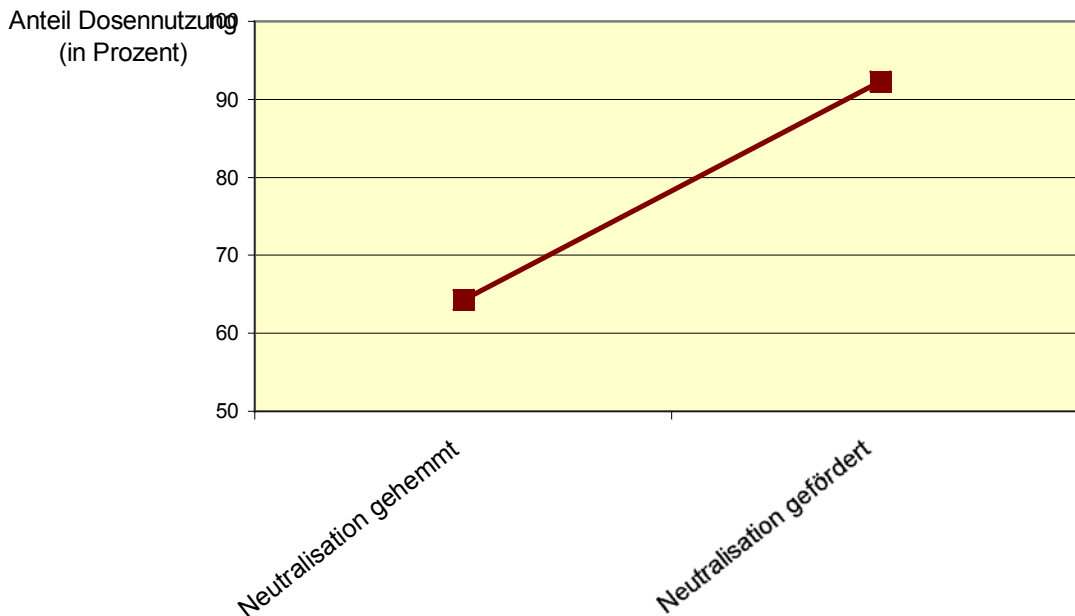


Abbildung 7: Relativer Anteil der Dosennutzung innerhalb der Versuchsbedingungen (N = 54)

Anhand dieser Daten sollten die drei oben genannten Moderatorhypothesen getestet werden.

Zur Testung des moderierenden Einflusses der individuellen Normakzeptanz auf den Neutralisationseffekt standen drei unterschiedliche Indikatoren dieser Normakzeptanz zur Verfügung. Die über eine erprobte 6-Item-Skala vor Neutralisationsepisode und Verhaltensgelegenheit erhobene allgemeine Umwelteinstellung stellt hierbei das theoretisch verhaltensfernste Merkmal dar. Ihr Einfluss sollte vermittelt sein über die individuelle Situationsauffassung (z.B. Bewertung der Dosennutzung als umweltschädigendes Verhalten, zugeschriebene individuelle Verantwortung), welche allerdings mit Rücksicht auf die Cover-Story des Experiments erst im Anschluss an die Verhaltensentscheidung erhoben werden konnten und deren Bestimmung daher anfällig war für konsistenzmotivierte Verzerrungen. Eine zweite für die Bestimmung der individuellen Normakzeptanz zur Verfügung stehende Variable war die Antwort auf das im Anschluss an den die Items zur allgemeinen Umwelteinstellung enthaltenden Fragenblock aufgeführte Item „Ich befürworte Umweltschutz.“. Dieses Item diente der ergänzenden Induktion der Normakzeptanz und wurde erwartungsgemäß von nur einer Person mit einem Wert unterhalb des Skalenmittelpunkts beantwortet. Die Antworten variieren also im Wesentlichen zwischen den Werten 3 und 4. Möglicherweise lässt sich aus dieser Unterscheidung eine unterschiedlich stark ausgeprägte Normakzeptanz ablesen. Die dritte im Sinn der individuellen Normakzeptanz interpretierbare Variable sollte am stärksten mit der spezifischen Verhaltenssituation korrespondieren und dem nahekommen, was Schwartz (1977) als

„spezifische persönliche Norm“ beschreibt. Die untersuchungsmethodischen Nachteile dieser Variable liegen in der Messung durch lediglich ein einzelnes Item sowie im späten Erhebungszeitpunkt (die spezifische persönliche Norm wurde erst nach Neutralisationsmanipulation und Getränkewahl erhoben). Der letztgenannte Umstand sollte aufgrund der bereits oben erwähnten Gefahr der konsistenzmotivierten Verzerrung am stärksten zur Vorsicht im Umgang mit dieser Variablen anhalten. Beide Nachteile sind den Zwängen geschuldet, die sich durch ein Experiment mit Cover-Story ergeben.

Zur Testung der Moderatorhypothese wurden für alle drei Variablen separate schrittweise logistische Regressionsanalysen (SPSS-Prozedur LOGISTIC REGRESSION) durchgeführt, in denen zunächst auf Haupteffekte der Neutralisations- sowie der Normakzeptanzvariablen und anschließend auf Interaktion dieser beiden Variablen getestet wurde (vgl. Baron & Kenny, 1986). Tab. 5 zeigt die Ergebnisse für diese Analysen.

Tabelle 5: Schrittweise logistische Regressionen (LOGISTIC REGRESSION, Methode FSTEP(COND) mit PIN(.10) POUT(.10), es wurde ein konstanter Term zugelassen); Ergebnisse für alle drei Schritte

	Neutralisation	Normakzeptanz	Neutralisation x Normakzeptanz	Nagelkerkes R <sup>2</sup>
<u>AV: Dosennutzung</u>				
<u>Kovariaten: Neutralisation, allgemeine Umwelteinstellung</u>				
Schritt 1: Neutralisation	B = -1,897 (p < .05)			.176
Schritt 2: Normakzeptanz	-	-		-
Schritt 3: Neutralisation x Normakzeptanz	-	-	-	-
<u>AV: Dosennutzung</u>				
<u>Kovariaten: Neutralisation, Induktionsitem Umwelteinstellung</u>				
Neutralisation	B = -1,854 (p < .05)			.17
Normakzeptanz	B = -2,046 (p < .05)	B = -1,547 (p < .05)		.26
Neutralisation x Normakzeptanz	-	-	-	-
<u>AV: Dosennutzung</u>				
<u>Kovariaten: Neutralisation, spezifische persönliche Norm</u>				
Neutralisation	B = -1,897 (p < .05)			.18
Normakzeptanz	B = -1,769 (p < .05)	B = -0,421 (p < .05)		.29
Neutralisation x Normakzeptanz	B = -7,964 (p = .17)	B = -1,639 (p = .13)	B = 1,407 (p = .20)	.38

Es zeigt sich, dass die Stärke der gelungenen Umwelteinstellungsinduktion sowie die im Nachhinein berichtete spezifische persönliche Norm jeweils statistisch signifikante Beiträge zur Erklärung der Dosennutzung im Experiment leisten, nicht jedoch die allgemeine Umwelteinstellung, das untersuchungsmethodisch anspruchsvollste Maß der Normakzeptanz.

Hinsichtlich der angenommenen Moderatorfunktion der Normakzeptanz auf den Neutralisationseffekt wird der entsprechende Interaktionsterm lediglich im Fall der spezifischen persönlichen Norm in die finale Regressionsgleichung aufgenommen. Dies verdankt er dem für diese Analysen mit  $PIN = .10$  relativ großzügig gewählten Aufnahmekriterium. Das Regressionsgewicht der Interaktion erreicht eine Irrtumswahrscheinlichkeit von  $p = .20$  und ist daher nur mit äußerster Vorsicht und allein unter explorativen Gesichtspunkten zu betrachten. Für eine Darstellung dieses möglichen Effekts wurde die Variable „spezifische persönliche Norm“ aus Gründen höherer Übersichtlichkeit durch Mediansplit dichotomisiert (Abb. 8). Beim Betrachten dieser Abbildung lässt sich keine bedeutsame Unterschiedlichkeit des Neutralisationseffekts unter den Bedingungen geringer vs. starker spezifischer persönlicher Norm feststellen.

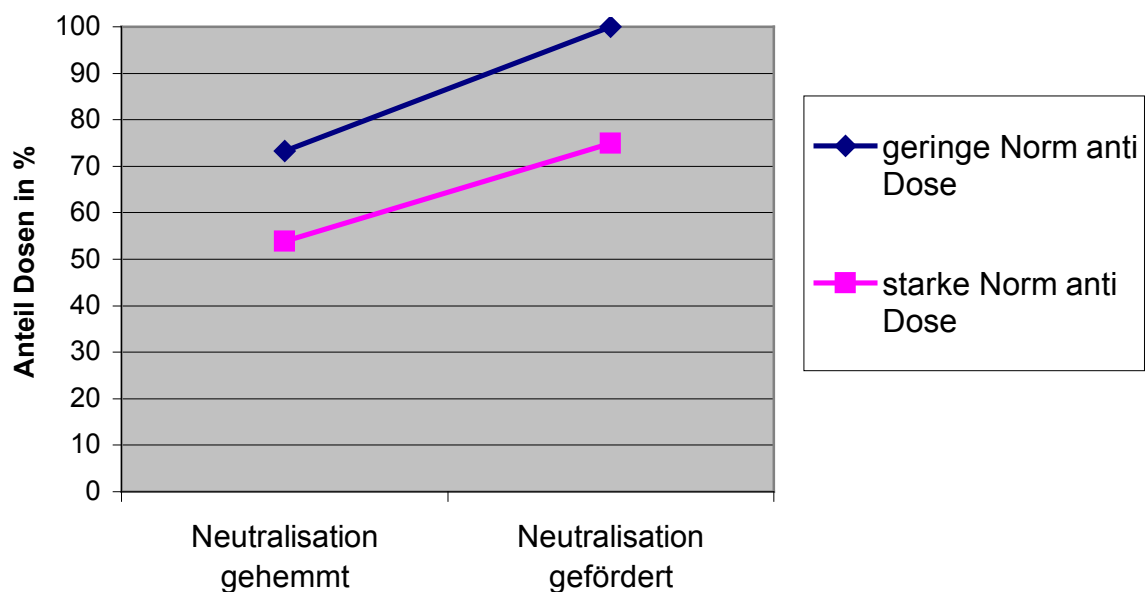


Abbildung 8: Relativer Anteil der Dosennutzung in Abhängigkeit von Neutralisation und spezifischer persönlicher Norm

In weiteren Tests auf Interaktionen höherer Ordnung zeigten sich explorativ interpretierbare Ergebnisse für die Interaktion von Neutralisation, spezifischer persönlicher Norm und allgemeiner Umwelteinstellung sowie alternativ zur letzten Variable auch des Einstellungsinduktionsitems. Die Ergebnisse für die erstgenannte – statistisch lediglich tendenziell bedeutsame – Dreifachinteraktion sind in Abb. 9 vereinfacht (stetige Variablen wurden am Median dichotomisiert) dargestellt. Bei der Betrachtung dieses Ergebnisses zeigt sich ein auf den ersten Blick überraschendes Bild: Der Neutralisationseffekt (Steigung der Geraden) ist in der Bedingung am stärksten, in der sich allgemeine Umwelteinstellung (= allgemeine persönliche Normakzeptanz) und spezifische persönliche Normakzeptanz in ihrer Ausrichtung widersprechen. Zieht man nun in Betracht, dass die allgemeine Normakzeptanz der spezifischen Normakzeptanz in der Erklärungskette des Verhaltens vorangeht, so hat hier möglicherweise eine Verhandlung der allgemeinen Norm stattgefunden, die ihren Niederschlag in einer entgegengerichteten spezifischen Norm gefunden hat. Entsprechend wichtig ist die Neutralisationsepisode für das Verhalten dieser Personen.

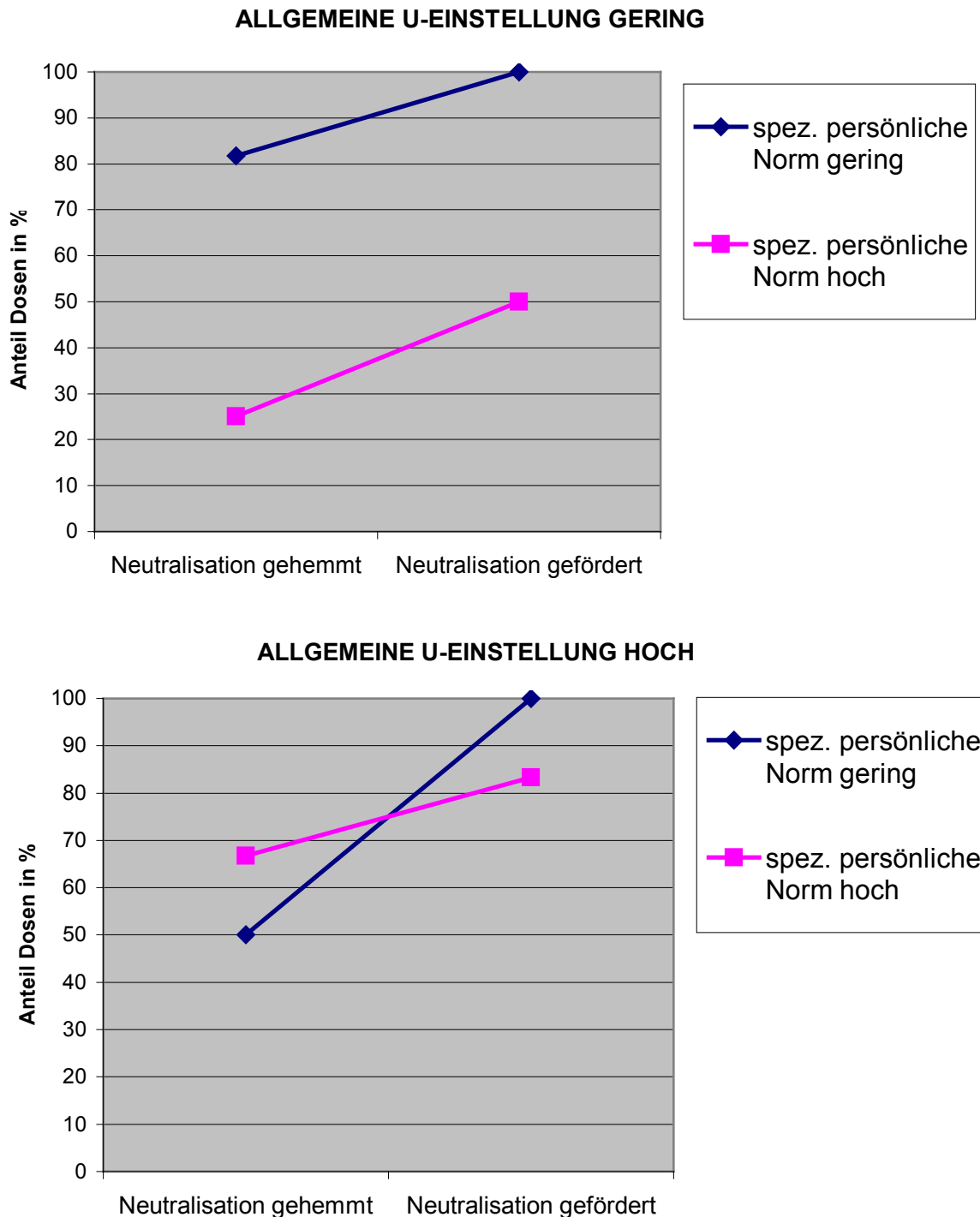


Abbildung 9: Relative Dosennutzung in Abhängigkeit von Neutralisation, spezifischer persönlicher Norm und allgemeiner Umwelteinstellung (beide letztgenannten Variablen für die Abb. am Median dichotomisiert). Logistische Regression (SPSS-Prozedur LOGISTIC REGRESSION; Methode: FSTEP(COND); ohne Zweifach-Interaktionen;  $B = 0,196$ ;  $p = .075$ ; Nagelkerkes  $R^2 = .41$ )

Eine weitere potentielle Moderatorvariable des Neutralisationseffekts ist die Schwere des Normbruchs. Im Rahmen einer Vorstudie wurde der Frage nachgegangen, ob das Trinken aus einer Einwegdose als ein prototypisches umweltschädigendes Verhalten wahrgenommen wird. Dies wurde anhand des zugeschriebenen Grades der Umweltschädlichkeit dieses Verhaltens im Kontext anderer Verhaltensweisen (von „Lösungsmittel in den Ausguss gießen“ bis „Zigarettenkippe auf dem Bürgersteig austreten“) entschieden. Die Dosennutzung stand hierbei unter 15 aufgeführten Verhaltensweisen mit einem Mittelwert deutlich oberhalb



des Skalenmittels auf Rang 5. Diese Einschätzung der relativen Umweltschädlichkeit der Dosennutzung wurde im Schlussfragebogen des zweiten Experiments mittels einer auf insgesamt fünf zu beurteilende Verhaltensweisen reduzierten Skala erhoben und kann so zum Neutralisationseffekt in Beziehung gesetzt werden. Allerdings ist diese Variable – wie viele andere der Informationen aus dem abschließenden Fragebogen – möglicherweise ebenfalls mit systematischen Verzerrungen behaftet. Nimmt man sie nach der Neutralisationsvariable in eine schrittweise logistische Regression auf, so zeigt sich weder ein Haupt- noch ein Interaktionseffekt auf die Dosennutzung. Dieses Ergebnis wiegt für die Diskussion des Einflusses der Tatschwere allerdings vergleichsweise leicht, da es sich bei Analysen unter Verwendung dieser Variable lediglich um eine methodisch unvollkommene Testung des moderierenden Einflusses der „Tatschwere“ handeln kann. Ein angemessenerer – aber im Rahmen der beiden aufwendigen Experimente nicht leistbarer – Zugang wäre der Vergleich des Neutralisationseffekts bei vorher als unterschiedlich schwerwiegend eingeschätzten Fehlverhaltensweisen.

Der letzte zu untersuchende potentielle Moderator des Neutralisationseffekts ist die Häufigkeit vorangegangener Normverstöße. Auch auf diese Variable trifft die oben schon mehrfach geäußerte Einschränkung der Erhebungsqualität durch Platzierung im Schlussfragebogen zu. Allerdings besitzt das eingesetzte Item einen klaren zeitlichen Bezugspunkt (Häufigkeit der Dosennutzung während der letzten drei Monate) und unterscheidet sich von einer im ersten Experiment eingesetzten Vorgängerversion aufgrund entsprechender Erfahrungen durch den Fokus auf den unteren Wertebereich. Möglicherweise sichern diese Merkmale der Erhebung der Häufigkeit vorangegangener Normverstöße eine annehmbare Validität. Ein Haupt- oder Interaktionseffekt auf die Dosennutzung lässt sich mittels der logistischen Regression für diese Variable jedoch nicht nachweisen.

## **Diskussion Experiment II**

Im Gegensatz zum ersten konnte im zweiten Experiment ein Effekt der Neutralisationsbedingungen auf die Dosennutzung der TeilnehmerInnen nachgewiesen werden. Die Orientierung an einem Arbeitsmodell der situativen Neutralisationswirkung bei der Umgestaltung des Versuchsaufbaus sowie die Erhöhung der Interventionsqualität und möglicherweise auch der -intensität führte zu Bedingungen, unter denen ein Neutralisationseffekt auftreten kann. Nachfolgende experimentelle Neutralisationsforschung sollte klären, welche der einzelnen Bedingungsveränderungen für die Verhaltenswirksamkeit von Rechenschaftskognitionen tatsächlich notwendig sind. Hierbei wäre – gerade vor einem Anwendungs- bzw. Interventionshintergrund – von besonderem Interesse zu erfahren, ob auch ein reines Verfügbarkeitstreatment den Neutralisationseffekt hervorrufen kann. Gründe dafür, dass ein solches Treatment nicht ausreicht, könnten in der Entwicklung von generalisierten sowie stark individuumsspezifischen Neutralisationsmustern liegen.

Die in der Neutralisationsliteratur vereinzelt berichteten Moderatoreffekte können weitgehend nicht bestätigt werden. Für die Variablen „Häufigkeit vergangener Normverstöße“ und „Schwere des Normbruchs“ ließen sich keinerlei Interaktions- oder Haupteffekte feststellen. Zumindest im Fall der letztgenannten Variable lässt sich dies möglicherweise auf eine unangemessene Erhebung durch Selbstbericht zurückführen. Vielmehr sollte zukünftig das Auftreten des Neutralisationseffekts in unterschiedlich schwer einzustufenden Verhaltensbereichen verglichen werden.

Hinsichtlich des moderierenden Einflusses der individuellen Normakzeptanz zeigt sich ein komplexes Ergebnismuster. Dies zeigt sich zunächst in unterschiedlichen Ergebnissen für unterschiedliche Operationalisierungen der potentiellen Moderatorvariable. Der Interaktionsterm von Neutralisation und Normakzeptanz trägt hierbei nur dann zur Vorhersage des Verhaltens bei, wenn es sich bei der Normakzeptanz um den Grad einer situationsspezifischen persönlichen Norm handelt. Im Fall zweier globaler Normakzeptanzmaße wird der Interaktionsterm nicht in die schrittweise gebildete Regressionsgleichung aufgenommen. Dieses Ergebnis kann mit der kausalanalytischen Distanz erklärt werden, welche die allgemeine Umwelteinstellung gegenüber dem Verhalten hat, stellt die experimentelle Neutralisationsforschung allerdings vor das Problem, spezifische Einstellungen im Vorfeld der experimentellen Manipulation erheben zu müssen, ohne dabei die Cover-Story in Gefahr zu bringen.

Bezieht man die Ergebnisse der ergänzenden explorativen Analysen in die Gesamtinterpretation ein, so zeigt sich ein Zusammenwirken dreier Variablen bei der Erklärung der Dosennutzung. Eine empirisch tendenziell signifikante Dreifachinteraktion von Neutralisationsbedingung, spezifischer persönlicher Norm und allgemeiner Umwelteinstellung deutet auf ein Bestehen des Interaktionseffekts von Neutralisation und Akzeptanz einer proximalen Norm lediglich unter der Bedingung hoher Akzeptanz der distalen Norm hin. Der Neutralisationseffekt ist hierbei unter der Bedingung besonders stark ausgeprägt, dass sich allgemeine und spezifische persönliche Norm widersprechen. Dieses Ergebnis entspricht einer Vorstellung von der Funktion von Neutralisationen, wie sie beispielsweise im Norm-Aktivationsmodell von Schwartz (1977) und Schwartz & Howard (1981) zum Ausdruck kommt. Rechenschaftskognitionen werden hier als Verdeutlicher widersprechender verhaltensbezogener Kognitionen betrachtet, die in Form eines internen Revisionsprozesses verhaltenssteuernde Kognitionen verdeutlichen.

## **Ausblick und Diskussion**

Die „Dosenuntersuchung“ zeigt die Darstellbarkeit der Verhaltenswirksamkeit von Neutralisationen im Experiment. Spezifische Bedingungen und Einschränkungen dieses Effekts sollten in Nachfolgeexperimenten geringerer untersuchungsmethodischer Aufwendigkeit abgeklärt werden. Es böte sich beispielsweise an, rechner- oder sogar internetgestützte Experimente zu entwerfen, innerhalb derer die von Experiment I auf Experiment II geänderten Bedingungen systematisch variiert werden. Vielversprechend dürfte hierbei die Überprüfung und Verfeinerung des präsentierten Arbeitsmodells der situativen Verhaltenswirksamkeit von Neutralisationen sein.

Ebenfalls sollte bei der nachfolgenden Untersuchung der Vergleich unterschiedlicher unter normativem Einfluss stehender Verhaltensbereiche im Mittelpunkt stehen.

## **Literatur**

Ajzen, I. & Madden, T. J. (1986). Prediction of goal-directed behavior: attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology*, 22, 453-474.  
Amelang, M., Schahn, J. & Kohlmann, D. (1988). Techniken der Neutralisation: Eine modelltestende Untersuchung auf der Basis offizieller und selbstberichteter Delinquenz. *Monatsschrift für Kriminologie und Strafrechtsreform*, 71 (3), 178-190.

- Amelang, M., Zahn, C. & Schahn, J. (1988). Empirische Prüfung einiger Elemente der Neutralisations-Theorie. In G. Kaiser, H. Kury & H.-J. Albrecht (Hrsg.), *Kriminologische Forschung in den 80er Jahren. Projektberichte aus der Bundesrepublik Deutschland* (S. 727-756).
- Ball, R. A. (1966). An empirical exploration of neutralization theory. *Criminologica*, 4, 22-32.
- Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer (Ed.), *The automaticity of everyday life: advances in social cognition* (S. 1-61). Mahwah, NJ: Lawrence Erlbaum Associates.
- Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51 (6), 1173-1182.
- Dodder, R. A. & Hughes, S. P. (1993). Neutralization of drinking behavior. *Deviant Behavior*, 14, 65-79.
- Egg, R. & Sponsel, R. (1978). „Bagatelldelinquenz“ und Techniken der Neutralisierung (Eine empirische Prüfung der Theorie von Sykes & Matza). *Monatsschrift für Kriminologie und Strafrechtsreform*, 61 (1), 38-50.
- Fritsche, I. & Mayrhofer, R. (2001, Juni). *Account strategies and normcontradictive behaviour: A new taxonomy exemplified in the context of environmentally harmful behaviour*. Posterpräsentation auf der IACM 2001, Paris.
- Hollinger, R. C. (1991). Neutralizing in the workplace: An empirical analysis of property theft and production deviance. *Deviant Behavior*, 12, 169-202.
- Kuckartz, U. (2000). Umweltbewusstsein in Deutschland. Ergebnisse einer repräsentativen Bevölkerungsumfrage. Berlin: Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit.
- Lauströer, A. (1996). *Sand ins Getriebe! Rechtfertigungen als soziales Schmiermittel im Umweltdiskurs*. IPU-Homepage, November.
- Minor, W. W. (1981). Techniques of neutralization: A reconceptualization and empirical examination. *Journal of Research in Crime and Delinquency*, 18, 295-318.
- Minor, W. W. (1984). Neutralization as a hardening process: Considerations in the modeling of change. *Social Forces*, 62, 995-1019.
- Norris, T. D. & Dodder R. A. (1979). A behavioral continuum synthesizing neutralization theory, situational ethics and juvenile delinquency. *Adolescence*, 55, 545-555.
- Opp, K.-D. (1974). *Abweichendes Verhalten und Gesellschaftsstruktur*. Darmstadt: Luchterhand.
- Schahn, J., Damian, M., Schurig, U. & Fuchsle, C. (2000). Konstruktion und Evaluation der dritten Version des Skalensystems zur Erfassung des Umweltbewußtseins (SEU-3). *Diagnostica*, 46 (2), 84-92.
- Schahn, J., Dinger, J. & Bohner, G. (1995). Rationalisierungen und Neutralisationen als Rechtfertigungsstrategien: Ein Vergleich zwischen Umwelt- und Delinquenzbereich. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 16 (3), 177-194.
- Schönbach, P. (1990). *Account episodes. The management or escalation of conflict*. Cambridge: University Press.
- Schwartz, S. H. (1977). Normative influences on altruism. In: L. Berkowitz (Hrsg.), *Advances in experimental social psychology*. Vol. 10 (S. 221-279). New York: Academic Press.
- Schwartz, S. H. & Howard, J. A. (1981). A normative decision-making model of altruism. In: J. P. Rushton, & T. M. Sorrentino (Hrsg.), *Altruism and helping behavior: Social, personality, and developmental perspectives* (189-211). Hillsdale: Lawrence Erlbaum.
- Schwarz, N. & Bayer, A. (1989). *Variationen der Fragenreihenfolge als Instrument der Kausalitätsprüfung: eine Untersuchung zur Neutralisationstheorie devianten Verhaltens*.

ZUMA-Arbeitsbericht Nr. 89/23. Mannheim: Zentrum für Umfragen, Methoden und Analysen.

Shields, I. W. & Whitehall, G. C. (1994). Neutralization and delinquency among teenagers. *Criminal Justice and Behavior*, 21 (2), 223-235.

Sykes, G. M. & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. *American Sociological Review*, 43, 643-656.

Thurman, Q. C. (1984). Deviance and the neutralization of moral commitment: An empirical analysis. *Deviant Behavior*, 5, 291-304.

Winkel, F. W. (1999?). Criminal behavior and the pre-victimization process: Three studies on neutralization, redefinition, and desensitization.

<sup>1</sup> ZUMA = Zentrum für Umfragen, Methoden und Analysen e. V., Mannheim

<sup>2</sup> Neben den Themen Umweltschutz und Schlechtes Gewissen standen folgende Themen zur Auswahl : (...)

# Survivability - Die Psychologie der Eigensicherung

Uwe Füllgrabe

Bildungsinstitut der Polizei Niedersachsen, Hann. Münden

## 1 Der Begriff der Survivability

Verschiedene Ereignisse wie z.B. die Ermordung von Polizeibeamten und Geiselnahmen in Gefängnissen zeigen, dass Personen, die im polizeilichen, forensischen u.ä. Bereichen arbeiten, ein besonders hohes Gefährdungspotenzial durch gewaltbereite Personen haben. In der Praxis zeigt sich aber auch, dass es große individuelle Unterschiede in der Fähigkeit gibt, gefährliche Situationen rechtzeitig zu erkennen, zu vermeiden und zu überleben. Dies drückte auch Siebert (1996) mit seinem Begriff der „Überlebenspersönlichkeit“ (Survivor personality) aus, wobei er nur verschiedene Falldarstellungen und eine unsystematische Aufzählung psychologischer Faktoren liefert, die seiner Erkenntnis nach eine Rolle spielen. Auch wird nicht immer genug deutlich, was die betreffenden Personen in der Gefahrensituation dachten, wie sie handelten.

Während es aber zahlreiche Untersuchungen zu der negativen Seite der Konfrontation mit Gefahren und Katastrophen gibt, nämlich posttraumatischen Symptomen, gibt es dagegen erstaunlicherweise kaum wissenschaftliche Untersuchungen zu den im wahrsten Sinne des Wortes überlebenswichtigen Fragen: wie kann man Gefahren rechtzeitig erkennen, bewältigen und überleben?

Die Gesamtzahl der psychologischen u.a. Faktoren, die hier eine Rolle spielen, möchte ich unter dem Begriff **Survivability** (von *to survive* und *ability*) zusammenfassen. Man kann feststellen, dass dieser Begriff auch durchaus metaphorisch gebraucht wird, etwa für posttraumatische Probleme von Vietnamesen, die einen Piratenüberfall überlebten bis hin zu der Suche nach den Unterschieden zwischen Lehrern, die die erste Runde der Delphi – Methode „überlebten“ und denen, die diese nicht „überlebten“, also einem kognitiven Problem. Doch das existenzielle Problem wird durch derartige Untersuchungen überhaupt nicht berührt. Es ist jedoch gerade für Personen, die beruflich mit gefährlichen Situationen konfrontiert werden könnten, z.B. Polizisten, wichtig, auch die Bewältigung der *körperlichen* Bedrohung unter diesem Begriff zu betrachten, also *vor* und *während* eines gefährlichen Ereignisses.

Was kann die Psychologie dazu beitragen, diesen leeren Fleck auf der Landkarte des Wissens auszufüllen und zwar in theoretischer Hinsicht *und* hinsichtlich konkreter Handlungsanweisungen ?

Damit eine wissenschaftliche Beschäftigung mit dem Thema „Überleben gefährlicher Situationen“ überhaupt sinnvoll ist, muss eine Voraussetzung erfüllt sein: Es darf nicht ausschließlich von *Zufallsfaktoren* abhängen, dass Menschen, die in gefährliche Situationen geraten, diese überleben. Vielmehr muss ihr Überleben von ihrem *eigenen* Verhalten, bestimmten psychologischen Faktoren usw. abhängig sein.

Natürlich ist es bei dieser Thematik- ähnlich anderen Extremphänomenen wie z. B. Panik - naturgemäß kaum möglich, dazu direkte, realitätsnahe experimentelle Untersuchungen anzustellen. Es gibt aber verschiedene Informationsquellen, die wichtige Informationen liefern, z.B.

⇒ Vergleich der Verhaltensweisen von Polizisten, die im Dienst getötet wurden mit denen, die eine Gefahr überlebten

⇒ Analyse der kognitiven Struktur von Tätern und überlebenden Polizisten

⇒ Erfahrungen von Personen, die die verschiedenartigsten Gefahren überlebt hatten.

Die könnte z.B. reichen von den systemtheoretischen Gedanken des japanischen

Schwertmeisters Musashi (1993) bis zu den Persönlichkeitsfaktoren der *Überlebenspersönlichkeit* Sieberts (1996).

Die konkreten Vorfälle kann man gemäß der *Methode der kritischen Vorfälle* analysieren und mit experimentellen Erkenntnissen vergleichen, wobei z.B. die Bindungsstilforschung (s. Miculincer, 1997, 1998) die psychologische Vernetzung dieser Ereignisse aufzeigt.

Bei der Darstellung der Thematik und der Verdeutlichung der Notwendigkeit der Risikovermeidung durch Beachtung von Sicherheitsbestimmungen hat sich gezeigt, dass es zum besseren Verständnis hilfreich ist, die Zuhörer zu bitten, den Ausgang folgender und ähnlicher Ereignisse zu erraten:

(1) Ein Serienvergewaltiger und Serienmörder bedrohte eine Frau in einem Fahrstuhl. Sie stieß ihn zurück, drückte den Türöffner, rannte heraus. Sie stolperte; der Täter fiel über sie, verlor sein Messer. Er lag neben der schreienden Frau.

(2) Ein drogenabhängiger Rauschgifthändler behauptete, dass er eine Botschaft von Gott erhalten habe, dass er einen Polizisten töten sollte, weil die Polizei seinen Drogenhandel ruinierte. Um dies zu vollenden, ging er zu einer Kreuzung in der Nähe seines Hauses, um einen Polizisten zu finden, den er töten konnte. An der Kreuzung beobachtete er einen uniformierten Sergeant an einer Tankstelle, der einen Reifen an seinem Streifenwagen reparieren ließ. Nach seinen eigenen Angaben näherte sich der Täter dem Sergeant, mit der Absicht, ihn zu töten.

(3) Ein junger Polizist hielt ein Auto wegen eines geringfügigen Vergehens an und näherte sich ihm sachgemäß. In dem Auto befanden sich vier Personen, die alle nicht besonders freundlich aussahen. Der Fahrer erzählte, dass er keinen Ausweis bei sich hatte, und die Körpersprache signalisierte dem Polizisten, dass etwas mit ihm nicht stimmte. Später stellte sich heraus, dass a) der Fahrer eine geladene Pistole vor seinem Sitz befestigt hatte und b) der Fahrer nur auf Bewährung frei war und zwei weitere der Personen im Auto wegen schwerer Verbrechen gesucht wurden.

Es zeigt sich, dass die Ausgänge dieser Fälle keineswegs leicht zu erraten sind. Im ersten Fall könnte man eigentlich erwarten, dass der Täter das Messer ergriff, um die Frau zu töten. Doch der Täter berichtete später:

„... ich lag auf dem Boden neben ihr, zu Tode erschrocken. Mein Geist war leer. Ich rannte aus dem Gebäude.“ Er wurde daraufhin gefaßt (Ressler u.a., 1983, S. 139).

Im zweiten Fall könnte man glauben, dass der Polizist nur eine geringe Überlebenschance gehabt hätte. Doch Pinizzotto und Davis (1999) stellten bei der genauen Analyse dieses und zweier weiterer gleich gelagerter Fälle, bei denen der Täter einen beliebigen Polizisten töten wollte, ein interessantes Muster fest: Die Täter töteten nicht den jeweils ersten Polizisten, den sie trafen, sondern denjenigen, dem sie als nächsten begegneten. In allen drei Fällen konnten die Täter als Grund für ihr Zögern, den ersten Polizisten zu töten, nur sagen: „Er sah so aus, dass es schwierig war, ihn zu überwältigen.“ Ähnlich war es beim dritten Beispiel.

Obwohl es sich nur um einen geringfügigen Verkehrsverstoß handelte, „schluckte der Polizist seinen Stolz hinunter“ und rief einen Streifenwagen zur Verstärkung. Als der Fahrer gefragt wurde, ob er daran gedacht hatte, seine Waffe zu benutzen, antwortete er offen: Ja, aber wegen der Art und Weise, wie sich der Polizist dem Auto näherte und den Kontakt aufgenommen hatte, hatte der Verdächtige nicht den Eindruck, dass er erfolgreich einen Schuss abgeben könnte, ohne selbst erschossen zu werden (Füllgrabe, 2001).

Man kann aus diesen Beispielen ableiten:

- **Überleben ist kein Zufall.**
- Entscheidend ist das *aktive Handeln*, vor allem das unerwartete Verhalten. Die Frau im ersten Beispiel hatte dadurch gewissermaßen das „Drehbuch“ des Täters zerrissen.
- Ein Konflikt kann häufig schon durch selbstsicheres (nicht überhebliches!) Auftreten des Polizisten auf der Grundlage von *sachgerechtem* Verhalten vermieden werden (Deeskalation).
- Dieses sachorientierte Verhalten lenkt und beeinflusst die Interaktion, denn es zeigt einer gewaltbereiten Person:
  - a) Dieser Polizist versteht sein Handwerk.
  - b) **ER** hat die Situation unter Kontrolle.
- Die gewaltbereite Person erkennt daraus: es ist besser für mich, wenn ich friedlich bleibe. Wenn die Situation dann doch gewalttätig wird, kann man sie durch sachgerechtes Verhalten überleben.
- Grundvoraussetzung für das Überleben ist ein „*Gefahrenradar*“.

## 2 Das richtige Weltbild

Warum ist es so schwer, den Ausgang der eingangs beschriebenen Ereignisse zu erraten? Dies hängt nicht nur mit einem Informationsmangel zusammen, etwa hinsichtlich der Psychologie gewaltbereiter Personen, z. B. *Kulturen der Ehre* (Cohen u.a., 1996), *Gesetz der Straße* (code of the streets, Anderson, 1994). Noch wichtiger ist, dass der von Kurt Lewin in seinem berühmten Vortrag vom 4. Februar 1930 angemahnte „Übergang von der aristotelischen zur galileischen Denkweise in Biologie und Psychologie“ in vielen Bereichen der Psychologie noch keineswegs stattgefunden hat. Dort findet man – ungeachtet aller Lippenbekenntnisse - immer noch eine recht statische Denkweise (s. z. B. die Konzeption der *Big Five* in der Persönlichkeitspsychologie), und die Umgebung spielt nur insofern eine Rolle, als sie „Störungen“ herbeiführen kann (was Lewin 1971, S. 33, als typisch für die aristotelische Begriffsbildung bezeichnet).

In Wirklichkeit spielt aber die Situation / Umgebung eine große Rolle, was z.B. Waddington (1957) in seiner *epigenetischen Landschaft* verdeutlicht: zu bestimmten kritischen Zeitpunkten kann sich ein System (Erbanlage, menschliches Schicksal usw.) in völlig unterschiedliche Richtungen hin entwickeln. Aber auch die Unkenntnis der tatsächlich wirkenden *individuellen* Faktoren verhindert die sachgerechte Betrachtungsweise der Ereignisse. Dies zeigt sich besonders am zweiten Beispiel deutlich, weshalb wir eine *zwischenmenschliche* Spieltheorie (Füllgrabe, 1997, 2001) brauchen. In zwischenmenschlichen Interaktionen sind nämlich die „Spielzüge“, die Entscheidungen nicht unbedingt rational, sondern werden durch Vertrauen, Täuschung, Angst, synergistisches Denken, persönlichen Mythen, den individuellen Bindungsstil usw. beeinflusst, so dass die Person unter Umständen sogar gegen ihre eigenen Interessen handeln kann, wie z.B. Frauen, die Mörder lieben oder Frauen, die bei einem Mann bleiben, der sie schlägt (Füllgrabe, 1997).

Beispiel 2 belegt die spieltheoretische Erkenntnis, dass unkooperative (hier aggressionsbereite) Personen ihr Gegenüber völlig anders als kooperationsbereite bewerten, nämlich nicht gemäß der Dimension *Freundlichkeit*, sondern gemäß einer *Machtdimension* (Beggan u.a., 1986). In Fall Nr. 2 sagte der Täter, dass dieser Polizist nicht besonders groß war oder bedrohlich im Aussehen, aber „so ausschaute, als ob er sich (bei einer Auseinandersetzung) gut selbst behaupten könnte.“ Dagegen strahlten die späteren Opfer in den von Pinizzotto und Davis (1999) untersuchten Fällen Zeichen der Unsicherheit aus. Dies zeigt auch, dass gerade in gewaltbereiten Umgebungen die TIT FOR TAT – Strategie wichtig ist: freundlich sein, aber sofort auf unkooperatives /

aggressives Verhalten reagieren können. Hier zeigt die Praxis auch, dass Personen, die gemäß der TIT FOR TAT – Strategie handelten, Gewalt vermieden, Geiselnahmen überstanden und selbst in gewaltbereiten Umgebungen ein kooperatives Beziehungsgeflecht aufbauen konnten (s. z. B. Bobinsky, 1994; Füllgrabe, 2001).

### 3 Die psychologischen Hintergrundfaktoren

Welche psychologischen Faktoren spielen bei der *Psychologie der Eigensicherung* eine Rolle?

Aus der Analyse konkreter Vorfälle lassen sich gemäß der „Methode der kritischen Vorfälle“ (critical incidents technique) die einzelnen Komponenten dieser Fähigkeit ermitteln: aktiver vs. passiver Lebensstil (Füllgrabe, 1999), synergistischer Denkstil vs. ICH - Zentrierung, gelassene Wachsamkeit vs. Angst, Humor, nichtsprachliche Signale der Selbstsicherheit usw. Diese Komponenten stehen in Verbindung mit bekannten psychologischen Konstrukten, z.B. *Bindungsstil*, *Monitoring / Blunting* (Miller, 1990), *Kontrollüberzeugung*, wobei z.B. in unterschiedlichen Situationen durchaus entgegengesetzte Strategien überlebenswichtig sein können: Informationssuche zur Gefahrenwahrnehmung (Monitoring) und Informationsabschottung (Blunting) in längerfristigen Gefahrenlagen, wodurch Gedanken der Hoffnungslosigkeit vermieden werden (Füllgrabe, 2001).

Dass die genannten Persönlichkeitsfaktoren durchaus einen Einfluss auf die Survivability haben, wird, obwohl es bisher keine *direkt* ermittelten Korrelationen gibt, durch folgende drei Untersuchungsergebnisse aufgezeigt:

a) lebensbedrohliche Situationen leichter überleben, wenn sie an Bezugspersonen denken (Füllgrabe, 1999, 2001), wodurch sie das **psychologische Immunsystem** aktivieren. Dieses besteht aus inneren Monologen, die beruhigen und zum Handeln aktivieren, wodurch Gedanken der Hilflosigkeit und Hoffnungslosigkeit vermieden oder beseitigt werden.

b) durch unkooperatives Verhalten schneller in Gefahr geraten. Sowohl deutsche (Sessar u.a., 1980) als auch amerikanische Untersuchungen (FBI, 1992) ermittelten, dass viele Polizisten, die im Dienst getötet wurden, keine Verstärkung riefen oder nicht auf die von ihnen gerufenen Kollegen als Verstärkung warteten.

c) leicht in Gefahr geraten, wenn ihre Wahrnehmung oder ihre Handlung durch „verfrühte kognitive Festlegungen“ beeinflusst werden. Langer (1991, S. 32) schreibt: „Da sich solche festen Einstellungen (*Mindsets*) ausbilden, bevor wir richtig nachgedacht haben, nennen wir sie *Premature cognitive commitments* – <<verfrühte kognitive Festlegungen>>.“

Eine solche „verfrühte geistige Festlegung“ ist z.B. wie in diesem Beispiel der Eindruck: Auf der Bettkante sitzt ein alter Mann. Die Situation ist *harmlos*. Doch der alte Mann greift plötzlich unter die Bettdecke, ergreift eine Pistole und verletzt zwei Polizisten, einen davon schwer. Durch eine genaue Durchsuchung und Beobachtung seines Verhaltens wäre dies zu verhindern gewesen.

Vergleicht man jetzt diese drei Untersuchungsergebnisse aus dem Bereich der Eigensicherung mit experimentellen Erkenntnissen, so kann man hier das Wirken des **Bindungsstils** erkennen.

Die Beziehung zwischen dem Bindungsstil und dem unterschiedlichen Erkennen, Vermeiden und Bewältigen gefährlicher Situationen wird besonders in den Untersuchungen von Miculincer (1997, 1998) deutlich. Er stellte nämlich fest, dass Personen mit sicherem Bindungsstil bessere Informationsverarbeiter sind: sie suchen aktiver nach Informationen, sind offener für neue Informationen. Wenn Miculincer (1997) daraus folgert, dass die Flexibilität ihrer kognitiven Strukturen der verbesserten Bewältigung und Anpassung an eine komplexe und sich verändernde Welt dient, sagt er



eigentlich damit aus, dass Personen mit einem sicheren Bindungsstil wohl weniger zu „Gedankenlosigkeit“ (Langer, 1982, 1991) neigen, also Gefahrensituationen vermutlich eher erkennen dürften.

Betrachtet man dagegen das wenig planvolle Handeln von Polizisten, die einen Angriff überlebten oder getötet wurden: Die Situation *war* anders und *entwickelte* sich anders, als es ihren kognitiven Schemata entsprach, ihre kognitiven Schemata waren zu starr und passten sich der neuen Situation nicht an. Dies entspricht genau dem, was Miculincer (1997) bei den beiden Gruppen von Bindungsunsicheren Personen feststellte. Sie vermeiden es, ihre kognitiven Schemata angesichts neuer Informationen zu öffnen und zu verändern. Sie schotten ihre kognitiven Schemata ab und vermeiden den Kontakt mit der Umwelt (also der Realität!), weil neue Informationen bei ihnen Ungewissheit und Verwirrung erzeugen könnten, und ihnen der Optimismus und das Gefühl fehlt, eine ungewisse Situation meistern zu können. Dies wäre aber gerade in Gefahrensituationen wichtig, um irrige Meinungen zu ändern, um ein realistisches Bild der Lage zu gewinnen, sich realistische Ziele zu setzen und realistische Handlungspläne zu entwickeln.

Von den beiden Bindungsunsicheren Gruppen vermeiden Personen mit meidendem Bindungsstil eher die Suche nach Informationen, spielen die Bedeutung von Informationen herunter u.ä., ängstlich – ambivalente Personen sind eher in einem Konflikt zwischen der Suche nach neuen Informationen und dem Bedürfnis nach positiven zwischenmenschlichen Beziehungen, was dann eher zu einer passiven Haltung („Nichts tun“) führt.

Einen direkten Bezug zur Survivability liefern Miculincers (1998) Untersuchungen zum Zusammenhang zwischen Bindungsstil und Vertrauen: Personen mit sicherem Bindungsstil erinnerten sich schneller an positive vertrauensbezogene Ereignisse, Personen mit unsicherem Bindungsstil haben dagegen schneller Zugriff auf negative Erinnerungen.

Angesichts der Tatsache, dass Menschen lebensbedrohliche Situationen deshalb überlebten, weil sie an Bezugspersonen dachten und sich so zum Handeln motivierten (Janis, 1971; Pinizzotto u.a., 1998), kann man die Hypothese wagen, dass Bindungssichere Personen gefährliche Situationen wohl eher bewältigen und überleben. Ihre Flexibilität bei der Informationsverarbeitung angesichts neuer und unvorhersagbarer Situationen verstärkt diese Hypothese noch. Dagegen dürfte die Abschottung Bindungsunsicherer Personen gegen neue Informationen eher zum Aufbau einer gefährlichen Situation beitragen und die fehlende oder geringe kooperative Orientierung das Abrufen positiver innerer Bilder oder Gedanken an Bezugspersonen und damit das Überleben in Gefahrensituationen erschweren.

Der Bindungsstil wirkt aber auch in einem häufigen Verhaltensmuster, das sowohl deutsche als auch amerikanische Polizisten zeigten, die angegriffen oder im Dienst getötet wurden: Sie riefen weder einen Kollegen um Hilfe, oder wenn sie es taten, warteten sie die Verstärkung nicht ab (FBI 1992, Pinizzotto u.a., 1997, Sessar u.a., 1980). Man findet hier das gleiche Verhalten wie im schulischen Bereich, wo manche Schüler bei der Lösung von Aufgaben um Hilfe ersuchen und andere nicht. Warum einige Menschen eine vorhandene „menschliche Ressource“ (z.B. Lehrer, andere Schüler) nicht in Anspruch nehmen und dadurch nur eine schlechtere Lösung eines Problems erreichen, hat besonders im pädagogischen Bereich unter dem Begriff „**die strategische Suche um Hilfe**“ umfangreiche Forschungen ausgelöst. Die Komplexität dieses Phänomens und dessen Bezug zu verschiedenen psychologischen Faktoren (z.B. Selbstachtung) wurde von Nadler (1998) dargestellt.

Das Hilfesuchen ist eine zwischenmenschliche Interaktion, umgekehrt kann man das Nichtersuchen um Hilfe als Weigerung ansehen, mit jemand in eine Interaktion einzutreten und eine „strategische Ressource“ zu nutzen. Aus der Bindungsstilforschung leitet Nadler (1998) deshalb z.B. ab: Personen mit **sicherem Bindungsstil** benutzen

Hilfe in angemessener Weise. Sie betrachten die Bitte um Hilfe als eine Bewältigungsstrategie, und wenn die Situation es erfordert, benutzen sie sie auch.

#### 4 Mentales Judo

Um zu wissen, wie man gefährliche Situationen rechtzeitig erkennt und bewältigt, bedarf es konkreter Handlungsanweisungen, was ich mit dem Begriff **Mentales Judo** ausdrücke, der praktischen Seite der **Survivability**. Obwohl die folgende Übersicht spezifisch auf den Streifenbeamten ausgerichtet ist, besitzt sie eine Allgemeingültigkeit für viele, besonders zwischenmenschliche, Gefahrensituationen.

Das **Mentale Judo** beinhaltet das Prinzip: Selbstverteidigung mit dem geringsten Aufwand und dem geringsten Grad an notwendiger Gewalt.

Der Begriff *Judo* wurde hier benutzt, weil Jigoro Kano, der Vater des modernen Judo, keineswegs nur eine Sportart erschuf, sondern damit auch das *Prinzip der drei Kulturen* verband, d.h. die Vernetzung der intellektuellen, moralischen und körperlichen Kultur (s. Maekawa & Hasegawa, 1963). Auch lernt man beim Judo, hart zu fallen, und dies ist metaphorisch wichtig für die Vorbereitung auf Gefahren, um posttraumatische Störungen zu vermeiden.

Das Modell behandelt u.a. die Bewältigung verschiedener Gefährdungsstufen:

1. *Eigensicherung* durch ein „*Gefahrenradar*“, wobei die Gefahrenwahrnehmung durch Angst oder *voreilige kognitive Festlegungen* beeinträchtigt werden kann, *Reaktionsbereitschaft usw.*
2. Krisenbewältigung (bei Schusswechseln usw.) durch: vorherige „*Stressimpfung*“, Abrufen automatisierter Verhaltensweisen.
3. Bei schweren Verletzungen und Bedrohungen des Lebens: *Aktivierung des psychologischen Immunsystems* und *schrittweises Handeln zum Entkommen aus der Gefahrensituation*.
4. *Nachbereitung* des Ereignisses zur Vermeidung von Ärger und posttraumatischen Symptomen.

Und dass gerade in längerfristigen Interaktionen mit gewaltbereiten Personen die TIT FOR TAT – Strategie die Überlebenschancen enorm erhöhen kann, weil sie den in Kulturen der Ehre (Cohen u.a., 1996) wichtigen Gesichtspunkt des Respekts erzeugt, belegt folgendes Beispiel:

Ein amerikanischer Gefängniswärter, der während einer Geiselnahme mit dem Tode bedroht war, formuliert aus den Erfahrungen während dieses Erlebnisses heraus: Der beste Aufbau von Beziehungen beginnt lange vor dem Ereignis, durch den alltäglichen Kontakt mit den Gefängnisinsassen. Im Umgang mit den Gefängnisinsassen hatte der Wärter immer versucht, sich daran zu erinnern, sie als Menschen zu respektieren. „Ich behandle jeden Insassen als Individuum, gleichgültig, welche (s) Verbrechen er begangen hat.“

Er ist sich sicher, dass das Vertrauen, das sich daraus ergeben hatte, ihm half, als Geisel zu überleben. Selbst als sie ihn mit dem Leben bedrohten, behandelten ihn die Geiselnahmer mit grundlegendem Respekt. Dagegen nannten sie die Namen von Gefängniswägtern, die sie als Gefangene misshandelt hatten und die sie - hätten sie sie in die Hand bekommen – getötet hätten.

## Mentales Judo bei der Eigensicherung

### I In kurzfristigen Interaktionen (z. B. Verkehrskontrolle)

1) *Vor* dem Umkippen des nichtaggressiven Zustandes einer Situation in Gewalt

**Eigensicherung:**

- *Nichtsprachliche und sprachliche Signale der Selbstsicherheit*
- *„Gefahrenradar“*
- *Reaktionsbereitschaft*
- *Mit Entschlossenheit handeln*

2) **Im Verlauf einer Krise** ( bei Schusswechseln usw.)

Bewältigung der Phasen einer Krise durch:

- *vorherige Stressimpfung*
- *Abrufen automatisierter Verhaltensweisen*

3) **Bei schweren Verletzungen und Bedrohungen des Lebens**

- *Aktivierung des psychologischen Immunsystems*
- *Schrittweises Handeln zum Entkommen aus der Gefahrensituation*

4) **Nachbereitung des Ereignisses**

- *Ärgerbewältigung*
- *Sachorientierung zur Vermeidung posttraumatischer Symptome*

**II Bei längerfristigen Interaktionen** ( im Kontakt mit gewaltbereiten Gruppen, Tätigkeit in gewaltbereiten Umgebungen)

- *Aufbau einer TIT FOR TAT – Kultur* ( eines kooperativen Beziehungsgeflechts)
- *Beachtung der Eigensicherung*

## Zusammenfassung

**Survivability**, die Fähigkeit rechtzeitig zu erkennen, zu bewältigen oder zu überleben, besteht aus drei Komponenten:

1. einer **Persönlichkeitsstruktur**, bei der aktive Informationssuche und – verarbeitung, eine realistische Kontrollüberzeugung, ein breites Verhaltensrepertoire, aktive Problemlösungen usw. vorherrschen.
2. der **TIT FOR TAT – Strategie**, die eine rechtzeitige Reaktion auf Gefahren beinhaltet und gleichzeitig geeignet ist, langfristig eine kooperierende Gemeinschaft aufzubauen.
3. **Handeln** gemäß den Prinzipien des Mentalen Judo.

## Literatur

- Anderson, E. (1994). The code of the streets. *The Atlantic Monthly*, 273, 5, 80-94.
- Beggan, J.K. et al. (1988). Social values and egocentric bias: Two tests of the might over morality hypothesis. *Journal of Personality and Social Psychology*, 55, 4, 606-611.
- Bobinsky, R. (1994). Reflecting on community – oriented policing. *FBI Law Enforcement Bulletin*, 63, 3, 15-19.
- Cohen, D. et al. (1996). Insult, aggression and the southern culture of honor: An „experimental ethnography“. *Journal of Personality and Social Psychology*, 70, 5, 945-960.
- FBI (Uniform Crime Reports Section. Federal Bureau of Investigation. United States Department of Justice.) (1992). *Killed in the line of duty: A study of selected felonious killings of law enforcement officers*. Washington D.C.: U.S. Department of Justice.
- Füllgrabe, U. (1997). *Kriminalpsychologie. Täter und Opfer im Spiel des Lebens*. Frankfurt: Edition Wötzel.
- Füllgrabe, U. (1999). Survivability: Überlebensfaktoren in gefährlichen Situationen – Zur Psychologie der Eigensicherung. *Praxis der Rechtspsychologie*, 9, 1, 28 – 52.
- Füllgrabe, U. (2001). *Die Psychologie der Eigensicherung – Überleben ist kein Zufall*. Stuttgart: R. Boorberg Verlag.
- Langer, E. I. (1991). *Aktives Denken*. Reinbek: Rowohlt.
- Lewin, K. (1971). *Der Übergang von der aristotelischen zur galileischen Denkweise in Biologie und Psychologie*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Maekawa, M. & Hasegawa, Y. (1963). Studies on Jigoro Kano – Significance of his ideals of physical education and Judo. *Bulletin of the Association for the Scientific Studies on Judo, Kodokan. Report II*, pp. 1 – 12. Tokyo: Kodokan.
- Miculincer, M. (1997). Adult attachment style and information processing: individual differences in curiosity and cognitive closure. *Journal of Personality and Social Psychology*, 72, 1217 – 1230.
- Miculincer, M. (1998). Attachment working models and the sense of trust: an exploration of interaction goals and affect regulation. *Journal of Personality and Social Psychology*, 74, 1209 – 1224.
- Miller, S.M. (1990). To see or not to see. Cognitive informational styles in the coping process. In M. Rosenbaum (Hrsg.), *Learned resourcefulness: on coping skills, self – regulation and adaptive behavior*. New York: Springer Press.
- Musashi, M. (1993). *Das Buch der fünf Ringe*. Düsseldorf: Econ Verlag.
- Nadler, A. (1998). Relationship, esteem and achievement perspectives on autonomous and dependent help seeking. In Karabenick (Hrsg.), *Strategic Help Seeking* (pp. 61–93). Mahwah: Lawrence Erlbaum.

- Pinizzotto, A. J. et al. (1997). *In the line of fire: Violence against law enforcement. A study of selected felonious assaults on law enforcement officers*. Washington: United States Department of Justice. Federal Bureau of Investigation. National Institute of Justice.
- Pinizzotto, A. J. & Davis, E. F. (1999). Offenders' perceptual shorthand. What messages are law enforcement officers sending to offenders? *FBI – Law Enforcement Bulletin*, 68, 6, 1-4.
- Ressler, R. K., Burgess, A. W. & Douglas, J. E. et al. (1983). Rape and rape – murder: one offender and twelve victims. *American Journal of Psychiatry*, 140, 1, January 1976, 137–141.
- Sessar, K. A., Baumann, U. & Müller, J. (1980). *Polizeibeamte als Opfer vorsätzlicher Tötungen*. Wiesbaden: Bundeskriminalamt.
- Siebert, A. (1996). *Erfolgreich Krisen bewältigen. Anleitung zum Überleben*. München: Hugendubel.
- Waddington, C. H. (1957). *The Strategy Of The Genes. A Discussion of Some Aspects of Theoretical Biology*. London: George Allen & Unwin Ltd.



## Kriminalitätsfurcht und Strafhärte:

Es kommt darauf an, ob man an "Prävention durch Strafe" glaubt

*Ute Gabriel & Colette Marti*

*Institut für Psychologie der Universität Bern, Schweiz*

Ob die Forderung nach harten Strafen Ausdruck von Kriminalitätsfurcht ist, ist eine in der kriminologischen Forschung seit langem diskutierte Frage. Die diesbezüglichen Forschungsergebnisse sind allerdings ambivalent: Manche Untersuchungen zeigen, dass die Furchtsameren härtere Strafen fordern (z.B. Sprott & Doob, 1997), in anderen Studien hingegen zeigt sich kein Zusammenhang zwischen Kriminalitätsfurcht und Punitivität (z.B. Ouimet & Coyle, 1991). Ziel dieses Beitrages ist es nun, nicht für oder gegen diesen Zusammenhang zu argumentieren, sondern ihn mittels der Unterscheidung zwischen einer verhaltenskontrollierenden und einer ausgleichenden Wirkung von Strafen (z.B. Vidmar & Miller, 1982) zu differenzieren: Wenn eine Person davon überzeugt ist, mittels Bestrafungen könne das *zukünftige Verhalten* des Täters (und Anderer) *kontrolliert* werden, dann könnte Furcht tatsächlich ein Grund dafür sein, vergleichsweise harte Strafen zu fordern und auf diese Weise die Quelle der Furcht zu verringern. Ist eine Person dagegen davon überzeugt, dass das Geschehene durch eine Bestrafung *ausgeglichen* werden könnte, dann sollte sie darum bemüht sein, die Höhe der geforderten Strafe der Tat anzupassen. Ihre persönliche Furcht davor, selbst Opfer zu werden, sollte hierauf keinen (unmittelbaren) Einfluss nehmen. Per Fragebögen wurden 236 Personen u.a. zu ihrer Kriminalitätsfurcht und ihrer Einstellung gegenüber Strafen befragt. Punitivität wurde anhand von Fallvignetten erhoben. Erwartungskonform ist der Zusammenhang zwischen Kriminalitätsfurcht und Strafhärte in der Gruppe der Personen, die von einer verhaltenskontrollierenden Wirkung einer Bestrafung überzeugt sind, signifikant höher als bei den Personen, die von einer ausgleichenden Wirkung überzeugt sind. Strafhärte kann demnach insbesondere dann ein Ausdruck von Kriminalitätsfurcht sein, wenn man an die präventive Wirkung von Strafen glaubt.

### **Einleitung**

Im Rahmen der Forschung zu Strafreaktionen werden von Rechts- und Sozialwissenschaftlern verschiedene Erklärungsansätze verfolgt. Einerseits werden Strafreaktionen z.B. in Abhängigkeit von Merkmalen des abweichenden Verhaltens (z.B. die Art des Verhaltens, ob jemand fahrlässig oder vorsätzlich gehandelt hat, welcher Schaden entstanden ist usw.) oder des Täters (Geschlecht, Vorstrafen usw.) betrachtet (z.B. Rossi, Simpson & Miller, 1985). Andererseits werden Strafreaktionen in Abhängigkeit von Merkmalen der ggf. strafenden Person betrachtet, wie z.B. Autoritarismus (z.B. Feather, 1996), alltagstheoretischen Annahmen über Kriminalitätsursachen (z.B. Oswald & Bilsky, 1991) oder Geschlecht (z.B. Hupfeld, 2000). Innerhalb dieser letztgenannten Perspektive existiert nun eine Argumentationslinie, die Strafbedürfnisse in einen Zusammenhang mit Kriminalitätsfurcht bringt. So wird verschiedentlich davon ausgegangen, dass Personen aus der Furcht heraus, selbst Opfer von Kriminalität zu werden, eine vergleichsweise härtere Bestrafung der Täter verlangen (z.B. Sheley, 1985; Tyler, Boeckmann, Smith & Huo, 1997, p.108). Diese Annahme lässt sich anhand empirischer Ergebnisse (z.B. Boers & Sessar, 1991; Langworthy & Whitehead, 1986; Ouimet & Coyle, 1991; Sprott & Doob, 1997) allerdings weder *eindeutig* belegen noch verwerfen, wenn auch verschiedene empirische Studien zumindest geringe bis mittlere Zusammenhänge zwischen Kriminalitätsfurcht und Strafbedürfnissen finden.

Ziel dieses Beitrages ist es nun nicht, für oder gegen diesen häufig postulierten Zusammenhang zu argumentieren, sondern ihn anhand der *Wirkung, die eine Person mittels Strafe zu erzielen glaubt*, zu differenzieren. Denn wenn *aus Furcht heraus* zu strafen einen Sinn ergeben soll bzw. Strafen in dieser Weise funktional sein soll, dann ja nur, wenn die strafende Person überzeugt ist, oder doch zumindest hofft, mit einer Strafe die Quelle ihrer Furcht in einer ihr zuträglichen Weise zu beeinflussen.

Im Folgenden wird zunächst Kriminalitätsfurcht als das zur Disposition stehende Motiv strafharter Reaktionen näher betrachtet. Daran an schliesst sich eine Diskussion von Strafe und den Strafzielen "Verhaltenskontrolle" und "Ausgleich". Hierbei soll geprüft werden, inwieweit diese beiden Strafziele ihrer hier vorgeschlagenen *vermittelnden Rolle* zwischen Kriminalitätsfurcht und strafender Reaktion gerecht werden können. In einem zweiten Teil werden die Ergebnisse einer Fragebogenuntersuchung vorgestellt, anhand derer diese Überlegungen überprüft werden.

### **Das Motiv: Kriminalitätsfurcht**

Kriminalitätsfurcht beschreibt die Furcht einer Person, *selbst* Opfer einer Straftat zu werden. Dies ist abzugrenzen von einer allgemeinen Wahrnehmung, die Kriminalität sei gestiegen oder von einer diesbezüglichen Besorgnis (Skogan, 1993). Weiter wird Kriminalitätsfurcht als ein Personenmerkmal verstanden, nämlich als die Tendenz, in relevanten Situationen mehr oder weniger furchtsam zu reagieren. Kriminalitätsfurcht wird verschiedentlich als multi-facettiert konzipiert (z.B. Bilsky, 1993; Boers, 1991; Hale, 1996). Ein solcher Ansatz, dem auch die weiteren Ausführungen folgen, unterscheidet eine affektiv-evaluative, eine kognitive und eine motivationale Komponente von Kriminalitätsfurcht (vgl. Gabriel, 1998, S. 20ff; Greve, 1998). Die affektiv-evaluative Komponente beschreibt dabei das *Furchterleben*, die kognitive Komponente die *Bedrohungswahrnehmung* (eine ausführliche Diskussion bzgl. dieser beiden Konstituenten der Furcht findet sich in Greve, 1998). Die dritte Seite der Furcht, die motivationale Seite, umfaßt mehr oder weniger autonome Reaktionen des vegetativen Systems, der Mimik, Gestik und des übrigen Verhaltensapparates aber auch gezieltes, eine Bedrohung antizipierendes Vermeidungs- und Schutzverhalten. "Sich furchtsam zu verhalten" kann somit als ein Indikator von Furcht gelten. Allerdings liegt hierbei das Problem gerade darin, zu entscheiden, ob beispielsweise jemand, der die Tür verriegelt, dies tatsächlich aus Kriminalitätsfurcht tut oder schlicht aus Gewohnheit. Die zu beantwortende Frage lautet dann, ob denn "hohe Bestrafung fordern" in diesem Sinne "Furcht"-Verhalten darstellt.

### **Die Vermittler: Überzeugungen hinsichtlich der Wirkung von Strafen**

Jemanden zu bestrafen beschreibt im Allgemeinen eine *Reaktion* einer Person (der strafenden Person) auf das normverletzende Verhalten einer anderen Person (der zu bestrafenden Person). Ganz egal, welche Ziele oder Absichten die strafende Person mit ihrer Reaktion außerdem verfolgt, zielt eine Strafe per Definition immer darauf ab, der zu bestrafenden Person ein wie auch immer geartetes "Übel" zuzufügen.

Dem in unserer Kultur vorherrschenden Verständnis entsprechend, muss das Verhalten einer Person (und ggf. dessen Konsequenzen) als *schuldhaft normwidrig* anzusehen sein, damit Strafe überhaupt eine mögliche (akzeptierte) Reaktionsform darstellt. Strafe impliziert also eine Schuldzuschreibung, wohingegen eine Schuldzuschreibung nicht notwendigerweise mit einer strafenden Reaktion einhergehen muss. Jemanden zu bestrafen ist *eine* Möglichkeit auf schuldhaft normwidriges Verhalten zu reagieren, aber bei weitem nicht die einzige.

Ob eine Person diese Möglichkeit ergreift, kann von verschiedenen Dingen abhängen, wie z.B. der wahrgenommenen Wichtigkeit der verletzten Norm, dem Ausmaß des entstandenen Schadens oder aber auch ganz generell von der Einstellung, die die strafende Person gegenüber Strafen hat. Hierbei sollte gelten, dass Menschen, die Strafen gegenüber positiv einge-



stellt sind, eher bereit sind, Strafe zu fordern, als Personen, die eine negative Haltung gegenüber Strafen haben.

Ein Aspekt dieser Einstellung gegenüber Strafen sind die *Überzeugungen* einer Person hinsichtlich der Wirkung, die Strafen zukommt. In Anlehnung an Vidmar und Miller (1982, vgl. auch Graham, Weiner & Zucker, 1997) lassen sich innerhalb der *erwünschten* Wirkannahmen Annahmen einer *kontrollierenden* Wirkung von Annahmen einer *ausgleichenden* Wirkung unterscheiden.

Zu den Annahmen einer *ausgleichenden* Wirkung von Strafen gehören Begriffe wie "ausgleichende Gerechtigkeit", "Sühne" und "Vergeltung". Damit eine Strafe ausgleichend wirken kann, muss sie so konzipiert sein, dass sie gerade so viel wiegt, wie die Schuld, die eine Person auf sich geladen hat (vgl. z.B. Montada, 1989). Die Wirkung der Strafe steht in umgekehrt u-förmiger Relation zur Strafhärte: Sowohl eine zu milde als auch eine zu harte Strafe verfehlt ihr Ziel (McFatter, 1982). Divergente Strafforderungen können hier Resultat dessen sein, dass (a) Personen Strafen (wie z.B. gemeinnützige Arbeit) als unterschiedlich hart bzw. mild bewerten (Sebba & Nathan, 1984) bzw. (b) die (Straf-)Sensibilität des Täters mit in ihre Strafbewertungen einbeziehen und sich in dieser voneinander unterscheiden, und/oder (c) dass Personen eine Tat (wie z.B. Sachbeschädigung) als unterschiedlich schwerwiegend ansehen.

*Wie könnte eine ausgleichende Wirkung nun im Hinblick auf Kriminalitätsfurcht von Nutzen sein?* Unmittelbar, so möchten wir behaupten, überhaupt nicht. Mittelbar ließe sich aber ein Bezug herstellen, indem man die affektiv-bewertende Komponente der Kriminalitätsfurcht mit der Schuldkomponente des zu bestrafenden Verhaltens in Verbindung setzt. Der Zusammenhang zwischen Kriminalitätsfurcht und Strafhärte wäre dann aber nicht der, dass Personen *aus* ihrer *Furcht* heraus strafen, sondern, dass sie dieselben Verhaltensweisen sowohl wenn sie gegen sie selbst gerichtet wären (affektiv-evaluative Komponente der Kriminalitätsfurcht) als auch wenn sie *nicht* gegen sie gerichtet wären (im Rahmen einer "neutralen" Schuld-schwerebewertung) in Folge negativerer Bewertung als schwerwiegend erachten. Zum einen muss das nicht notwendigerweise der Fall sein, und zum anderen stellt die affektiv-evaluative Komponente eben nur *eine* Komponente der Kriminalitätsfurcht dar, so wie auch die Bewertung der Schuldschwere weitere Facetten umfassen kann (bzw. sollte).

Annahmen einer *kontrollierenden* Wirkung von Strafe beziehen sich sowohl auf das zukünftige Verhalten des Täters selbst, als auch auf das zukünftige Verhalten anderer, sofern diese Kenntnis der Bestrafung des Täters haben. Zu strafen dient hier vor allem dem Aufbau von *Straferwartungen*, welche neben der erwarteten Strafhöhe wesentlich von der Wahrscheinlichkeit, dass ein entsprechendes Verhalten tatsächlich geahndet werden wird, bestimmt wird. Sowohl die delinquente Person als auch andere, "beistehende" Personen sollen auf diesem Wege dazu gebracht werden, aus Furcht vor einer Bestrafung normwidriges Verhalten zu unterlassen. Die angenommene präventive Wirkung einer Strafe steigt mit der Höhe der Strafe monoton an (McFatter, 1982).

*Wie könnte eine kontrollierende Wirkung nun im Hinblick auf Kriminalitätsfurcht von Nutzen sein?* Hier zeigt sich eine Funktionalität im Hinblick auf Kriminalitätsfurcht unmittelbar: Wenn Strafen präventive Wirkung zukommt, dann lässt sich die wahrgenommene Bedrohung durch kriminelles Verhalten Anderer mittels Strafen mildern. Wenn gilt, dass der präventive Effekt mit der Strafhärte ansteigt, dann könnte sich Kriminalitätsfurcht tatsächlich in der Forderung von hohen Strafen niederschlagen.

Zusammenfassend lässt sich diese Argumentation wie folgt darstellen:

Unter der Voraussetzung, dass eine Person der Überzeugung ist, mittels Bestrafungen könnte das zukünftige Verhalten sowohl des Täters als auch Anderer in kontrollierender (präventiver) Weise beeinflusst werden, kann die eigene Kriminalitätsfurcht ein Motiv für vergleichsweise harte Strafforderungen darstellen.

Unter der Voraussetzung, dass eine Person der Überzeugung ist, mittels (gerechter) Bestrafungen ließe sich die Schuld, die der Täter auf sich geladen hat, ausgleichen, können vergleichsweise harte Strafen *nicht* Ausdruck der Kriminalitätsfurcht der Person sein. Vergleichsweise geringe empirische Zusammenhänge sind aber zu erwarten, da Deliktsbewertungen sowohl einen Aspekt der Kriminalitätsfurcht als auch einen Aspekt der Bewertung der Schuldschwere darstellen.

Empirisch müsste sich dies folgendermaßen niederschlagen:

1. Wenn Personen von einer verhaltenskontrollierenden Wirkung von Strafen überzeugt sind, dann korreliert Kriminalitätsfurcht stärker positiv mit Punitivität als wenn sie davon nicht überzeugt sind, bzw.
2. wenn Personen von einer verhaltenskontrollierenden Wirkung von Strafen, *aber nicht* von einer ausgleichenden Wirkung von Strafen überzeugt sind, dann korreliert Kriminalitätsfurcht stärker mit Punitivität als wenn Personen von einer ausgleichenden Wirkung, *nicht aber* von einer verhaltenskontrollierenden Wirkung überzeugt sind.

Diese Hypothesen werden anhand einer Fragebogenuntersuchung überprüft, deren Inhalt und Ablauf im Weiteren dargestellt wird.

## **Datenerhebung und Stichprobe**

Die Datenerhebung erfolgte im Sommer 2000 mittels eines Fragebogens (Marti & Greve, 2000). Um in der Stichprobe möglichst verschiedene Bevölkerungsgruppen abzubilden, wurde der Fragebogen in unterschiedlichen Kursen und Vorlesungen der Volkshochschule Bern und Zürich ausgeteilt. Weiter wurden Fragebögen in verschiedenen Schulklassen einer privaten Weiterbildungsschule sowie an einer Höheren Schule für Wirtschaft und Verwaltung verteilt. Teilnehmer eines Beschäftigungsprogramms für Arbeitslose (IPA) füllten die Fragebögen im Rahmen eines Bildungstages aus. Schliesslich wurden Fragebögen per Schneeballsystem über Bekannte abgegeben. Zusammen mit dem Fragebogen wurde jeweils ein voradressiertes und vorfrankiertes Kuvert ausgehändigt. Von den insgesamt 575 Fragebogenexemplaren, die verteilt wurden, kamen 236 verwertbare Fragebögen zurück, was einer Ausschöpfungsrate von 41 % entspricht.

In die Auswertung gehen die Angaben von 236 Personen im Alter zwischen 18 und 86 Jahren ein. Das Altersmittel liegt bei 43 Jahren. Der Frauenanteil beträgt 50.8 %.

## **Erhebungsinstrument**

Der Fragebogen umfasst 18 Seiten. Nur ein Teil der im Rahmen dieser Befragung erhobenen Variablen finden bei der Überprüfung der hier interessierenden Hypothesen Verwendung. Die Operationalisierung dieser Variablen wird im weiteren beschrieben.

### *Kriminalitätsfurcht*

Die affektiv-evaluative Komponente der Kriminalitätsfurcht wurde als *Furchthäufigkeit* operationalisiert, nämlich als die Häufigkeit, mit der die Person Furcht empfindet hinsichtlich neun verschiedener Deliktstypen (Handtaschenraub, Diebstahl, Betrug, Gewalt mit/ohne Waffe, Raub, Einbruch, sexuelle Nötigung, Vergewaltigung). Die entsprechenden Fragen lauteten z.B.: *Wie häufig fürchten Sie, mit einer Waffe angegriffen oder verletzt zu werden?* Als Antwortformat wurde eine 5-stufige Ratingskala, die mit den verbalen Ankern *nie, sehr selten, eher selten, eher häufig, sehr häufig* versehen war, vorgegeben. Als Skalenwert wurde der Mittelwert über die neun Fragen berechnet ( $M = 2.33$ ,  $SD = 0.63$ , Cronbachs  $\alpha = 0.85$ ).

Die kognitive Komponente wurde als subjektive Risikoeinschätzung ebenfalls über die o.g. neun Deliktstypen erhoben. Die einleitende Frage lautete: *Für wie wahrscheinlich halten Sie es, dass Ihnen persönlich innerhalb der nächsten zwölf Monate so etwas tatsächlich passiert?* Diese sollte dann jeweils in Bezug auf die verschiedenen Deliktsbeschreibungen (z.B. ... *dass Ihnen die Handtasche geraubt wird*) auf einer 4-stufigen Ratingskala beantwortet werden. Die Ratingskala war mit den verbalen Ankern *sehr unwahrscheinlich, eher unwahrscheinlich, eher wahrscheinlich und sehr wahrscheinlich* versehen. Als Skalenwert wurde der Mittelwert über die neun Fragen berechnet ( $M = 1.76$ ,  $SD = 0.42$ , Cronbachs  $\alpha = 0.81$ ).

Zur Erfassung des Furchtverhaltens wurden im Anschluss an die Aufforderung *Bitte kreuzen Sie nun an, wie häufig Sie persönlich die folgenden vorbeugenden Massnahmen ergreifen, um sich persönlich vor Kriminalität zu schützen* 15 verschiedene Verhaltensweisen vorgegeben, bezüglich der die Personen auf einer 5-stufigen Skala angeben sollten, wie häufig sie die entsprechenden Verhaltensweisen durchführen (*nie, selten, manchmal, häufig, immer*). Eine dieser Verhaltensweisen lautete beispielsweise: *Ich trage etwas bei mir, womit ich mich wehren kann*. Als Skalenwert wurde der Mittelwert über diese 15 Verhaltensweisen berechnet ( $M = 2.57$ ,  $SD = 0.73$ , Cronbachs  $\alpha = 0.86$ ).

Schliesslich wurde Kriminalitätsfurcht mit der sogenannten Standardfrage (*Wie sicher fühlen Sie sich oder würden Sie sich fühlen, wenn Sie in Ihrer Wohngegend nachts draussen alleine sind?*) erfasst, die auf einer 4-stufigen Ratingskala (*sehr sicher, ziemlich sicher, ziemlich unsicher, unsicher*) zu beantworten war ( $M = 2.0$ ,  $SD = 0.70$ ). Die vielfach kritisierte Standardfrage kann als ein globales Mass zur Erfassung von Kriminalitätsfurcht angesehen werden, welches allerdings den Verhaltensaspekt vergleichsweise stark widerspiegelt (vgl. Hale, 1996).

Die vier Maße der Kriminalitätsfurcht korrelieren zwischen  $r = .49$  (Verhalten & subjektive Risikoeinschätzung) und  $r = .71$  (Furchthäufigkeit & subjektive Risikoeinschätzung) miteinander.

### *Strafhärte*

Jede Person musste insgesamt acht Fallbeschreibungen beurteilen. Diese hatten alle ein Körperverletzungsdelikt zum Thema, welches jeweils die Konsequenz nach sich zog, dass das Opfer mehrmals ambulant behandelt werden musste. Eine solche Vignette lautete beispielsweise:

*Thomas F. fährt im Auto durch die Stadt und programmiert gleichzeitig seinen CD-Player. Dabei entgeht ihm, dass der Autolenker vor ihm bremst, und so fährt Thomas F. direkt in die-  
sen hinein. Der Vordermann muss mehrmals ambulant behandelt werden.*

Die Befragten waren aufgefordert, hinsichtlich jeder der acht Fallbeschreibungen ihre Strafreaktion auszudrücken. Als einzige Reaktionsform stand ihnen dabei die Geldstrafe zur Verfü-

gung. Den Befragten wurde zunächst die Bedeutung des Begriffes "Tagessatz" an einem Beispiel erläutert. Anschliessend sollten sie für jede der Geschichten eine Strafe in Form einer Anzahl von Tagessätzen angeben, die ihrer Reaktion entsprechen würde. Wenn sie keine Strafe verhängen würden, sollten die Befragten "0 Tagessätze" notieren. Mit diesem Vorgehen wird das Skalierungsproblem von Sanktionen somit in einer Weise gelöst, die zumindest ein Ordinalskalenniveau garantiert. Die resultierenden Werte können mit Einschränkung sogar als intervallskaliert betrachtet werden. Als Skalenwert wurden die acht Einzelwerte aufsummiert ( $M = 344.60$ ,  $SD = 360.52$ , Cronbachs  $\alpha = 0.78$ ). Die Werte sind extrem schief verteilt (Schiefe = 2.13, Median = 205).

#### *Ausgleichende vs. verhaltenskontrollierende Wirkung von Strafen*

Den Personen wurden verschiedene Wirkungen, die Strafen haben könnten, mit dem einleitenden Satz vorgegeben: *Meiner Meinung nach kann Strafe vor allem ...*. Je fünf Aussagen repräsentierten dabei eine ausgleichende (Itembeispiel: ... *das geschehene Unrecht wieder ausgleichen.*) bzw. eine verhaltenskontrollierende Wirkung von Strafen (Itembeispiel: ...*bewirken, dass der Straftäter aus Angst vor Strafe zukünftige kriminelle Handlungen unterlässt*). Zu jeder dieser zehn Aussagen sollten die Personen mittels einer 4-stufigen Ratingskala den Grad ihrer Zustimmung äußern. Als Skalenwert wurde jeweils der Mittelwert über die fünf entsprechenden Aussagen berechnet (Skala *Ausgleich*:  $M = 2.40$ ,  $SD = 0.77$ , Cronbachs  $\alpha = 0.84$ ; Skala *Verhaltenskontrolle*:  $M = 2.56$ ,  $SD = 0.74$ , Cronbachs  $\alpha = .87$ ). Die Werte der beiden Skalen korrelieren zu  $r = 0.45$  miteinander, was bedeutet, dass Personen eher *sowohl* von einer verhaltenskontrollierenden *als auch* einer ausgleichenden Wirkung von Strafen überzeugt sind, bzw. *von beiden nicht* überzeugt sind.

## **Ergebnisse**

Um zu prüfen, ob Kriminalitätsfurcht dann stärker mit Strafhärte korreliert, wenn Personen von einer verhaltenskontrollierenden Wirkung von Strafen überzeugt sind (Hypothese (1)) wurde die Stichprobe anhand des (theoretischen) Skalenmittelwertes in zwei Gruppen aufgeteilt. Die erste Gruppe ( $VK^+$ ) bilden diejenigen Personen, die den Aussagen der Skala *Verhaltenskontrolle* im Mittel zugestimmt haben ( $M > 2.5$ ,  $n = 121$ ), die zweite Gruppe ( $VK^-$ ) bilden diejenigen Personen, die den Aussagen der Skala *Verhaltenskontrolle* im Mittel *nicht* zugestimmt haben ( $M < 2.5$ ,  $n = 112$ ).

Die beiden Gruppen wurden bzgl. Strafhärte sowie den Indikatoren der Kriminalitätsfurcht miteinander verglichen. Demnach zeigt die  $VK^+$ -Gruppe tendenziell mehr Furchtverhalten ( $t = -1.84$ ,  $p = .067$ ) sowie ein stärkeres Unsicherheitsgefühl bei der Standardfrage ( $t = -1.67$ ,  $p = .096$ ).

Für jede der beiden Gruppen wurden die Korrelationskoeffizienten zwischen Strafhärte und den Indikatoren der Kriminalitätsfurcht berechnet und miteinander verglichen (vgl. Bortz, 1993, S. 202). Tabelle 1 zeigt die Korrelationen sowie die z-Werte der Differenzen der jeweiligen Korrelationspaare. Aufgrund der schiefen Verteilung der Strafhärte-Werte wurde Spearmans rho berechnet.

Tabelle 1: Zusammenhänge zwischen Strafhärte und Kriminalitätsfurcht in Abhängigkeit von der Verhaltenskontrollüberzeugung

Überzeugung:		Furchthäufigkeit	subjektives Risiko	Furchtverhalten	Standardfrage
VK <sup>+</sup>	Spearman's Rho	<b>.14</b>	<b>.12</b>	<b>.19</b>	<b>.00</b>
(n = 121)	p (zweiseitig)	.15	.20	.04*	.98
VK <sup>-</sup>	Spearman's Rho	<b>-.16</b>	<b>.02</b>	<b>.03</b>	<b>.06</b>
(n = 112)	p (zweiseitig)	.09	.86	.76	.54
Vergleich:	z-Werte	<b>2.25</b>	<b>0.76</b>	<b>1.23</b>	<b>-0.43</b>
	p (einseitig)	.01*	.22	.11	.33

Anmerkung: \* =  $p \leq .05$ ; \*\* =  $p \leq .01$ ; \*\*\* =  $p \leq .001$

Die Hypothese bestätigt sich lediglich beim Vergleich der Zusammenhänge zwischen Furchthäufigkeit und Strafhärte: Die Korrelation zwischen Furchthäufigkeit und Strafhärte unterscheidet sich signifikant in Abhängigkeit von der Überzeugung, Strafe habe eine verhaltenskontrollierende Wirkung. Hierzu trägt bei, dass in der VK<sup>-</sup>-Gruppe Kriminalitätsfurcht mit Strafhärte *negativ* korreliert, diese Personen also mit zunehmender Furchthäufigkeit der Tendenz nach *strafmilder* reagieren. Immerhin in die richtige Richtung weisen die Ergebnisse hinsichtlich der subjektiven Risikoeinschätzung sowie des Furchtverhaltens: Die Zusammenhänge zwischen den Kriminalitätsfurchtindikatoren und Strafhärte sind in der VK<sup>+</sup>-Gruppe (wenig) höher als in der VK<sup>-</sup>-Gruppe.

Zur Prüfung der Annahme, dass Kriminalitätsfurcht dann stärker mit Kriminalitätsfurcht korreliert, wenn Personen von einer verhaltenskontrollierenden Wirkung von Strafen, *aber nicht* von einer ausgleichenden Wirkung von Strafen überzeugt sind (Hypothese (2)) wurde die Stichprobe anhand des Skalenmittelwertes der Skala *Ausgleich* in zwei Gruppen geteilt. Die erste Gruppe (A<sup>+</sup>) umfasst diejenigen Personen, die den Aussagen der Skala *Ausgleich* im Mittel zugestimmt haben ( $M > 2.5$ ,  $n = 89$ ), die zweite Gruppe (A<sup>-</sup>) umfasst diejenigen Personen, die den Aussagen der Skala *Ausgleich* im Mittel *nicht* zugestimmt haben ( $M < 2.5$ ,  $n = 145$ ). Kreuztabelliert man diese Gruppeneinteilung mit der Gruppeneinteilung bzgl. Verhaltenskontrolle resultieren vier Zellen, von denen zwei im weiteren miteinander verglichen werden sollen:

- (1) Personen, die zu VK<sup>+</sup> und A<sup>-</sup> gehören, also Personen, die von der verhaltenskontrollierenden Wirkung von Strafe überzeugt sind, nicht aber von der ausgleichenden und
- (2) Personen, die zu VK<sup>-</sup> und A<sup>+</sup> gehören, also Personen, die von der ausgleichenden Wirkung von Strafen überzeugt sind, nicht aber von einer verhaltenskontrollierenden.

Hierbei zeigt sich, dass der größte Anteil der Befragten ( $n = 84$ , das entspricht 35.6 % der Stichprobe) *weder* von einer verhaltenskontrollierenden Wirkung von Strafe *noch* von einer ausgleichenden Wirkung überzeugt ist, und so der Kategorie VK<sup>-</sup> & A<sup>+</sup> lediglich 26 Personen angehören.

Die beiden ausgewählten Gruppen wurden auf Unterschiede bzgl. Kriminalitätsfurcht und Strafhärte geprüft. Hier zeigen sich keine bedeutsamen Unterschiede (alle  $p \geq .18$ ). Analog zur Überprüfung von Hypothese 1 wurde für jede der beiden Gruppen die Korrelationskoeffizienten (Spearman's rho) zwischen Strafhärte und den Indikatoren der

Kriminalitätsfurcht berechnet und miteinander verglichen. Die Ergebnisse sind in Tabelle 2 dargestellt.

Tabelle 2: Zusammenhänge zwischen Strafhärte und Kriminalitätsfurcht in Abhängigkeit der Verhaltenskontrollüberzeugung *und* der Ausgleichsüberzeugung

Überzeugung:		Furchthäufigkeit	subjektives Risiko	Furchtverhalten	Standardfrage
VK <sup>+</sup> & A <sup>-</sup> (n = 59)	Spearman's rho	<b>.31</b>	<b>.21</b>	<b>.21</b>	<b>.09</b>
	p (zweiseitig)	.02*	.12	.12	.49
VK <sup>-</sup> & A <sup>+</sup> (n = 26)	Spearman's rho	<b>-.24</b>	<b>-.21</b>	<b>-.08</b>	<b>-.31</b>
	p (zweiseitig)	.24	.30	.70	.13
Vergleich:	z-Werte	<b>2.37</b>	<b>1.79</b>	<b>1.23</b>	<b>1.73</b>
	p (einseitig)	.01**	.04*	.11	.04*

Anmerkung: \* =  $p \leq .05$ ; \*\* =  $p \leq .01$ ; \*\*\* =  $p \leq .001$

Bei dieser Gruppenauswahl (wobei jede Gruppe eine Teilmenge der in der ersten Berechnung verwendeten Gruppen darstellt) unterscheiden sich die Zusammenhänge zwischen den beiden Gruppen bzgl. drei der vier Kriminalitätsfurchtsindikatoren signifikant voneinander. Lediglich der Unterschied zwischen Furchtverhalten und Strafhärte erreicht keine statistische Signifikanz. Auch hier gilt, dass der Unterschied zwischen den Korrelationskoeffizienten durch die *Gegenläufigkeit* der Zusammenhänge zustande kommt und nicht (wie erwartet) durch die Höhe der (positiven) Zusammenhänge.

## Resümee

Ausgangspunkt war die Frage, ob Kriminalitätsfurcht als ein Motiv für straharte Reaktionen auf normabweichendes Verhalten angesehen werden kann. Es wurde dafür argumentiert, dass dies mit der Einschränkung, die Person sei von einer verhaltenskontrollierenden Wirkung von Strafen überzeugt, gelten kann.

Die empirische Prüfung zeigt, dass der erwartete Effekt unterschiedlicher Korrelationsmuster insbesondere dann auftritt, wenn die Überzeugung bzgl. verhaltenskontrollierender *und* ausgleichender Wirkung von Strafen berücksichtigt wird. Bei Personen, die von einer verhaltenskontrollierenden Wirkung von Strafen, *aber nicht* von einer ausgleichenden Wirkung von Strafen überzeugt sind, unterscheiden sich die Zusammenhänge zwischen drei der vier Indikatoren der Kriminalitätsfurcht und Strafhärte signifikant von denjenigen bei Personen, die von einer ausgleichenden Wirkung, *nicht aber* von einer verhaltenskontrollierenden Wirkung überzeugt sind. Diese Unterschiede gehen mit einem Richtungswechsel der Korrelationskoeffizienten einher: Bei Personen, die von einer ausgleichenden Wirkung, *nicht aber* von einer verhaltenskontrollierenden Wirkung überzeugt sind, zeigt sich ein schwach negativer Zusammenhang zwischen Kriminalitätsfurchtindikatoren und Strafhärte. Dieser Befund sollte zunächst nicht überbewertet werden, da die diesen (schwachen) Effekt produzierende Stichprobe vergleichsweise gering ist (n = 26).

Festgehalten werden kann, dass Kriminalitätsfurcht eher im Kontext präventiver Strafwirkung einen Ausdruck in der *Strafhärte* findet als im Kontext ausgleichender Strafwirkung. Die Höhe, der unter der Voraussetzung "Strafe kann Verhalten kontrollieren" gefundenen Zusammenhänge, verweist allerdings darauf, dass sich in der Unterschiedlichkeit der Strafhärte weit mehr als Unterschiede in der Furcht, selbst Opfer einer Straftat zu werden, widerspiegelt.

## Literatur

- Bilsky, W. (1993). Blanks and open questions in survey research on fear of crime. In W. Bilsky, C. Pfeiffer & P. Wetzels (Hrsg.), *Fear of crime and criminal victimization. Interdisziplinäre Beiträge zur kriminologischen Forschung, Neue Folge, 3* (S. 9-19). Stuttgart: Ferdinand Enke Verlag.
- Boers, K. (1991). *Kriminalitätsfurcht*. Pfaffenweiler: Centaurus.
- Boers, K. & Sessar, K. (1991). Do people really want punishment? On the relationship between acceptance of restitution, needs for punishment, and fear of crime. In K. Sessar & H.-J. Kerner (Hrsg.), *Developments in Crime and Crime Control Research. German Studies on Victims, Offenders, and the Public* (S. 126-149). New York: Springer.
- Bortz, J. (1993). *Statistik für Sozialwissenschaftler*. Berlin: Springer.
- Feather, N. T. (1996). Reactions to penalties for an offense in relation to authoritarianism, values, perceived responsibility, perceived seriousness, and deservingness. *Journal of Personality and Social Psychology, 71*(3), 571-587.
- Gabriel, U. (1998). *Furcht und Strafe. Kriminalitätsfurcht, Kontrollüberzeugungen und Strafforderungen in Abhängigkeit von der Erfahrung krimineller Viktimisierung*. Baden-Baden: Nomos.
- Graham, S., Weiner, B., & Zucker (1997). An attributional analysis of punishment goals and public reactions to O. J. Simpson. *Personality and Social Psychology Bulletin, 23*, 331-346.
- Greve, W. (1998). Fear of crime among the elderly: Foresight, not fright. *International Review of Victimology, 5*, 277-309.
- Hale, C. (1996). Fear of crime: A review of the literature. *International Review of Victimology, 4*, 79-150.
- Hupfeld, J. (2000). Frauen, eine Gefahr für das Strafrechtssystem? Geschlecht und Reaktionen auf Kriminalität. In Abt. für die Gleichstellung von Frauen und Männern der Universität Bern (Hrsg.), *Frauen im Recht: Kindsmörderinnen und Richterinnen* (S. 43-59). Bern: eFeF.
- Langworthy, R. H. & Whitehead, J. T. (1986). Liberalism and fear as explanations of punitiveness. *Criminology, 24*, 3, 575-591.
- Marti, C. & Greve, W. (2000). *Fragebogen: Meinungen, Erfahrungen und Befürchtungen zur Kriminalität*. Bern: Universität Bern.
- McFatter, R. M. (1982). Purposes of punishment: effects of utilities of criminal sanctions on perceived appropriateness. *Journal of Applied Psychology, 67*, 255-267.
- Montada, L. (1989). Strafzwecküberlegungen aus psychologischer Sicht. In C. Pfeiffer & M. Oswald (Hrsg.), *Strafzumessung: Empirische Forschung und Strafrechtsdogmatik im Dialog* (S. 137-141). Stuttgart: Enke.
- Oswald, M. E. & Bilsky, W. (1991). Subjektive Theorien über Kriminalitätsursachen und richterliche Schuldzuschreibung. *Monatsschrift für Kriminologie und Strafrechtsreform, 3*, 129-145.
- Ouimet, M., & Coyle, E. J. (1991). Fear of crime and sentencing punitiveness: Comparing the general public and court practitioners. *Canadian Journal of Criminology, 33*, 149-162.
- Rossi, P. H., Simpson, J. E., & Miller, J. L. (1985). Beyond crime seriousness: fitting the punishment to the crime. *Journal of Quantitative Criminology, 1*, 59-90.
- Sebba, L. & Nathan, G. (1984). Further explorations in the scaling of penalties. *The British Journal of Criminology, 23*, 221-247.
- Sheley, J. F. (1985). *America's "Crime Problem": An Introduction to Criminology*. Belmont, CA: Wadsworth.
- Skogan, W. G. (1993). The various meanings of fear. In W. Bilsky, C. Pfeiffer & P. Wetzels (Hrsg.), *Fear of crime and criminal victimization. Interdisziplinäre Beiträge zur kriminologischen Forschung, Neue Folge, 3* (S. 131-140). Stuttgart: Ferdinand Enke Verlag.

- Sprott, J. B., & Doob, A. N. (1997). Fear, victimization, and attitudes to sentencing, the courts, and the police. *Canadian Journal of Criminology*, 39, 275-291.
- Tyler, T. R., Boeckmann, R. J., Smith, H. J., & Huo, Y. J. (1997). *Social Justice in a Diverse Society*. Boulder, CO: Westview Press.
- Vidmar, N. & Miller, D. T. (1982). Social psychological motives underlying punishment reactions. In H. Hiebsch (Hrsg.), *Social Psychology* (S. 195-202). Amsterdam: North-Holland Publishing.



## Empirisch fundierte Prognosestellung im Maßregelvollzug nach § 63 StGB (EFP-63)

*Lutz Gretenkord*

*Klinik für forensische Psychiatrie, Haina*

Neulich habe ich ein Autohaus betreten, um mich nach einem neueren Auto umzusehen. Ich fragte den Verkäufer, wie viel er mir für mein altes geben würde. Er setzte sich an seinen Computer und gab allerhand ein, z.B. den Typ, das Baujahr und die Kilometerleistung. Nach einiger Zeit eröffnete er mir, daß mein Auto noch einen Wert von 550 DM habe.

Es gibt Aufstellungen, in denen die in letzter Zeit gezahlten durchschnittlichen Gebrauchtwagenpreise aufgeführt sind. Niemand gibt Ihnen die Garantie, daß Sie Ihr Fahrzeug tatsächlich zu diesem Preis verkaufen können. Dennoch werden solche Listen von vielen Menschen als nützliche Information angesehen, so daß man sich die Mühe macht, sie zu erstellen.

Wenn Sie als Gutachter zum Risiko des Rückfalls bei einem Maßregelvollzugspatienten Stellung nehmen sollen, haben Sie nicht die Möglichkeit, in entsprechenden Listen nachzuschlagen. Für Aufstellungen, aus denen z.B. hervorgeht, wie viele Patienten mit der Diagnose Persönlichkeitsstörung, die im Maßregelvollzug untergebracht waren, rückfällig geworden sind, gibt es offenbar nicht genug Interessenten.

Die Abneigung gegen solche "Prognosetafeln" – oder wie immer man sie nennen will – wird oft damit begründet, daß jeder Fall anders sei, daß diese Statistiken über den Einzelfall gar nichts aussagen.

Aber auch jeder Gebrauchtwagen ist einzigartig. Trotzdem hat man einen guten Anhaltspunkt für den Preis, wenn man in einer einschlägigen Liste nachschaut. Ausgehend von dieser Basisrate können dann Abschläge für zerbeulte Kotflügel und Zuschläge etwa für den Fall, daß einer der Vorbesitzer ein Prominenter war, gemacht werden.

Hier wird die Meinung vertreten, daß man auch bei der Prognose künftigen kriminellen Verhaltens analog vorgehen sollte. Es sollte systematisch erfaßt werden, wer rückfällig wird und wer nicht, und mit welchen Faktoren die Rückfälligkeit in Beziehung gesetzt werden kann.

Im folgenden soll eine Möglichkeit dargestellt werden, die Ergebnisse von entsprechenden Untersuchungen so auszuwerten und darzustellen, daß sie in der Praxis genutzt werden können (ausführliche Darstellung und Literaturhinweise s. Gretenkord, 2001).

Dazu werden Daten aus einer eigenen Untersuchung zur Rückfälligkeit herangezogen. In Tabelle 1 sind einige Kennwerte für die Untersuchungsgruppe wiedergegeben.

Tabelle 1: Grunddaten.

<i>Probanden</i>	188 nach § 63 StGB eingewiesene Männer, die 1977-1985 aus der Klinik für gerichtliche Psychiatrie Haina entlassen wurden.
<i>Unterbringungsdauer</i>	M = 8 Jahre (maximal 31 Jahre)
<i>Alter bei Entlassung</i>	M = 39 Jahre
<i>Auszüge aus dem BZR</i>	August 1989
<i>Katamnesezeit</i>	M = 8 Jahre (4-12 Jahre)

Bei der Untersuchung der Rückfallrate für die Gesamtgruppe, erfaßt nach drei Kriterien, ergab sich folgendes:

- 107 der 188 männlichen Patienten (56,9%) hatten nach ihrer Entlassung keinen Eintrag im Bundeszentralregister. Die anderen 81 (43,1%) hatten irgendeinen Eintrag, wobei es sich aber auch um Bagatelldelikte oder etwa den Widerruf der bedingten Entlassung wegen Nichteinhaltung von Bewährungsauflagen handeln kann.
- 56 Patienten (29,8%) kamen erneut in den Freiheitsentzug; entweder wurde eine Freiheitsstrafe verhängt oder es erfolgte der Widerruf der bedingten Entlassung oder es wurde wieder Maßregelvollzug angeordnet.
- Hiervon begingen 25 Patienten (13,3%) nach ihrer Entlassung Gewalt- und/oder Sexualdelikte, darunter ein Tötungsdelikt.
- Drei dieser Rückfälligen begingen Sexualdelikte ohne Anwendung körperlicher Gewalt, demnach wurden 22 Patienten (11,7%) mit einem Gewaltdelikt rückfällig.

Diese letztere Gruppe (gravierendstes Kriterium) wird bei den folgenden Analysen herangezogen.

Stellt man Beziehungen zwischen einzelnen Faktoren und der Rückfälligkeit auf, so lassen sich durchaus statistisch signifikante Zusammenhänge ermitteln.

Es wurde versucht, mit multivariaten Auswertungsmethoden Variablen zu finden, die geeignet sind, dem praktisch tätigen Prognostiker im Maßregelvollzug empirisch gesicherte Hilfestellung zu geben.

Eine der größten Schwierigkeiten dabei ist, die "besten" Variablen für eine solche Auswertung zu finden. Folgende Voraussetzungen sollten gegeben sein:

- Die Variablen müssen mit dem Rückfallkriterium einen statistisch bedeutsamen Zusammenhang aufweisen.
- Die Variablen sollten auch in anderen Untersuchungen einen Zusammenhang mit der Rückfälligkeit aufweisen.
- Die Zusammenhänge der Variablen mit der Rückfälligkeit sollten interpretierbar sein ("face validity").
- Die Definition der Variablen sollte möglichst eindeutig sein.
- Die Variablen sollten ohne zu großen Aufwand ermittelbar sein.
- Die Variablen sollten nicht spezifisch für eine bestimmte Klinik sein.
- Die Variablen sollten nicht spezifisch für einen bestimmten Zeitraum sein.

Unter Berücksichtigung dieser Gesichtspunkte wurden schließlich die folgenden zehn Variablen in die Rechnung aufgenommen:

- Zumindest teilweise Heimunterbringung (mindestens ein Jahr) bis zum 18. Lebensjahr.
- Weniger als ein Jahr versicherungspflichtige Arbeit.
- Das Vorliegen der Diagnose Persönlichkeitsstörung, unabhängig davon, ob noch weitere Diagnosen gestellt wurden. (Dies war die einzige Diagnose, bei der ein signifikanter Zusammenhang mit der Rückfälligkeit gefunden wurde.)
- Rückführbarkeit der Delinquenz auf ein kriminelles Umfeld.
- Jugendrichterliche Maßnahmen (Ermahnungen, Erziehungsmaßnahmen, Fürsorgeerziehung).
- Gewaltdelikt in der Vorgeschichte (Vorstrafe oder MRV wegen eines Gewaltdelictes im weiteren Sinne, also auch gewaltsame Sexual- und Eigentumsdelikte).
- Mindestens eine Entweichung während der Unterbringung.
- Mindestens einmal unerlaubter Alkoholkonsum während der Unterbringung.

- Mindestens zweimal aggressives Verhalten während der Unterbringung, also körperliche Auseinandersetzungen mit Mitpatienten oder Angriff auf das Personal. Die Dichotomisierung "jemals Gewalt versus keine Gewalt" hat sich als weniger trennscharf erwiesen.
- Entlassungsalter. Diese Variable hat sich in zahlreichen Rückfälligkeitserhebungen als relevant erwiesen, zudem ist sie sehr leicht und zuverlässig zu erheben.

Als statistisches Verfahren wurde die "logistische Regression" verwendet. Sie hat den Vorteil, daß man – bezogen auf diese Untersuchung - für jeden Patienten einen Wert ermitteln kann, der die Wahrscheinlichkeit für das Begehen eines Gewaltdelikt nach der Entlassung aus dem Maßregelvollzug schätzt.

Ein solcher Prozentwert ist brauchbarer als solche Angaben wie "die Prognose ist recht gut" oder "mit Wahrscheinlichkeit sind weitere Straftaten zu befürchten". Die Aussage "Dieses Auto kostet 14.000 DM" gibt mehr Information, als wenn es heißt: Dieser Wagen ist "sehr preisgünstig" oder "recht teuer".

Von den zehn nach den oben aufgeführten Kriterien ausgesuchten Variablen, die in die Rechnung eingegangen sind, sind letztlich vier für die Endauswertung übrig geblieben; die anderen sechs Variablen bringen keinen zusätzlichen signifikanten Erkenntnisgewinn.

Diese vier Variablen sind:

- Persönlichkeitsstörung (ja/nein)
- Bereits vor dem Indexdelikt Vorbelastung mit Gewaltdelikt (ja/nein)
- Mindestens zweimal gewalttätig während der Unterbringung (ja/nein)
- Entlassungsalter (in Jahren)

Wie erwähnt, sollten die in der Praxis genutzten Variablen auch in anderen Untersuchungen rückfallrelevant sein. Bei der Auswertung von 46 Studien, in denen Angaben zur Rückfälligkeit von psychisch gestörten Rechtsbrechern gemacht wurden, wurden acht gefunden, die einen Zusammenhang zwischen Diagnosen und der Rückfälligkeit mit einem Gewaltdelikt beinhalteten (Tabelle 2).

Tabelle 2: Literaturlauswertung: Diagnosen und Rückfälligkeit mit einem Gewaltdelikt

Land	Untersuchung	N	GR	Rückfälle	
				häufiger	seltener
USA	Koppin (1990)	109	+	"keine Psychose"	
CAN	Rice, Harris, Lang & Bell (1990)	253	+	Persönlichkeitsstörung	
CAN	Harris, Rice & Cormier (1991)	169	+++	Psychopathie	
CAN	Rice & Harris (1992)	192	0		
CAN	Webster et al. (1994)	618	+	Persönlichkeitsstörung	Schizophrenie
GB	Bailey & MacCulloch (1992)	112	+	Psychopathie	Psychose
GB	Tennent & Way (1984)	617	++	Persönlichkeitsstörung	Schizophrenie
D	Jockusch & Keller (1993)	101	0		
D	Gretenkord (2001)	188	+	Persönlichkeitsstörung	

Anmerkung. GR = Gewaltrückfall: 0 = keine statistisch signifikante Beziehung; +: p<.05; ++: p<.01; +++: p<.001; bei negativer Korrelation entsprechend mit Minuszeichen (nach Gretenkord, 2001, S. 248f.).

In den meisten Studien wurde ein positiver Zusammenhang zwischen dem Vorliegen einer Persönlichkeitsstörung und dem Gewaltrückfall gefunden, wobei in einigen Fällen diese auf eine „Psychopathie“ im Sinne von Hare (1991) eingegrenzt war. Koppin (1990) definierte

diese Variable als „keine Psychose“, wobei es sich bei diesen Probanden überwiegend um Persönlichkeitsgestörte handelt.

Auch die Vorbelastung mit einem Gewaltdelikt ergibt überwiegend positive Zusammenhänge mit dem Gewaltrückfall (Tabelle 3).

Tabelle 3: Literaturlauswertung: Früheres Gewaltdelikt und Rückfälligkeit mit einem Gewaltdelikt

Land	Untersuchung	N	GR	Bemerkungen
GB	Tennent & Way (1984)	617	+++	Gewaltsame Vordelikte
USA	Steadman & Coccozza (1974)	98	n.a.	Legal Dangerousness Score
USA	Thornberry & Jacoby (1979)	414	+	Festnahmen wg. Gewaltdelikt
USA	Koppin (1990)	109	(.06)	Gewalttat vor Indexdelikt
CAN	Rice, Harris, Lang & Bell (1990)	253	+	Akman-Normandeau-Skala
CAN	Harris, Rice & Cormier (1991)	169	+++	Gewaltsames Vordelikt
CAN	Rice & Harris (1992)	192	0	Früheres Gewaltdelikt
CAN	Webster et al. (1994)	618	+	Akman-Normandeau-Skala
D	Gretenkord (2001)	188	+	Gewaltsames Vordelikt

Anmerkung. GR = Gewaltrückfall: 0 = keine statistisch signifikante Beziehung; +: p<.05; ++: p<.01; +++: p<.001; n.a.: Signifikanz wurde nicht angegeben. Nach Gretenkord, 2001, S. 254.

Es wurden lediglich zwei Studien gefunden, in denen ein Zusammenhang zwischen dem Verhalten während der Unterbringung und dem Rückfall mit einem Gewaltdelikt hergestellt wurde (Tabelle 4). Zwar kann hier nicht der Anspruch erhoben werden, daß wirklich alle publizierten Untersuchungen erfaßt wurden, aber offensichtlich gibt es nur wenige Forscher, die sich für diesen – in der Praxis doch sehr bedeutsamen – Punkt interessieren.

Tabelle 4: Literaturlauswertung: Gewalttätigkeit während der Unterbringung und Rückfälligkeit mit einem Gewaltdelikt

Land	Untersuchung	N	GR	Bemerkungen
CAN	Harris, Rice & Cormier (1991)	169	0	Angriffe im letzten Jahr (s. Text)
CAN	Rice & Harris (1992) (nur Schizophrene)	96	0	Angriffe im ersten Jahr
CAN	dto.	96	+	Angriffe im letzten Jahr
D	Gretenkord (2001)	188	+	Gewalttätigkeit gegen Patienten und/oder Personal

Anmerkung. GR = Gewaltrückfall: 0 = keine statistisch signifikante Beziehung; +: p<.05; ++: p<.01; +++: p<.001; n.a.: Signifikanz wurde nicht angegeben. Nach Gretenkord, 2001, S. 256.

In den kanadischen Untersuchungen wurde zwischen Angriffen („assaults“) im ersten bzw. letzten Jahr der Unterbringung differenziert. Dies hat den Vorteil, daß eventuelle Therapieerfolge mit eingehen. Auf der anderen Seite reduziert sich dadurch die Anzahl der Aggressionen, was die Wahrscheinlichkeit reduziert, statistisch signifikante Zusammenhänge zu ermitteln; so fanden Harris, Rice und Cormier (1991) zwar, daß die Rückfälligen dreimal häufiger als die Nicht-Rückfälligen während des letzten Jahres der Behandlung aggressiv waren, dies war jedoch nicht statistisch signifikant.

Zum Zusammenhang des Alters bei der Entlassung aus der stationären Unterbringung mit dem Gewaltrückfall wurden sechs Untersuchungen gefunden (Tabelle 5). Nicht nur bei Straftätern allgemein, sondern auch bei psychisch gestörten Rechtsbrechern wurden überwiegend – wenn auch nicht immer – Korrelationen gefunden.

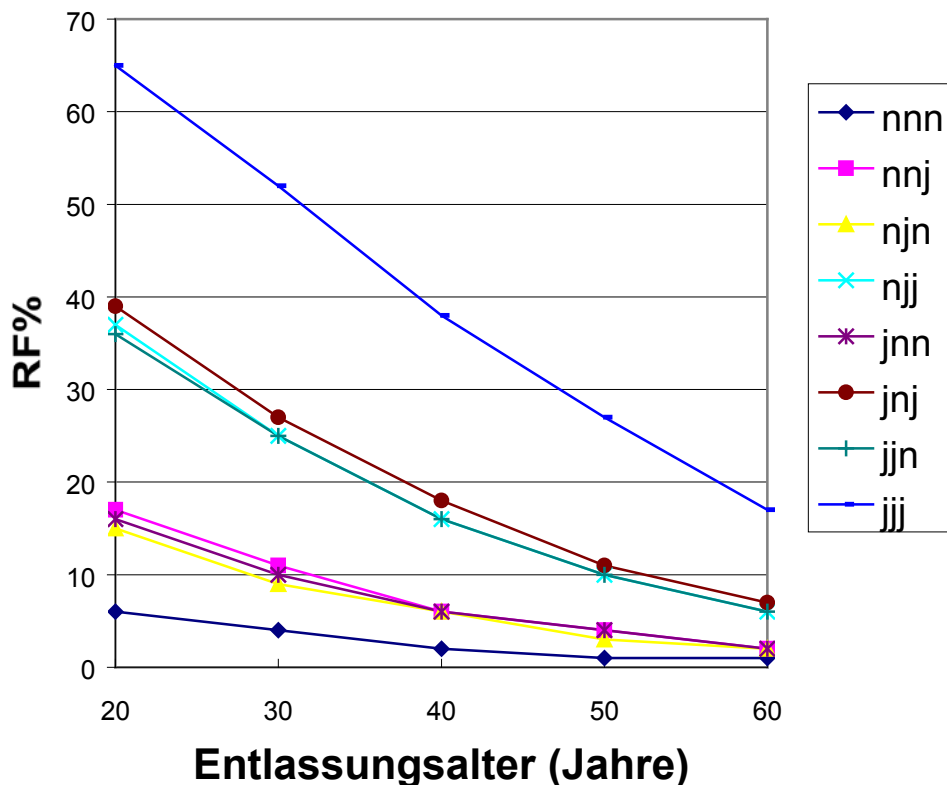
Tabelle 5: Literaturlauswertung: Entlassungsalter und Rückfälligkeit mit einem Gewaltdelikt

Land	Untersuchung	N	GR	Bemerkungen
USA	Steadman & Coccozza (1974)	98	n.m.	Offensichtlich signifikant
USA	Thornberry & Jacoby (1979)	414	+	
USA	Koppin (1990)	109	0	n.s., aber Jüngere häufiger rückfällig
CAN	Rice, Harris, Lang & Bell (1990)	253	+++	
CAN	Harris, Rice & Cormier (1991)	169	0	
CAN	Rice & Harris (1992)	192	0	
D	Gretenkord (2001)	188	+	

Anmerkung. GR = Gewaltrückfall: 0 = keine statistisch signifikante Beziehung; +:  $p < .05$ ; ++:  $p < .01$ ; +++:  $p < .001$ ; n.a.: Signifikanz wurde nicht angegeben. Nach Gretenkord, 2001, S. 259.

Im folgenden werden die Ergebnisse der Gleichung zunächst als Grafik dargestellt.

### Wahrscheinlichkeit eines Gewaltrückfalls nach Logistischer Regression



1. Stelle: "Persönlichkeitsstörung"
2. Stelle: "Vorbelastung mit Gewaltdelikt"
3. Stelle: "mindestens zweimal während der Unterbringung gewaltdelikt"

Anmerkung. Die Abkürzungen bedeuten folgendes: "j" steht für ja, also der Faktor ist gegeben. "n" steht für nein, der Faktor ist nicht gegeben.

Abbildung 1: Grafischer Überblick über die Wahrscheinlichkeit eines Gewaltrückfalls

Es wird deutlich, daß die Rückfallwahrscheinlichkeit mit höherem Entlassungsalter geringer wird. Die Gruppe mit der größten Rückfallwahrscheinlichkeit ist die Gruppe, bei der alle drei

Merkmale (persönlichkeitsgestört, Vorbelastung mit Gewaltdelikt, aggressiv während Behandlung) zutreffen, die Gruppe mit der geringsten Rückfallwahrscheinlichkeit ist die, bei der keines der drei Merkmale gegeben ist. Die anderen Gruppen liegen dazwischen.

Je übersichtlicher und nachvollziehbarer die Untersuchungsergebnisse dargestellt werden, desto größer ist die Chance, daß sie in der Praxis genutzt werden. Deshalb hat es nicht nur kosmetische Gründe, wenn eine möglichst gute Darstellung angestrebt wird.

Daher werden dieselben Ergebnisse noch einmal in Tabellenform dargestellt, so daß die Prozentwerte für die Rückfallwahrscheinlichkeit direkt abgelesen werden können. Das Entlassungsalter ist wieder in Zehnjahresschritten dargestellt (Tabelle 6).

Tabelle 6. Rückfallwahrscheinlichkeit nach Regressionsgleichung

Faktor	Rückfallwahrscheinlichkeit bei einem Entlassungsalter von ...						
	Vorstrafe mit Gewaltdelikt	Mind. 2x gewalttätig	20 J.	30 J.	40 J.	50 J.	60 J.
nein	nein	nein	6%	4%	2%	1%	1%
		ja	17%	11%	6%	4%	2%
	ja	nein	15%	9%	6%	3%	2%
		ja	37%	25%	16%	10%	6%
ja	nein	nein	16%	10%	6%	4%	2%
		ja	39%	27%	18%	11%	7%
	ja	nein	36%	25%	16%	19%	6%
		ja	65%	52%	38%	27%	17%

Ein Proband, der weder persönlichkeitsgestört ist, noch wegen Gewaltdelikten vorbelastet ist und der auch nicht während der Unterbringung gewalttätig war, hat je nach Entlassungsalter eine Rückfallwahrscheinlichkeit zwischen 1% und 6%. Die höchste Rückfallwahrscheinlichkeit hat ein Pb, bei dem alle drei Merkmale zutreffen und der bei der Entlassung noch jung ist, bei einem Zwanzigjährigen beträgt sie 65%.

Ein Skalpell in der Hand eines qualifizierten, erfahrenen und verantwortungsbewußten Chirurgen kann sehr nützlich sein. Ein Skalpell in der Hand eines Menschen, der diese Voraussetzungen nicht mitbringt, kann sehr schädlich sein. Dasselbe gilt für viele andere Werkzeuge, auch für die hier dargestellte Prognosestabelle. Es wird daher für wichtig gehalten, mögliche Fehler bei der Handhabung einer solchen Übersicht aufzulisten.

- Überinterpretation der Prozentangaben. In der Tabelle sind Prozentwerte aufgeführt, weil sie anschaulich und nachvollziehbar sind. Es handelt sich aber nur um die Prozentwerte für eine Gruppe von psychisch gestörten Rechtsbrechern, die in einem bestimmten Zeitraum aus einer Klinik entlassen worden sind. Niemand behauptet, daß für eine andere Gruppe genau dieselben Prozentwerte gelten. Der Informationsgehalt dieser Prozentangaben wird aber deshalb für wertvoll gehalten, weil sie empirisch ermittelt wurden und weil die Voraussetzungen, unter denen sie zustande gekommen sind, dargestellt wurden.
- Anwendung auf Gruppen, die sich in wesentlichen Merkmalen von der Untersuchungsgruppe unterscheiden. Die Ergebnisse sind an einer Gruppe von Probanden gewonnen worden, die nach § 63 StGB untergebracht waren. Sie sind selbstverständlich nicht einfach übertragbar etwa auf nach § 64 StGB Unterbrachte oder auf Probanden in Justizvollzugsanstalten.

- Anwendung auf ein anderes Rückfallkriterium. Der Inhalt der Prognosetabelle gilt für die Rückfälligkeit mit einem Gewaltdelikt, etwa Körperverletzung, Raub oder Vergewaltigung. Selbstverständlich können für die Rückfälligkeit etwa mit Sexualdelikten allgemein oder mit Brandstiftung andere Kriterien sinnvoller sein.
- Beschränkung auf die Variablen in der Tabelle. Selbstverständlich bilden die enthaltenen Variablen nur einen kleinen - aber doch wesentlichen - Teil der "Wirklichkeit" ab. Es wird keineswegs bestritten, daß andere Variablen ebenfalls eine - vielleicht noch wesentlichere - Rolle spielen können. Die Variablen wurden ausgewählt, weil sie in dieser Untersuchung - mit einer beschränkten Datenbasis und einer beschränkten Probandenzahl - ermittelt werden konnten. Es gibt beispielsweise nicht wenige Untersuchungen, in denen eine Suchtproblematik mit der Rückfälligkeit in Beziehung steht, dies sollte man im Auge behalten. Wenn etwa in einem Prognosegutachten andere Variablen eingeführt werden, wäre es allerdings gut, deutlich zu machen, warum man diese für relevant hält (Literaturhinweise, eigene Erfahrung, vergleichbare Fälle).
- Ausschalten des "gesunden Menschenverstandes". Sollte der Eindruck entstanden sein, daß hier für die schematische Anwendung computerunterstützter Modelle eine Lanze gebrochen wird, so ist dieser falsch. Auch bei einem Probanden, der nach den hier dargestellten Merkmalen sehr rückfallgefährdet ist, sollte die Entlassung gewagt werden, wenn er z.B. ein therapeutisches Programm erfolgreich durchlaufen hat, bei ausgedehnten Lockerungsmaßnahmen gezeigt hat, daß er seine Alkoholproblematik im Griff hat und ein Entlassungsumfeld bereit steht, das die in seinem Fall bestehenden Risiken kennt und bei Gefahr im Verzug geeignete Gegenmaßnahmen trifft. Und bei einem unkorrigierbar Wahnkranken, der einen vermeintlichen Verfolger umgebracht hat, der weiterhin paranoid ist und Drohungen ausstößt, wird man vorsichtig sein, auch wenn er nicht persönlichkeitsgestört ist, bisher nur ein Delikt begangen hat und sich in der stationären Behandlung an die Regeln gehalten hat.
- Verzicht auf die "klinische Prognose". Eine Prognosetafel ersetzt in der Praxis keineswegs die gründliche Analyse des Einzelfalles. Selbstverständlich müssen auch Kenntnisse über bestimmte Störungsbilder (z.B. typischer Verlauf bestimmter Formen der Schizophrenie) oder Therapiemethoden (z.B. Wirksamkeit antidepressiver Medikamente bei affektiver Psychose) oder die Situation nach der Entlassung (Kooperationsfähigkeit und -bereitschaft des sozialen Empfangsraumes) bei der Entscheidung über Entlassung oder Nichtentlassung eine Rolle spielen.
- Verzicht auf den Fachmann. Wenn man ein gebrauchtes Auto kauft, ist es kein Fehler, jemanden mitzunehmen, der etwas von Autos versteht. Ein erfahrener Kraftfahrzeugingenieur kann eher erkennen, ob beispielsweise eine fatale Roststelle geschickt übertüncht ist oder ob das Geräusch des Motors für die Prognose von dessen Lebensdauer etwas hergibt.

Nun kann man gegen das hier vorgestellte Konzept sicherlich einiges vorbringen. Auf einen Punkt soll kurz eingegangen werden. So könnte ein Kritiker sagen:

Die Daten zur Rückfälligkeit stammen von Patienten aus einer einzigen Klinik nur eines Bundeslandes. Außerdem sind sie auch schon ziemlich alt. Sie können daher nicht auf Maßregelvollzugspatienten in anderen Bundesländern angewendet werden, und auch nicht auf solche, die erst demnächst entlassen werden sollen.

Das ist im Prinzip richtig. Um die Verwendung der Prognosetafel für den Maßregelvollzug allgemein dennoch zu ermöglichen, wurden Variablen ausgesucht, die nicht spezifisch für eine Klinik oder ein Bundesland sind und sich auch in anderen Untersuchungen als prognoserelevant erwiesen haben. So lange keine besseren Daten zur Verfügung stehen, können diese einen praxisrelevanten Maßstab darstellen. Denn die Alternative – freie

Schätzung des jeweiligen Prognostikers aufgrund seiner Fachkenntnisse und Erfahrungen – ist noch weitaus weniger überzeugend. Selbstverständlich kann die hier entwickelte Tabelle nur ein Anfang sein. Kreuzvalidierungen sind erforderlich. Wenn weitere Daten aus möglichst vielen Institutionen in einer Datei zusammengefaßt werden, kann der Nutzen um ein Vielfaches gesteigert werden.

Abschließend sollen einige wichtige Punkte zusammengefaßt werden:

- Bei jedem gemäß § 63 StGB Untergebrachten ist jährlich zu entscheiden, ob er entlassen werden kann oder nicht.
- Die Kosten einer Fehlentscheidung sind enorm (erneute Delinquenz oder ungerechtfertigter und kostspieliger Freiheitsentzug).
- Es sollten alle Möglichkeiten ausgeschöpft werden, die Fehlerrate so gering wie möglich zu halten.
- Empirische Untersuchungen haben gezeigt, daß es bedeutsame Zusammenhänge zwischen einzelnen Faktoren und der Rückfälligkeit gibt.
- Die Ergebnisse solcher Untersuchungen sollten so aufbereitet werden, daß sie in der Praxis genutzt werden können (z.B. als "Prognosetafel").
- Ein solches Instrument gehört in die Hand einer qualifizierten, erfahrenen und verantwortungsbewußten Fachkraft, welche die Ergebnisse als "Basisrate" oder "Ankerreiz" mit in ihre Überlegungen einbezieht.

## Literatur

- Bailey, J., & MacCulloch, M. (1992). Characteristics of 112 cases discharged directly to the community from a new Special Hospital and some comparisons of performance. *The Journal of Forensic Psychiatry*, 3(1), 91-112.
- Gretenkord, L. (2001). *Empirisch fundierte Prognosestellung im Maßregelvollzug nach § 63 StGB - EFP-63*. Bonn: Deutscher Psychologen Verlag.
- Hare, R. D. (1991). *The Hare Psychopathy Checklist - Revised. Manual*. New York, Toronto: Multi-Health Systems.
- Harris, G. T., Rice, M. E., & Cormier, C. A. (1991). Psychopathy and violent recidivism. *Law and Human Behavior*, 15, 625-637.
- Jockusch, U., & Keller, F. (1993). Differenzierung und Spezialisierung einer Abteilung des Maßregelvollzugs in der psychiatrischen Klinik - mögliche Auswirkungen auf Behandlungsdauer und strafrechtliche Rückfälligkeit. *Psychiatrische Praxis*, 20 (Sonderheft I), 12-15.
- Koppin, M. K. (1990). *Prediction of dangerous behavior for release of the criminally-insane in Colorado*. Ann Arbor, Michigan: University Microfilms International.
- Rice, M. E., & Harris, G. T. (1992). A comparison of criminal recidivism among schizophrenic and nonschizophrenic offenders. *International Journal of Law and Psychiatry*, 15(4), 397-408.
- Rice, M. E., Harris, G. T., Lang, C., & Bell, V. (1990). Recidivism among male insanity acquittees. *Journal of Psychiatry and Law*, 18 (3-4), 379-403.
- Steadman, H., & Coccozza, J. (1974). *Careers of the Criminally Insane*. Lexington, Mass.: Lexington Book.
- Tennent, G., & Way, C. (1984). The English Special Hospital - a 12-17 year follow-up study: A comparison of violent and non-violent re-offenders and non-offenders. *Medicine, Science and Law*, 24 (2), 81-91.



Thornberry, T., & Jacoby, J. (1979). *The Criminally Insane: A Community Followup of Mentally Ill Offenders*. Chicago: University of Chicago Press.

Webster, C. D., Harris, G. T., Rice, M. E., Cormier, C., & Quinsey, V. L. (1994). *The Violence Prediction Scheme. Assessing dangerousness in high risk men*. Toronto: University of Toronto.



# Prognose von Täterverhalten bei Geiselnahmen

*Everhard von Grootte*

*Polizeifortbildungsinstitut Neuss*

In Geiselnahmen stellt sich für die Polizei das Problem, mit nur wenigen und unvollständigen Informationen Entscheidungen mit weitreichenden Konsequenzen treffen zu müssen. Besondere Probleme ergeben sich in der Prognose von Täterverhalten im weiteren Tatverlauf. Die psychologischen Entsprechungen polizeilicher Probleme werden hier betrachtet, wobei die Prognose von Täterverhalten als Sonderfall psychologischer Diagnostik verstanden wird. Angelehnt an die Verhaltensgleichung von Westhoff und Kluck (1991) werden Variablengruppen untersucht, die Hinweise zur Prognose von Täterverhalten erstens in Bezug auf Fragestellungen zu Gewaltbereitschaft und zweitens zur Suizidalität des Täters geben. Das aus diesen Überlegungen resultierende Instrument ist als PC-gestütztes System zur Entscheidungsunterstützung realisiert. Erläutert werden Aufbau und Arbeitsweise dieses Programms. Eine Validierung des Systems ist aufgrund der kleinen Fallzahlen schwierig. Es werden Ansätze zur Überprüfbarkeit der Testgütekriterien diskutiert.

## Hintergrund

Geiselnahmen sind seltene und komplexe Ereignisse, in denen die Polizei eine Vielzahl von Informationen verarbeiten und Entscheidungen treffen muss. Von besonderem psychologischen Interesse ist die Prognose von Täterverhalten. Da der Begriff der Geiselnahme im allgemeinen Sinne einige Unschärfen birgt und auch die juristische Definition im StGB (Tröndle & Fischer, 1999) sowohl für die praktische polizeiliche Arbeit als auch für die Betrachtung der psychologischen Dimensionen des Geschehens wenig brauchbar ist, wird im folgenden auf eine polizeiinterne Definition nach Gatzke (1996) zurückgegriffen. Danach „handelt es sich um eine Geiselnahme im polizeitaktischen Sinne dann, wenn Täter unter Verwirklichung der Tatbestände §§ 239a oder 239b StGB Personen zur Durchsetzung ihrer Ziele an einem der Polizei bekannten Ort in ihrer Gewalt halten“ (S. 430). Davon zu unterscheiden ist die Bedrohungslage, die eine Geiselnahme im familiären Kontext darstellt. Diese ereignet sich gewöhnlich dann, wenn Paare in Trennungssituationen sind und der Mann über keine ausreichenden gewaltfreien Konfliktlösungsmuster verfügt. Die polizeiliche Kriminalstatistik weist für 1999 103 Fälle von erpresserischem Menschenraub und Geiselnahme aus (Bundeskriminalamt, 2000). Klassische Geiselnahmen im Sinne der o.g. Definition ereignen sich in Deutschland etwa 4 bis 10 mal pro Jahr (Landeskriminalamt NRW, 1999), Bedrohungslagen etwa 35 mal jährlich allein in NRW (Landeskriminalamt NRW, 1999). Jede Geiselnahme stellt ein einmaliges und sehr komplexes Geschehen dar (Prinz, 1995). Um darauf angemessen, das heißt dem konkreten Einzelfall gerecht werdend, reagieren zu können, muss die Polizei eine möglichst genaue Vorstellung davon entwickeln, welche Faktoren tatrelevant sind und wie die jeweilige Ausprägung dieser Faktoren ist. Die Lagebeurteilung umfasst daher verschiedene Aspekte (beispielsweise Orts- und Zeitfaktoren, Beurteilung der involvierten Personen usw.). Ein besonders wichtiger Aspekt der Lagebeurteilung betrifft den oder die Täter. Die Polizei muss versuchen, eine möglichst zutreffende Vorstellung von der Gefährlichkeit des Täters bzw. der Täter zu entwickeln. Die wichtigste Frage ist, ob (oder genauer: unter welchen Bedingungen) davon auszugehen ist, dass der oder die Täter ihre Geiseln tatsächlich verletzen oder töten werden. Diese Frage ist sowohl für die aktuelle Situation als auch für den weiteren Fortgang der Lage relevant. Täter, deren Gefährlichkeit als besonders hoch einzustufen ist, erfordern andere polizeiliche Konzepte und Maßnahmen als Täter von eher geringer Gefährlichkeit.

## **Fragestellung**

Um Täterverhalten kategorisieren und vorhersagen zu können, wurde polizeiliches Erfahrungswissen in polizeilichen Verhandlungsgruppen gesammelt und in beschränktem Maße publiziert (Fuselier, 1981; Gist & Perry 1985; Greenstone, 1995; Herbertz, 1979; Krummsiek, 1979; Steffen & Polz, 1991). Gemeinsam ist den Erfahrungsberichten und Modellen, dass sie auf wenig oder keiner empirischen Forschung beruhen, sondern von Praktikern aufgrund ihres Erfahrungsschatzes erstellt wurden. Wissenschaftliche empirische Untersuchungen fanden in Deutschland bislang noch nicht statt, was einerseits auf polizeiliche Vorbehalte gegenüber einer Öffnung dieses sensiblen Bereichs für Forscher zurückzuführen ist, andererseits aber auch auf geringes wissenschaftliches Interesse. In der polizeilichen Praxis werden Tätereinschätzungen von Polizeibeamten und in einigen Bundesländern von Psychologen in Zusammenarbeit mit Polizisten durchgeführt. Dabei ergeben sich aus psychologischer Sicht verschiedene Probleme. Ein Problem besteht darin, dass die Einschätzungen verschiedener Beamter und Psychologen sehr unterschiedlich ausfallen können, da unterschiedliche Kriterien zur Einschätzung herangezogen werden. Auch die Gewichtung der Kriterien ist in das Belieben der mit dieser Aufgabe betrauten Personen gestellt. Es fehlen bislang praktikable, reliable und valide Arbeitshilfen, die die hier zu leistende Arbeit unterstützen können. Das hier vorgestellte Verfahren versucht, durch eine Systematisierung und Strukturierung dieses Aufgabenfeldes analog zu einem üblichen diagnostischen Vorgehen ein Hilfsmittel zur Tätereinschätzung zu entwickeln.

## **Methode**

Wie bei jeder psychodiagnostischen Aufgabe gilt es zuerst, die zu beantwortenden psychologischen Fragen genau zu definieren. Gefordert ist neben der Beschreibung oder Klassifizierung des Täters vor allem eine Prognose des von ihm zu erwartenden Verhaltens. In der Tätereinschätzung sind besonders zwei Fragen von großer Bedeutung: Wie wahrscheinlich ist es, dass der oder die Täter physische Gewalt gegen die Geiseln anwenden? Dies ist eine zentrale Einschätzungsfrage in jeder Geiselnahme, da sie unmittelbar mit dem polizeilichen Ziel verbunden ist, das Leben der Geiseln zu schützen. Definitionswürdig ist hier der Begriff der Gewalt: Dieser Begriff ist nicht eindeutig belegt und wird weder im StGB, das den Gewaltbegriff in einer Vielzahl von verschiedenen Paragraphen verwendet, definiert, noch in der Wissenschaft homogen verwendet. Ein für das Thema tauglicher Ausgangspunkt zur Definition von Gewalt ist das Abstellen auf den Eingriff in die physische Integrität eines Menschen. Unter Gewalt soll hier die Ausübung solcher physischer Gewalt verstanden werden, die das Opfer schwerwiegend verletzt oder tötet. Somit fallen aus dem engen Gewaltbegriff, wie er hier verwendet wird, Handlungen heraus, die zwar in ihrer Art auch mit Angriff oder Aggression verbunden sind, sich aber in ihrem Effekt gegen Dinge richten. Zweitens wird die Frage untersucht, wie wahrscheinlich ein Suizid oder ein erweiterter Suizid des Geiselnahmers ist. Diese Frage scheint zunächst vielleicht weniger relevant. Es gibt jedoch verschiedene Geiselnahmen und Bedrohungslagen, die mit Suiziden oder erweiterten Suiziden geendet haben (Landeskriminalamt NRW, 1999), woraus sich die hohe polizeiliche Relevanz des Themas ergibt.

Zur Entwicklung eines diagnostischen Instruments, das die oben genannten Fragen beantworten kann, wurde auf die von Westhoff und Kluck (1991) entwickelte Verhaltensgleichung zurückgegriffen. Diese teilt die zur Erklärung, Beschreibung, Vorhersage und Beeinflussung von Verhalten relevanten Variablen in Variablengruppen ein. Dies dient nach Westhoff (1991) mehreren Zwecken: "Sie [die Verhaltensgleichung] faßt

[sic] die Variablen, die bei der Erklärung, Vorhersage und Beeinflussung individuellen Verhaltens wichtig sind, zu wenigen Gruppen zusammen und erlaubt somit eine Strukturierung dieses ansonsten kaum übersehbaren Bereichs. Sie unterstützt die Prüfung, ob bei der Begutachtung alle wichtigen Bereiche bedacht wurden. Sie ist also als Hilfsmittel für das praktische Arbeiten gedacht“ (S. 25). Die Verhaltensgleichung umfasst folgende Variablengruppen:

$$V = f_i (O,K,E,M,S,U)$$

Dabei bedeuten die einzelnen Abkürzungen:

Verhalten (V) ist eine Funktion (f) von

- Organismusvariablen (O),
- kognitiven Variablen (K),
- emotionalen Variablen (E),
- motivationalen Variablen (M),
- soziale Variablen (S),
- Umgebungsvariablen (U)
- und deren Wechselwirkungen (Subskript i).

Organismusvariablen beziehen sich auf alle körperlichen Bedingungen wie beispielsweise Krankheiten, Drogenmissbrauch und -abhängigkeit, Hunger und Durst oder Müdigkeit. Die Gruppe von Variablen, die als kognitive Variablen bezeichnet werden, beziehen sich auf die geistige Leistungsfähigkeit und umfassen beispielsweise Konzentrationsfähigkeit oder Intelligenz. Emotionale Variablen erfassen die jeweils relevanten Gefühle, die Verhalten mitbestimmen, begleiten und färben. Motivationale Variablen umfassen nicht nur die unmittelbare Motivation zum Verhalten, sondern auch Werte, Ziele, Wünsche und Überzeugungen. Soziale Variablen betrachten die sozialen Bezüge und ihre Auswirkungen auf das Verhalten. Umgebungsvariablen beziehen sich auf die äußeren Bedingungen von Verhalten.

Zur Vorhersage von konkretem Täterverhalten wurden die einzelnen Variablengruppen daraufhin abgesehen, ob hier für das vorherzusagende Verhalten bedeutsame Variablen existieren. Dies geschah in zwei Schritten. Erstens wurde ein Workshop mit polizeilichen Spezialisten für Geiselnahmen (Mitglieder einer Verhandlungsgruppe) durchgeführt. Zweitens wurde eine Literatursichtung zu den relevanten Themen vorgenommen. Der einwöchige Workshop mit Mitgliedern einer Verhandlungsgruppe befasste sich mit dem Auffinden von Prädiktorvariablen für den Bereich der Vorhersage von Gewaltanwendung gemäß der o.g. Definition in Geiselnahmen. In dem Workshop wurde methodisch so vorgegangen, dass alle Variablengruppen nacheinander von den polizeilichen Verhandlern vor dem Hintergrund ihrer Erfahrungen in verschiedenen Einsätzen diskutiert wurden. Ziel war, der Systematik der Verhaltensgleichung folgend, die Prädiktorvariablen zu beschreiben, die den Verhandlern auf Grund ihrer Erfahrungen wichtig sind. Dazu haben die Verhandler zunächst in Einzelarbeit alle ihnen wichtig scheinenden Variablen der jeweiligen Variablengruppe notiert. Die Einzelnennungen wurden daraufhin gesammelt und in der Gruppe vor dem Hintergrund der Einsatzerfahrungen diskutiert. Letztlich wurden die Variablen in eine abschließende Liste aufgenommen, über deren Relevanz nach der Diskussion in der Gruppe weitgehende Einigkeit bestand. Variablen, deren Wirksamkeit von mindestens zwei der zehn Verhandler als nicht relevant angesehen wurden, wurden nicht aufgenommen.

Die Ergebnisse des Workshops spiegeln die sich auf Einsatzerfahrungen gründenden Alltagstheorien der Verhandler wider, die die Form von vor dem beruflichen Hintergrund reflektierten Überzeugungen und Behauptungen haben. Diese können als alleinige Informationsquelle zur inhaltlichen Füllung eines Prognoseinstruments nicht ausreichen. Deshalb schloss sich an den Workshop eine Literaturrecherche an, in der die einzelnen Variablen daraufhin überprüft wurden, inwieweit sie durch Forschungsergebnisse gestützt werden. Die Psychologie kann zum Thema „Gewalt“ auf einen recht großen Fundus einschlägiger Forschung zurückgreifen (Überblick z.B. bei Selg, Mees & Berg, 1997). Auch zum Thema Suizid liegen zahlreiche wissenschaftliche Publikationen vor, die relevantes Handlungswissen zur Beantwortung dieser Fragestellung anbieten (Überblick z.B. bei Welz & Möller, 1984; Van Egmond & Diekstra, 1990).

Problematisch an der Verwendbarkeit der vorliegenden Forschungsergebnisse in beiden Bereichen ist, dass hier meist andere Facetten des jeweiligen Verhaltens untersucht wurden. Im Fall der Verhaltensprognose während einer Geiselnahme steht die Vorhersage von folgenden Facetten der Gewalt im Vordergrund: Welche Arten von Gewalt können vorkommen? Hier muss zwischen Gewalt gegen Sachen, andere Personen (Geiseln) und sich selbst (Selbstverletzung, Suizid) unterschieden werden. Eine andere Facette betrifft die Schwere der Folgen der Gewaltanwendung: Ist eine ernste Verletzung (z.B. Beinschuss) oder gar eine Tötung zu erwarten? Eine weitere Facette betrifft die Häufigkeit: Wie oft wird Gewalttätigkeit vorkommen? Eine andere Facette beleuchtet den Zeitpunkt: Wie nah steht eine Gewalttat bevor? Als letzte Facette ist schließlich die Wahrscheinlichkeit zu nennen: Wie groß ist die Wahrscheinlichkeit, dass die beschriebenen Facetten von Gewalttätigkeit überhaupt vorkommen werden? Vor allem auf den Punkt der Vorhersage der Wahrscheinlichkeit des Auftretens gewalttätigen Verhaltens hat die rechtspsychologische Forschung einen Schwerpunkt gelegt, während die anderen Punkte eher vernachlässigt wurden (Hart, 1999). Es liegen zahlreiche Forschungsarbeiten und Instrumente vor, die Aussagen darüber zulassen, wie groß die Wahrscheinlichkeit ist, dass eine Person gewalttätig wird. Problematisch für die polizeiliche Täterereinschätzung in Geiselnahmen ist jedoch die Übertragbarkeit dieser Forschungsergebnisse auf die konkrete Situation einer Geiselnahme. Vor allem die Zeit-Facette unterscheidet eine Vorhersage in einer Geiselnahme grundlegend von den wissenschaftlichen Vorhersagestudien. Während die wissenschaftlichen Studien und die zur Unterstützung der Justiz erstellten Prognoseverfahren lange Prognosezeiträume abdecken (Monate bis Jahre), muss während einer Geiselnahme eine Aussage über die Gewaltwahrscheinlichkeit innerhalb der nächsten Minuten und Stunden getroffen werden. Es ist anzunehmen, dass sich zumindest einige der dafür heranzuziehenden Prädiktoren grundlegend von denen der Langzeitvorhersagen unterscheiden werden. So ist ein extrem aufgeregter Geiselnahmer nach Ansicht der Verhandler im Hinblick auf seine Gewaltbereitschaft anders einzuschätzen als ein sehr ruhiger. Aktuelle Erregung spielt hingegen bei der Langzeitvorhersage keine bedeutende Rolle. Ein weiterer Unterschied liegt in der Bewertung des Kontextes: Eine Geiselnahme stellt auch für den Täter einen hochgradig unnormalen Kontext dar. Auch deshalb sind Zweifel angebracht, ob die wissenschaftlich erforschten Risikofaktoren für Gewalttätigkeit auch in dieser Ausnahmesituation Gültigkeit haben. Die Kenntnis des Kontextes bietet für die Vorhersage jedoch auch einen bedeutenden Vorteil gegenüber der Vorhersage der Gewalttätigkeit über einen längeren Zeitraum, in der der Kontext kaum berücksichtigt werden kann. Ein weiterer Vorteil besteht darin, dass der Vorhersagezeitraum kurz ist und lediglich wenige Stunden umfasst. In dieser Zeit befindet sich der Täter in einem recht genau definierten Kontext. Es sind der Ort, die Ausgangslage, die Anzahl der Geiseln, die Drohungen und Ultimaten und viele andere für die Beurteilung des Kontexts bedeutsame Situationsfaktoren bekannt. Diese Kontextinformationen können ebenfalls zur Vorhersage genutzt werden.

So konnte erst in der Zusammenschau der Ergebnisse des Workshops und der Literaturrecherche eine Auswahl der Variablen getroffen werden, die einerseits wissenschaftlich abgesichert sind und andererseits von den Verhandlern als praktisch relevant eingestuft werden. Die Auswahl und Gewichtung der Variablen geschah durch Expertenrating. Ein Team aus zwei Diplompsychologen und zwei erfahrenen Verhandlungsgruppenmitgliedern hat die Kriterien einzeln vor dem Hintergrund der Literaturlage und der praktischen Erfahrungen diskutiert und jedem einzelnen Kriterium jeweils ein Gewicht in Bezug auf die Relevanz der vorherzusagenden Verhaltensweisen zugeordnet. Die Gewichte, mit denen das System arbeitet, berechnen sich aus den gemittelten und gerundeten Einzelgewichten dieser vier Experten. Diese Vorgehensweise stellt sicher, dass sowohl die wissenschaftlichen Befunde zu den Themen "Gewalt" und "Suizid" als auch die praktischen Erfahrungen in den entsprechenden Einsätzen in gleicher Stärke einfließen. Kritisch anmerken lässt sich hier, dass dieses Verfahren nicht notwendigerweise sicherstellt, dass die vergebenen Gewichte tatsächlich ein genaues Abbild der realen Einflüsse der Variablen auf die Kriterien darstellt. Aufgrund der geringen Fallzahlen sehe ich jedoch derzeit keine andere Möglichkeit, als hier auf Expertenratings zurückzugreifen, die aufgrund der jeweiligen Einsatzerfahrungen reflektiert und ggf. angepasst werden müssen.

Tabelle 1 zeigt die auf diese Weise generierten zur Verhaltensvorhersage relevanten Variablen mit ihren Gewichten im Überblick. Die Variablen werden gemäß der Verhaltensgleichung in Organismus-, emotionale, motivationale, soziale und Umgebungsvariablen unterschieden. Zusätzlich werden relevante Informationen aus der Geschichte des Täters und psychische Erkrankungen mit Vorhersagewert dargestellt.

Tabelle 1: Variablen zur Verhaltensvorhersage in Geiselnahmen

Variable	Gewicht Suizidalität	Gewicht Gewaltbereitschaft in Bedrohungslagen	Gewicht Gewaltbereitschaft in klassischen Geiselnahmen
<b>Organismusvariablen</b>			
Alter des Täters	1	3	3
Geschlecht männlich	2	1	3
Starke Schmerzen	4	4	4
Lethale Erkrankung	2	2	2
Akute Alkoholisierung	5	5	4
Akuter Entzug von Alkohol oder Drogen	2	3	3
<b>Kognitive Variablen</b>			
Kognitive Überforderung	4	4	3
Traditionelles Rollenverständnis	0	3	0
Internaler Locus of Control	0	3	0
Verminderte Fähigkeit, mit Druck umzugehen	3	3	3
Verminderte Intelligenz	2	3	3
Auf der Stelle tretendes Denken mit gewaltsamen Inhalten	2	4	3
Politisch motivierter Täter	0	0	4
Besondere mit Gewalt verbundene Hobbys und Fähigkeiten	1	3	2
Nichts zu verlieren haben	2	3	3
<b>Emotionale Variablen</b>			
Impulsivität	3	3	3
Ängstlichkeit	1	3	1

Feindseligkeit	0	3	3
Geringes Selbstwertgefühl	3	3	1
Hohes aktuelles Erregungsniveau	3	5	4
Hohe Erschöpfung	1	3	3
Akuter Verlust einer geliebten Person	4	2	2
Akute Hoffnungslosigkeit	5	4	3
Hass, Wut, Rachegefühle	4	4	4
Akut hohe Angst	5	3	4
<b>Motivationale Variablen</b>			
Täter will seine Macht beweisen	0	3	3
Täter will Handlungsstillstand durchbrechen	3	3	3
<b>Soziale Variablen</b>			
Intime Beziehung zwischen Täter und Geisel	3	4	1
Negative Beziehung zwischen Täter und Geisel	1	5	4
Positive Beziehung zwischen Täter und Geisel	0	0	-3
Geisel provoziert Täter	0	6	5
Täter ist allein, alle Geiseln konnten fliehen	4	0	0
Schlechte Beziehung zum Polizeisprecher	3	4	4
Täter hat Kriegs- oder Bürgerkriegserfahrung	2	3	3
Täter ist Mitglied in einer besonders gewaltbereiten Gruppe	1	3	3
Konfliktreiche Herkunftsfamilie	0	3	3
Psychische Störungen in der Familie	5	0	0
Jugendliche: Probleme mit den Eltern oder der Eltern	4	0	0
Jugendliche: Schulprobleme	4	0	0
Partnerschaftsprobleme	4	4	1
Fehlende soziale Unterstützung	5	4	2
Soziale Umbruchsituation	4	0	0
Verhaltensauffälligkeiten in Kindheit und Jugend	3	3	2
<b>Umgebungsvariablen</b>			
Täter verfügt über mehr als eine Geisel	0	0	2
Täter verfügt über Handgranate/Sprengstoff	3	3	2
Funktionsfähiges Drohmittel	3	3	3
Täter verfügt über Schusswaffe	3	3	3
<b>Biographische Variablen</b>			
Vorstrafen und Vortaten	3	5	5
Geringer oder fehlender Schulabschluss	3	3	3
Psychotherapeutische Hilfe in Anspruch genommen	3	0	0
Frühere Suizidversuche	5	2	0
Gewalt gegen Geisel in laufender Lage	1	6	6
<b>Psychische Erkrankungen</b>			
Erkrankungen aus dem depressiven Formenkreis	4	0	0
Alkoholmissbrauch und/oder Abhängigkeit	4	4	4
Drogenmissbrauch und/oder Abhängigkeit	4	4	4



Psychotische Erkrankungen	4	3	3
Borderline-Störung	4	3	2
Antisoziale Persönlichkeitsstörung	4	4	4
Täter mit emotional instabiler Persönlichkeit	4	3	4
Paranoide Persönlichkeitsstörung	3	3	3

Die Prädiktorvariablen wurden mit ihren jeweiligen Gewichten in ein Decision Support System (DSS), also ein PC-gestütztes Expertensystem zur Entscheidungsunterstützung, eingestellt (eine Einführung in das Thema Expertensysteme findet sich bei Amelingmeyer & Stahringer, 1999).

Die Eingaben in das System erfolgen in der Eingabemaske. Die Benutzer können die Variablen, zu denen sie über Informationen verfügen, bewerten und die entsprechenden Werte in das Feld „Einschätzung“ jeweils neben der Variablen eingeben.

Der Nutzer kann zur Ausgabe der Ergebnisse verschiedene Wege nutzen. Die sogenannte „Ergebnisseite“ hält nur die Ergebnisse vor, ohne Informationen über das Zustandekommen anzugeben. Diese Seite zeigt dem Nutzer die in Abbildung 1 gezeigten Informationen an.

14.01.2001 18:31		
<b>Tätereinschätzung Geiselnahme</b>		
<b>Ergebnisse</b>		
Gewaltbereitschaft:		
Die momentan vorliegenden Informationen sind	<b>ausreichend</b>	aussagekräftig für eine Bewertung
Eine Gewaltanwendung des Täters gegenüber Geiseln ist momentan	<b>wenig wahrscheinlich</b>	
Suizidgefahr:		
Die momentan vorliegenden Informationen sind	<b>eingeschränkt</b>	aussagekräftig für eine Bewertung
Eine Suizidhandlung des Täters ist momentan	<b>wenig wahrscheinlich</b>	

Abbildung 1: Ergebnisseite des DSS „Geiselnahme“

Oben werden Datum und Uhrzeit der Einschätzung festgehalten. Dies scheint zunächst vielleicht trivial, ist aber zur Dokumentation eines Verlaufs im realen Einsatz wichtig. So lassen sich die Einschätzungen zu verschiedenen Stadien der Lage miteinander vergleichen. Hierunter werden Einschätzungen angegeben. Für jeden der Fälle werden Aussagen zu Suizidalität und Gewaltbereitschaft getroffen.

Der Benutzer erhält weiterhin Informationen darüber, wie aussagekräftig die vorliegenden Informationen für eine Bewertung der Suizidalität und der Gewaltbereitschaft des Täters gegenüber den Geiseln sind. Diese Information wird dem Benutzer (wie alle folgenden Informationen dieser Seite auch) in Form eines Satzes gegeben, in dem das Wort, das die Ausprägung der Aussagekraft ausdrückt, farblich abgesetzt ist und je nach Erkenntnisstand variiert. Der Satz lautet:

*Die momentan vorliegenden Informationen sind ( umfassend; ausreichend; eingeschränkt; unzureichend ) aussagekräftig für eine Bewertung.*

Hier wurde eine 4stufige Skala gewählt, um die Informationsdichte verbal zu beschreiben. Welche konkrete Einstufung erfolgt, hängt von einem intern errechneten Quotienten ab, der angibt, wie viele der zur Vorhersage relevanten Variablen bewertet wurden und in die Berechnung mit eingeflossen sind. Da allein die Summierung der Anzahl der Variablen jedoch wegen der verschiedenen Gewichtungen der Variablen in die Irre führen würde, werden die Variablen jeweils mit ihren Gewichten in diese Berechnung eingestellt. Der so berechnete Quotient gibt also ein Maß dafür, wie viel Prozent der dem System maximal verfügbaren Informationen von den Ratern eingegeben wurden und damit zur Verrechnung zur Verfügung standen. Wurden dem System weniger als 50% der relevanten Informationen zur Bewertung von Suizidalität und Gewaltbereitschaft eingegeben, so wird der o.g. Satz mit “unzureichend” ergänzt werden. Stehen zwischen 51 und 70% der Informationen zur Verfügung, so lautet die Einstufung “eingeschränkt”. Bei 71 bis 80% wird die Einstufung “ausreichend” gewählt, während im günstigsten Fall, 81 bis 100% Information, die Einstufung “umfassend” gewählt wird. Die hier beschriebenen Schwellen sind wiederum im Abgleich und der Diskussion mit einer Verhandlungsgruppe entstanden und spiegeln die Vorsicht wider, mit der das System Einschätzungen vornimmt. Die Einschätzung von Suizidalität und Gewaltbereitschaft erfolgt nach folgenden Regeln: Im ersten Fall ( “unzureichend”) werden keine Aussagen gemacht. In allen anderen Fällen wird eine Aussage hierzu getroffen, und zwar in folgender Form:

*„Eine Suizidhandlung des Täters ist momentan (eher unwahrscheinlich; wenig wahrscheinlich; wahrscheinlich; sehr wahrscheinlich).“*  
und

*“Eine Gewaltanwendung des Täters gegenüber den Geiseln ist momentan (eher unwahrscheinlich; wenig wahrscheinlich; wahrscheinlich; sehr wahrscheinlich).“*

Tabelle 2 gibt einen Überblick über die gewählten Schwellen.

Tabelle 2: Schwellen zur verbalen Codierung der Vorhersagen für Suizidalität und Gewaltbereitschaft

<b>Prozent der für die Vorhersage verfügbaren Informationen</b>	<b>Einstufung</b>
0 bis 25 %	eher unwahrscheinlich
26 bis 50 %	wenig wahrscheinlich
51 bis 75 %	wahrscheinlich
76 % bis 100 %	sehr wahrscheinlich

Für den Nutzer und letztlich auch für den Polizeiführer ist es sehr bedeutsam, das Zustandekommen der Ergebnisse nachvollziehen zu können. Daher hat der Nutzer die Möglichkeit, genaue Informationen darüber aufzurufen, wie eine Einstufung zustande kam. Er kann bei Bedarf im Detail nachvollziehen, welche Werte in die Berechnung eingingen und welche Rechenschritte vollzogen wurden.

## **Diskussion**

Das hier vorgestellte Modell birgt methodisch zahlreiche noch unbeantwortete Fragen, die hier kurz diskutiert werden sollen. So stellt sich die Frage, ob die Prädiktorvariablen sinnvoll gewählt sind. Sie entstammen, wie oben geschildert, einer Expertenbefragung und einer Literatursichtung und wurden weitgehend additiv zusammengestellt. Eine Analyse

hinsichtlich der Trennschärfe und der Homogenität und, dies scheint mir besonders wichtig, eine Überprüfung der Dimensionalität haben bisher noch nicht stattgefunden. Ich halte es durchaus für möglich, dass sich über eine Betrachtung der Interkorrelationen mehrere homogene Cluster identifizieren lassen, die sich auch theoretisch fassen und begründen lassen. Noch ist unklar, welche Konstrukte hier genau operationalisiert werden. Wenn dies genauer bekannt wäre, ließe sich die Vorhersageleistung u.U. auch mit einer wesentlich geringeren Anzahl von Prädiktorvariablen in gleicher Güte gewährleisten.

Da das Modell auch als Test im Sinne von Lienerts Definition eines Tests als „... Verfahren zur Untersuchung eines oder mehrerer empirisch abgrenzbarer ... Merkmale mit dem Ziel einer ... Aussage über den relativen Grad der individuellen Merkmalsausprägung“ (Lienert, 1969) verstanden werden kann, möchte ich die Probleme anhand der klassischen Gütekriterien für Tests diskutieren.

Nach Bortz (1995) gibt die Objektivität oder Anwenderunabhängigkeit eines Verfahrens an, in welchem Ausmaß die Testergebnisse vom Testanwender unabhängig sind. Er unterscheidet zwischen Durchführungs-, Auswertungs- und Interpretationsobjektivität. Weil das hier diskutierte Verfahren kein klassischer Test ist, bei dem die Objektivität meist vergleichsweise wenig Probleme bereitet, soll der Unterschied betrachtet werden: Auf der Eingabeseite stehen hier nicht die Testwerte eines Probanden, sondern Einschätzungen von Anwendern. Diese Einschätzungen oder Ratings bilden einen wesentlichen potenziellen Schwachpunkt des Verfahrens. Hier müssen Maßnahmen zur Sicherung der Durchführungsobjektivität getroffen und deren Erfolg überprüft werden.

Zwei Maßnahmen zur Sicherung der Durchführungsobjektivität müssen getroffen werden: Erstens müssen die Nutzer durch kontextsensitive Hilfetexte unterstützt werden. Dies ist auf der Eingabeseite realisiert. Zweitens müssen alle Nutzer hinlänglich eingewiesen und geschult sein. Dazu kann an Beispielsachverhalten geübt und anschließend gemessen werden, wie hoch die Übereinstimmung von Nutzern auf der Seite der Dateneingabe ist. Erste Vortests, die jedoch wissenschaftlichen Kriterien noch nicht standhalten, geben Anlass zur Hoffnung, dass diese notwendige Voraussetzung einer hohen Übereinstimmung in ausreichendem Maße erreichbar sein könnte. Hier besteht dringender Handlungsbedarf, dies genauer zu messen. Wichtig in diesem Zusammenhang ist die Frage, wer mit welchem Training als Nutzer dieses Systems zugelassen wird. Eine Bedienung des Systems nur durch Polizeibeamte ist hier, wie bei jedem psychologischen Test, abzulehnen. In Frage kommt lediglich eine Benutzung unter Mitwirkung eines Sozialwissenschaftlers, der sowohl mit dem Instrument als auch mit Tests allgemein vertraut ist und so die nötige kritische Distanz zu den jeweiligen Ergebnissen hat. Eine Schulung für potenzielle Nutzer sollte sich nur an Sozialwissenschaftler und Polizeibeamte richten, die Erfahrung im Bereich „Verhandlungen“ haben. Auswertungs- und Interpretationsobjektivität sind durch die automatische Verrechnung der eingegebenen Werte als gegeben anzusehen.

Die Reliabilität oder Zuverlässigkeit eines Verfahrens gibt den Grad der Messgenauigkeit eines Instrumentes an (Bortz, 1995). Zur Bestimmung der Reliabilität des Verfahrens bieten sich die Bestimmung der Testhalbierungs-Reliabilität (Split-Half-Reliabilität, Äquivalenz) an. Bei der Testhalbierungs-Reliabilität werden pro Proband zwei Werte berechnet, die jeweils auf der Hälfte aller Items beruhen. Die Testhalbierung kann dabei unterschiedlich realisiert werden (Zufallsauswahl von allen Items, erste und letzte Hälfte, gerade und ungerade Nummer, ...). Die Testhalbierungs-Reliabilität entspricht dann der Korrelation der Testwerte beider Testhälften. Im Fall des hier vorgeschlagenen Systems kann mit dieser Methode anhand verschiedener Beispielsachverhalte berechnet werden, in welchem Grad jeweils gleiche Vorhersagen zustande kommen.

Die Validität eines Tests gibt an, wie gut der Test in der Lage ist, genau das zu messen, was er zu messen vorgibt (Bortz, 1995). Dies ist die Kernfrage und im vorliegenden Fall nicht leicht

zu messen. Methodisch gesehen stellt das Modell eine multivariate Zusammenhangshypothese auf, die jedoch in dieser Arbeit noch nicht empirisch untersucht wird. Behauptet wird ein Zusammenhang zwischen einer Anzahl von Prädiktorvariablen und Kriteriumsvariablen. Nach Bortz (1995) erfolgt die Überprüfung einer multiplen Zusammenhangshypothese, die eine Beziehung zwischen mehreren Prädiktorvariablen und einer Kriteriumsvariablen behauptet, über die multiple Korrelation. Dazu werden alle Merkmale der Prädiktoren und des Kriteriums an einer repräsentativen Stichprobe erhoben. Dieses methodisch eigentlich nicht sehr schwierige Vorgehen trifft hier in der Praxis auf erhebliche Probleme.

Erstens sind die Fallzahlen vergleichsweise gering. Viele der Lagen, die in der Statistik erfasst sind, sind bereits beendet, bevor Spezialeinheiten alarmiert und einsatzbereit sind. Auch in den Fällen, in denen die Polizei mit Spezialeinheiten rechtzeitig vor Ort und arbeitsfähig ist, sind nicht alle geeignet, das System zu testen. Gerade bei den sogenannten Bedrohungslagen ist der polizeiliche Kräfteansatz oft so gewählt, dass eine parallele Bearbeitung nicht sichergestellt werden kann. Selbst wenn der Fall eintritt, dass die aktuelle Lage geeignet ist, Prädiktor- und Kriteriumsvariablen zu messen, so hilft dies nur bedingt weiter. Dies ist vor allem dadurch bedingt, dass alle drei Kriteriumsvariablen auch vom Verhandlungsverlauf und anderen polizeilichen Interventionen abhängen. Wenn die Polizei zu dem Ergebnis kommt (mit oder ohne das EUSiG), daß eine Gewaltanwendung seitens des Täters gegen sich oder die Geiseln bevorstehen könnte, wird sie ihr taktisches Vorgehen daran orientieren und entweder eine der Situation angepasste Verhandlungsstrategie wählen oder gewaltsam intervenieren. Es kann also sein, dass das System richtigerweise eine bevorstehende Gewalttätigkeit des Geiselnegers gegenüber einer Geisel anzeigt, zu der es aber nicht kommt, weil die Polizei vorher das Gebäude stürmt und die Lage gewaltsam beendet, um die Geiseln zu schützen. Wie hoch Prädiktoren und Kriterien tatsächlich miteinander interagieren, lässt sich also im Feld nicht messen.

Auch die Messung der Konstruktvalidität ist im vorliegenden Fall schwierig, da hier im Gegensatz zu klassischen Tests keine Konstrukte, sondern konkretes Verhalten vorhergesagt werden soll. Alle Variablen, von denen aber ein Zusammenhang mit dem vorherzusagenden Verhalten bekannt ist, sind selbst Bestandteil des Tests.

Eine Möglichkeit, sich dennoch einer Validitätsmessung zu nähern, besteht in der Messung der konvergenten Validität. Dies könnte realisiert werden, indem Experten verschiedene Fälle zur Beurteilung der fraglichen Kriterien vorgelegt werden. Dieselben Fälle werden dann mithilfe des Systems bearbeitet. Problematisch bei diesem Ansatz ist jedoch, dass die mir bekannten Herangehensweisen der Experten Eingang in dieses System gefunden haben und somit integraler Bestandteil sind. Insofern wären die Messungen nicht gänzlich unabhängig voneinander.

Was weiterhin bleibt, ist die Bestimmung der Inhaltsvalidität (Face Validity, Augenscheinvalidität, Logische Validität). Diese ist nach Bortz (1995) gegeben, wenn der Inhalt der Test-Items (bzw. hier der Prädiktorvariablen) das zu messende Konstrukt in seinen wichtigsten Aspekten erschöpfend erfasst. Inwieweit dies gegeben ist, ließe sich beispielsweise über Expertenratings erheben. Das Ergebnis einer solchen Experteneinschätzung besteht jedoch lediglich aus aufaddierten subjektiven Einschätzungen und bildet kein gutes Maß für die tatsächliche Validität des Verfahrens.

Im Hinblick auf das Thema dieser Arbeit wäre eine auf lange Zeit angelegte Validierungsstudie hilfreich. Diese Studie könnte versuchen, Einsätze nachzubereiten und zur Messung der vorherzusagenden Kriterien auch Daten aus Befragungen der Täter und Geiseln sowie des Polizeiführers und der Verhandler mit einzubeziehen. So könnten auch Verzerrungen durch polizeiliche Interventionen zumindest teilweise erfasst werden. Um eine ausreichende Anzahl von Fällen erfassen zu können, wäre eine bundesweit angelegte Studie

vorteilhaft. Dies wird jedoch aufgrund der länderspezifisch unterschiedlichen Polizeistrukturen und des damit verbundenen hohen Aufwands nur schwer realisierbar sein.

## Literatur

- Amelingmeyer, J. & Stahringer, S. (1999). Expertensysteme als Werkzeuge für das Wissensmanagement. *Praxis der Wirtschaftsinformatik*, 36, 208, 80-92.
- Bortz, J. & Döring, N. (1995). *Forschungsmethoden und Evaluation*. Berlin: Springer.
- Bundeskriminalamt (2000). *Polizeiliche Kriminalstatistik Bundesrepublik Deutschland, Berichtsjahr 1999*. Wiesbaden: BKA.
- Fuselier, G.D. (1981). A Practical Overview of Hostage Negotiations. *FBI Law Enforcement Bulletin*, 10 –15.
- Gatzke, W. (1996). Geiselnahmen, Entführungen, Bedrohungslagen. In M. Knietsch (Hrsg.), *Handbuch für Führungskräfte der Polizei* (S. 427 – 449). Lübeck: Schmidt-Römhild.
- Gist, R. & Perry, R. (1985). Perspectives on Negotiation in Local Jurisdictions. *FBI Law Enforcement Bulletin*.
- Greenstone, J.L. (1995). Tactics and negotiating techniques (TNT): the way of the past and the way of the future. In M.I. Kurke, & E.M. Scrivner (Hrsg.), *Police psychology into the 21<sup>st</sup> century*. Hillsdale: Lawrence Erlbaum.
- Hart, S. (1999). *Approaches to violence risk assessment*. Paper presented at the Pre-Conference Programme of applied courses, International Conference Psychology and Law, Dublin.
- Herbertz, C.W. (1979). *Akzeptieren, Sprechen, Beruhigen. Verhaltensmaßregeln bei Geiselnahmen*. *WS 2*, 45 – 49.
- Krummsiek, L. (1979). Kontaktaufnahme und Verhandlungsführung bei Geiselnahmen. *Kriminalist*, 3, 114 – 118.
- Landeskriminalamt NRW (1999). *Geiselnahmen in der Bundesrepublik Deutschland von 1991 - 1998*. Düsseldorf: LKA NRW.
- Lienert, G.A. (1969). *Testaufbau und Testanalyse*. Weinheim: Beltz.
- Prinz, H. (1995). Geiselnahmen. *Der Kriminalist*, 10, 469 – 471.
- Rueth, T.W. (1993). Onsite psychological evaluation of a hostage taker. *Psychological Reports*, 73, 659-664.
- Selg, H., Mees, U. & Berg, D. (1997). *Psychologie der Aggressivität*. Göttingen: Hogrefe.
- Steffen, W. & Polz, S. (1991). *Familienstreitigkeiten und Polizei: Befunde und Vorschläge zur polizeilichen Reaktion auf Konflikte im sozialen Nahraum*. München: Bayerisches LKA.
- Tröndle, H. & Fischer, T. (Hrsg.) (1999). *Strafgesetzbuch und Nebengesetze*. München: Beck.
- Van Egmond, M. & Diekstra, R.-F. (1990). The predictability of suicidal behaviour: the results of a meta-analysis of published studies. *Crisis*, 11, 2, 57 – 84.
- Welz, R. & Möller, J. (1984). *Bestandsaufnahme der Suizidforschung*. Regensburg: Roderer.
- Westhoff, K. & Kluck, M.L. (1991). *Psychologische Gutachten schreiben und beurteilen*. Berlin: Springer Verlag.



# Maßgeschneiderte psychometrische Verfahren für die familienrechtliche Begutachtung<sup>1</sup>

*Wilfried Hommers*

*Lehrstuhl für Psychologie I der Universität Würzburg*

Der Beitrag untersucht, welche Implikationen die Kindschaftsrechtsreform vom Juli 1998 für die familienrechtliche Diagnostik hat. Dabei wird der Fall-Typ zweier über Sorge oder Aufenthalt eines oder mehrerer Kinder streitender Eltern betrachtet. Es zeigt sich, dass zwei bestehende Mängel in der familienrechtlichen Psychodiagnostik verstärkt wurden, da es schon bisher an maßgeschneiderten Verfahren und derartigen psychometrischen Verfahren fehlte. Die Analysen führen zu dem weiteren Ergebnis, dass sich aus der durch die Reform gestellten forensischen Aufgabenstellung ein neuer psychometrischer Ansatz ergibt, der die direkte Bestimmung von Reliabilitäten für die Vergleiche von auf Mutter und Vater bezogenen Daten erfordert. Für die familienrechtliche Praxis hat dies zur Folge, dass sich die zufallskritische Prüfbarkeit von Testergebnissen gegenüber der indirekten Bestimmung der notwendigen Reliabilitäten wegen möglicher positiver Auswirkungen von Fehlerkorrelationen verbessern kann.

## **Zur maßgeschneiderten Diagnostik im Familienrecht**

Das Fehlen von maßgeschneiderten Verfahren für die familienrechtliche Diagnostik wurde von der Praxis bemängelt (Ell, 1990). Die Beantwortung der Frage, warum familienrechtlich maßgeschneiderte Verfahren fehlten, hat mehrere Aspekte zu berücksichtigen, die sich z.T. durch die Rechtslage ergeben.

Das Fehlen maßgeschneiderter Diagnostik im Falle eines Streits der sich trennenden Eltern um das Kind (Fall-Typ) war zunächst einmal zu bemängeln, weil in der sorgerechtlchen Begutachtung Verfahren angewendet wurden, die aus anderen diagnostischen Kontexten stammten. Der wegen seiner hohen Flexibilität brauchbare FRT (Bene & Anthony, 1985) stammte aus der psychoanalytisch geprägten Erziehungsberatung. Weiterhin wird eine Übertragung der Bindungstheorie nach Bowlby und Ainsworth befürwortet (Schwabe-Höllein, 1996, 2001). Auch wäre an die Anwendung von Verfahren aus der Sozialisationsforschung (z.B.: FDTS, ESI vgl. unten) zu denken. Schließlich sind Verfahren zur Erfassung der mütterlichen Erziehung (z.B.: HAMEL) oder der Erziehungseinstellungen (z.B.: Marburger Erziehungsstilinventar) zu nennen, die aus der Erziehungsberatung oder der Sozialisationsforschung unter Bezug auf die Lewinschen Erziehungsstile bzw. einfache lernpsychologische Maximen entstanden sind.

Der Mangel dieser Ansätze bestand schon vordergründig wegen der unterschiedlichen diagnostischen Aufgabenstellungen bei klinischem Helfen in der Erziehungsberatung, bei wissenschaftlichem Erkennen im Rahmen von Sozialisationsforschung usw. und bei gerichtlichem Entscheiden aufgrund einer familienrechtlichen Kinderdiagnostik bzw. einer familienrechtlichen Elterndiagnostik für den hier betrachteten Fall-Typ. Wegen Untersuchungen unter Trennungsstress, natürlicher Variabilität der Rollenverschiedenheit von Müttern und Vätern, Kontaktungleichheiten aufgrund vorläufiger Regelungen, der Möglichkeit von Beeinflussungen, des zukünftigen Kindeswohls als Zielsetzung trotz Streits um das Kind ist die diagnostische Aufgabe im Familienrecht verschieden von derjenigen bei Erziehungsberatungs- oder anderem klinischen oder wissenschaftlichen Erkenntnis-, Handlungs- oder Diagnosebedarf. Die Frage, wodurch sich "maßgeschneiderte" Verfahren in

der familien- oder sorgerechtlichen Diagnostik auszeichnen oder sogar auszeichnen müssen, hat aber zwei weitere grundsätzliche Momente.

Das erste Moment folgt aus einem der drei besonderen Leitziele der Forensischen Diagnostik (Hommers, 1993), der Kriterienausschöpfung. Demnach sind, was selbstverständlich erscheinen mag, die juristisch vorgegebenen Kriterien auszuschöpfen. Wegen der in diesem Sinne zu erfüllenden inhaltlichen Sachrichtigkeit kommen in den Untersuchungen nur familiendiagnostische Verfahren in Frage, die sich auf die gesetzlich bestimmten Kriterien (FGG § 50b: Bindung, Neigung, Wille des Kindes) oder durch die Rechtsprechung gegebenen Kriterien beziehen (z.B. gemäß immer wieder zu beobachtenden Urteilen des BGH die Kooperationsbereitschaft der Kindseltern bzw. die in den familiengerichtlichen Aufträgen direkt oder indirekt angesprochene Erziehungsfähigkeiten der Kindseltern).

Das zweite Moment ist spezifisch kennzeichnend für die maßgeschneiderte familienrechtliche Diagnostik. Es ergibt sich aus der im Zivilrecht allgemein bestehenden Gleichbehandlungspflicht der Parteien. Mit beiden Elternteilen müsste demnach dasselbe in der psychologischen Untersuchung geschehen sein, damit es sich um eine gerichtlich akzeptable Vergleichsdiagnostik der beiden Zuordnungsalternativen (einer der Eltern) handelt. Das geht z.B. aus der Formulierung hervor, die in der in Bayern üblichen familiengerichtlichen Auftragsstellung enthalten ist: "Die Aufnahme von anderen Kontakten mit nur einer Partei ist geeignet, den Anschein der Parteilichkeit zu erwecken; sie sollte daher vermieden werden". Das muss natürlich auch für die Daten gelten, die vom Kind erlangt werden: Vater und Mutter müssen in den Aufgaben, die das Kind bearbeitet, im Prinzip diagnostisch die gleichen Chancen haben. Daraus ergibt sich, dass das entscheidende Moment maßgeschneiderter familienrechtlicher Diagnostik für den betrachteten Fall-Typ eine im Sinne der Gleichbehandlung faire Vergleichsdiagnostik ist. Das lässt sich erreichen durch Vorgabe gleicher Items für Mutter und Vater beim Kind als Datenquelle bzw. an Mutter und Vater, wenn jene die Daten liefern.

Als Konsequenz ergibt sich weiterhin für die familienrechtliche Diagnostik, dass eine diagnostische Nullhypothese zu prüfen ist: "Vater und Mutter sind aufgrund der erhobenen Daten als gleich geeignet anzusehen". Darin unterscheidet sich maßgeschneiderte Diagnostik im Familienrecht prinzipiell in zweierlei Weise von maßgeschneiderter Intelligenzdiagnostik. Einerseits geht es dort lediglich um die technisch effiziente (Zeit bzw. Mühe sparende) Erfassung, was durch eine der jeweiligen Intelligenzausprägung angepasste Itemvorgabe erfolgen soll (vgl. AID oder die altersabhängigen Einstiege in die Itemvorgaben im HAWIK oder im Binetarium). Andererseits bleibt die familienrechtliche Vergleichsdiagnostik grundsätzlich familiär-ipsativ (innerhalb der Familie) und nicht normativ im Sinne eines Vergleichs mit einer empirischen Norm wie z.B. beim Vergleich mit einer durchschnittlichen Intelligenz. Die familienrechtliche Aufgabe lässt sich etwas pointiert damit umschreiben, dass es bei dem gewählten Fall-Typ unwichtig ist, wo ein Unterschied zwischen den Eltern auf einer Skala liegt, es zählt nur die Existenz eines Unterschieds zwischen den Eltern.

Weiterhin ergibt sich eine methodisch-technische Konsequenz. Bei Vorliegen maßgeschneiderter Verfahren würden jeweils Itempaarungen für Mutter und für Vater existieren, aus denen nach Überführung der Antworten in Zahlencodes Itemdifferenzen „Mutter - Vater“ berechnet und für weitere Auswertungen verwendet werden könnten. Daraus entstehen u.U. weitreichende Folgerungen für die psychometrische Seite der Diagnostik im Familienrecht bei Verwendung maßgeschneiderter Verfahren.



## **Zur psychometrischen Diagnostik im Familienrecht**

Wenn Verfahren den beschriebenen Anspruch der Gleichbehandlung erfüllen, dann können sie zur Widerlegung der diagnostischen Nullhypothese geeignet sein. Zur hinlänglichen Prüfung der Nullhypothese gehört aber die Einbeziehung der zufallskritischen Fragestellung, ob die eventuell aufgewiesenen Unterschiede nach fairer Untersuchung auch statistisch abgesichert werden können. Das weist auf den Wert der Bestimmung psychometrischer Eigenschaften der Verfahren hin.

Aufstellungen über in diesem Anwendungsbereich verwendete Tests zeigen, dass die allgemein anerkannte Strategie der psychometrischen Verfahren bislang einen sehr geringen Stellenwert in der familienrechtlichen Begutachtung besaß (Salzgeber, Müller-Baumgärtel). Das lag z.T. an den üblicherweise verwendeten semi-projektiven Verfahren zur Bindungsdiagnostik mit Kindern, auf deren lediglich die Exploration unterstützenden Charakter Wegener (1981) verwies. Ergebnisse von Exploration und semi-projektiven Verfahren halten in der Regel zufallskritischer Prüfung nicht stand. Das mag aufgrund der den Fällen innewohnenden weiteren multi-methodalen Befunde und der Begutachtungsziele, ein Problem zwischen Eltern lösen zu helfen, wenig bedeutsam erscheinen. Jedoch nimmt man dann in Kauf, dass sich ein derartiges „klinisches“ Vorgehen als „soft psychology“ darstellt (Meehl, 1978).

Die psychometrische Strategie hat aber durch die Kindschaftsrechtsreform 1998 einen besonderen Stellenwert erlangt. Aus der veränderten Rechtslage ergibt sich neben der Gleichbehandlung eine weitere normative Forderung. Danach besteht die Gemeinsame Elterliche Sorge fort, solange keinem Antrag auf Abänderung stattgegeben wird oder werden muss.

Die Kindschaftsrechtsreform hat zunächst einmal zur Folge, dass nun über den Aufenthalt des Kindes die alten Fragen nach der Präferenz eines Elternteils wegen Bindung, Neigungen und Willen des Kindes zu klären sind, da es einen gemeinsamen Aufenthalt nicht geben kann, allenfalls einen gleichverteilt wechselnden, was selbst wohl eher die Ausnahme sein dürfte. Für die angemessene Begutachtung zum Aufenthalt gelten wieder die diagnostische Nullhypothese als Ausgangspunkt und die Norm des Gleichbehandlungsprinzips.

Aber die Kindschaftsrechtsreform hat die geforderte Prüfung der diagnostischen Nullhypothese für die sorgerechtliche Begutachtung durch die Einführung einer expliziten gesetzlichen Nullhypothese verändert, was Konsequenzen für die Beweislast hat. Früher mußte man einen vermeintlich besseren Elternteil quasi zwangsweise ausmachen, weil aufgrund der Verweigerung der Gemeinsamen Elterlichen Sorge seitens eines Elternteils eine Entscheidung für die Alleinige Elterliche Sorge eines Elternteils unumgänglich war. Möglicherweise konnte das aufgrund der multi-trait-multi-method-Befunde überzeugend geschehen. Jetzt geht es aber um die Prüfung der gesetzlich vorgegebenen Nullhypothese einer im Prinzip trotz Trennung oder Scheidung fortbestehenden Gemeinsamen Elterlichen Sorge. Die gesetzliche Alternative der Alleinigen Elterlichen Sorge tritt nur ein, wenn die Befunde dafür stichhaltigen Anlass geben, von der Widerlegung der Nullhypothese der Gemeinsamen Elterlichen Sorge oder des gleichgünstigen Aufenthalts auszugehen.

Daraus ergibt sich als Konsequenz der Kindschaftsrechtsreform, dass sich damit für die Alleinige Sorge und für den Aufenthalt unterschiedliche Beweislastanforderungen ergeben

haben. Das müsste sich in einem jeweils anderen statistischen Risiko für ein Präferenzergebnis zugunsten eines Elternteils in der durchzuführenden sorgerechtlichen Vergleichsdiagnostik ausdrücken. Wenn früher etwa, wie bei klinischen Fragestellungen (Huber, 1973; Davis, 1959) eine festgestellte Präferenz auf dem 20%-Niveau für eine Alleinige Sorge zählen durfte, weil man Begründungen in eine Richtung finden musste, spricht die Interpretation der Gemeinsamen Elterlichen Sorge als gesetzlicher Nullhypothese dafür, dass eine Präferenzfeststellung pro Alleiniger Sorge nun nur noch mit mindestens dem 5%-Niveau wie in der Experimentalpsychologie üblich begründet werden könnte. Anders dagegen wäre beim strittigen Aufenthalt zu verfahren, wenn überhaupt nur ein überwiegender Aufenthalt bei einem Elternteil in Frage kommen würde. Da könnte man sich schon mit dem 20%-Niveau als Begründung eines Präferenzergebnisses begnügen.

Man beachte im übrigen den Einklang zwischen gesetzlicher und diagnostischer Nullhypothese einerseits sowie experimentalpsychologischer Verwendung des Begriffs der Nullhypothese andererseits. In der Experimentalpsychologie besteht der Regelfall der Prüfung der Nullhypothese aus einem Vergleich zweier oder mehrerer empirischer Statistiken (z.B. t-Test, Varianzanalyse). Daher entspricht der regelhafte Gebrauch des Begriffs der Nullhypothese für die familienrechtliche Fragestellung völlig dem Vergleich von Daten über die beiden Elternteile.

### **Analyse familiendiagnostischer Verfahren**

Unter der Perspektive der aus dem Gleichbehandlungsprinzip erwachsenden fairen Vergleichsdiagnostik sind bestehende familiendiagnostische Test- und Untersuchungsverfahren unter exemplarischer Auswahl darauf zu prüfen, ob sie familienrechtlich maßgeschneidert sind und psychometrischen Konstruktionsprinzipien genügen. Diese Fragen lassen sich am Beispiel der testpsychologischen Erfassung der emotionalen Beziehungen zwischen Kind und Eltern und der Erziehungsfähigkeiten der Kindseltern untersuchen. Neben Verfahren für Kinder und Verfahren für Eltern werden auch integrative Verfahren, die von Seiten des Untersuchers getrennt für die beiden Elternteile zu beantworten sind und in linearer Weise durch Summenwerte ausgewertet werden, bezüglich der zugleich maßgeschneiderten und psychometrischen Qualitätsanforderung untersucht. Das diesbezügliche Fazit vorwegnehmend ergibt sich, dass aus dem formalen Grund der geforderten Gleichbehandlung nur einige deutschsprachige Verfahren grundsätzlich in Frage kommen, während in den USA dieser Anforderung entsprechende Verfahren schon erstellt worden sind.

#### **US-amerikanische maßgeschneiderte Ansätze**

Hierzu (vgl. Tabellen 1 und 2) sind zunächst von Bricklin mehrere Verfahren entwickelt worden, in denen die Kinder (PORT für Kinder im Vorschulalter, BPS für Kinder ab sechs Jahren) oder die Kindseltern (PASS) oder der Begutachter (PCCP) als Datenquellen wirkten. Für diese Verfahren wurden keine Reliabilitätsangaben gemacht. Jedoch wurde das Gleichbehandlungsprinzip weitgehend realisiert. Dann gab es weitere Entwicklungen, in denen auch die Reliabilität als Testgütekriterium mitgeteilt wurde: Der PCRI mit den Kindseltern als Datenquelle und einer großen, vermutlich repräsentativen Eichstichprobe, der Custody Quotient und der ASPECT mit dem Gutachter als Datenquelle.

Alle US-amerikanischen Verfahren zeichnen sich dadurch aus, dass das zivilrechtliche Fairnessgebot der Gleichbehandlung beachtet wurde. Also beurteilten die Kinder bezüglich

Vater oder Mutter dieselbe Fragen oder bearbeiteten dieselben Aufgaben (teils semi-projektiv PORT, teils direkt skalierend BPS). Auch beantworteten die Eltern dieselben Fragen oder beurteilte der Begutachter dieselben Merkmale bezüglich beider Elternteile (diese Merkmale waren teilweise sehr komplex, z.B. "alle klinischen Skalen des MMPI unauffällig").

Bemerkenswert ist, dass interne Konsistenzen (im Sinne von Cronbachs Alpha) getrennt für Vater, Mutter und für die Differenz "Mutter minus Vater" bestimmt werden könnten, da jeweils die gleichen Items zur Anwendung kamen. Jedoch geschah das auch in den psychometrisch weiter ausgearbeiteten Verfahren des PCRI, Custody Quotient und ASPECT nicht. Die Reliabilitätsschätzungen erfolgten soweit nachvollziehbar global, d.h. in Kombination aller Daten, also nicht getrennt für Väter oder Mütter. Für Einschätzungen von Vätern oder Müttern durch ihre Kinder wurden allerdings bislang überhaupt keine Reliabilitäten angegeben. Nur Bricklin erhob direkt von Kindern Daten, untersuchte aber die Reliabilität seiner Verfahren nicht.

### Deutschsprachige Verfahren

Eine Unterteilung nach dem Alter der Kinder erscheint nötig. Zwei Verfahren erfüllen die zivilrechtliche Anforderung hinsichtlich maßgeschneiderter Diagnostik, der Familien- und Kindergarten-Interaktionstest, FIT-KIT, von Sturzbecher & Freytag (2000) und das Erziehungsstilinventar, ESI, von Krohne & Pulsack (1995), obwohl beide nicht direkt mit dieser familienrechtlichen Perspektive konstruiert wurden. Das Familiendiagnostische Testsystem, FDTS, von Schneewind et al. (1985) jedoch nicht (vgl. die Tabellen 3 und 4).

Das FDTS für Familien mit 9- bis 14-jährigen Kindern (vgl. Tabelle 3) steht für den Fall der Verletzung des zivilrechtlichen Gleichbehandlungsgebots. Die Verfahren für Väter und die für Mütter als Datenquellen und die beiden Verfahren über Vater und Mutter aus Sicht von 9-14-jährigen Probanden unterscheiden sich jeweils in den Skalen und in den Items. Daher sind trotz ansonsten sorgfältiger psychometrischer Entwicklung des FDTS die Differenzen "Mutter minus Vater" pro Item nicht bestimmbar. Aber es bleibt vor allem unklar, wie man faire Vergleiche zwischen Vater und Mutter auf den Skalen des FDTS für die familienrechtlichen Fragestellungen, insbesondere bei dem Fall-Typ "Sorge und Aufenthalt" auf wissenschaftlicher Basis unter der Gleichbehandlungspflicht durchführen sollte.

Das ESI (vgl. Tabelle 4 oben) kennzeichnet sich dagegen dadurch aus, dass das mindestens 8 Jahre alte Kind die gleichen Fragen über beide Elternteile beantwortet. Das ESI ist also ein faires familiengerichtliches Verfahren. Es macht auch auf Itemebene die Differenzen "Mutter minus Vater" bestimmbar, was allerdings nicht durchgeführt wurde.

Ein sehr neues Verfahren für die Untersuchung von 4 bis 8-jährigen Kindern ist der FIT-KIT (vgl. Tabelle 4 unten). Es ist am FRT angelehnt und hat Ähnlichkeiten zum US-amerikanischen BPS von Bricklin, weil das Kind Kompetenzen des Elternteils einschätzt. Der FIT-KIT erfüllt wie das ESI die Gleichbehandlungsforderung, da das Kind für jeden Elternteile dieselben Fragen beantwortet. Aber auch bei ihm sind die Differenzen "Mutter minus Vater" pro Item zwar bestimmbar, aber nicht für eine Bestimmung diesbezüglicher interner Konsistenz verwendet worden.

Auch für den anders ansetzenden Familien-Identifikations-Test (FIT) von Remschmidt & Mattejat (1999), wo korrelativ Ähnlichkeiten von Real- und Ideal-Bild zwischen Kind und Eltern bestimmt werden, fragt man über Vater und Mutter die gleichen Inhalte ab. Auch der FIT ist insofern ein familienrechtlich maßgeschneidertes Verfahren. Jedoch ist der FIT in der Vorgehensweise nicht an dem üblichen Summierungskonzept von Messungen orientiert.

## Differenzreliabilität und korrelierte Fehler

Eine erste gemeinsam zutreffende Feststellung für alle besprochenen Verfahren ist, dass die angegebenen Reliabilitäten der einzelnen vorangehend beschriebenen Skalen gering ausfallen. Daher würden sie nach den Kriterien von Lienert & Raatz (1994) nicht für die Einzelfalldiagnostik in Frage kommen. Eine weitere gemeinsam zutreffende Feststellung für alle Verfahren, die das Gleichbehandlungsprinzip erfüllen, ist, dass sie die Möglichkeiten nicht ausnutzen, die sich aus der Bildung von Itemdifferenzen ergeben können. Zunächst wäre zu erläutern, wie die Bildung von Itemdifferenzen zu einer Verbesserung der psychometrischen Möglichkeiten familienrechtlicher Einzelfalldiagnostik führt.

Ein Ansatzpunkt hierfür liegt in der Korreliertheit der Daten über Vater und Mutter. Wenn z.B. ein Kind, wie zivilrechtlich zu fordern wäre, sich über beide Elternteile anlässlich der gleichen Fragen äußert, dann ist diese "Datenquelle korreliert" und es ist möglich, dass nicht nur die Messergebnisse, sondern auch die Fehler dieser Datenquelle selbst miteinander korreliert sind. Es fragt sich, welche Auswirkungen dieser Umstand haben könnte. Dazu muss man die vollständige Formel der Klassischen Testtheorie bei korrelierten Fehlern für die Bestimmung der Reliabilität von Differenzen betrachten.

Die vollständige Formel der Klassischen Testtheorie für die Bestimmung der Reliabilität von Differenzen (Tabelle 5) zeigt, dass sich aufgrund positiv korrelierter Fehler die Reliabilität der Differenzen erhöhen kann (Irtel, 1996; Williams & Zimmermann, 1977). Daher könnten sich als Folge korrelierter Fehler trotz der geringen Reliabilitäten der separat berechneten Summenscores (für Vater- bzw. Mutter-Einschätzungen oder für die Angaben von Vätern oder Müttern oder für die jeweiligen Teil-Beurteilungen durch den Gutachter) u.U. höhere Reliabilitäten für den Vergleich der Vater- bzw. Mutter-Einschätzungen ergeben. Dadurch würden die Konfidenzintervalle kleiner werden, so dass die diagnostischen Schlussfolgerungen eher der Zufallskritik standhalten würden.

Da aber die Fehlerkorrelation nicht direkt bestimmbar ist, folgt aus der Realisierung des Gleichbehandlungsprinzips durch Verwendung gleicher Items der einzig gangbare Weg, diese theoretischen Implikationen der Formel in Tabelle 5 über eine direkte Bestimmung der Differenzreliabilität aus den Itemdifferenzen auszunutzen.

Die Differenzreliabilitäten aus den Itemdifferenzen ließen sich über Cronbachs Alpha direkt bestimmen. Aber weder ESI noch FIT-KIT nutzten diesen Umstand, obwohl die Tabelle 8 des Manuals des ESI (Krohne & Pulsack, 1995, S. 29) die Korrelation zwischen Mutter- und Vaterversionen berichtete. Durch die psychometrischen Normen der weiteren dortigen Tabelle 4 wären zwar die aufgrund der KTT bei unkorrelierten Fehlern berechenbaren Reliabilitäten für Vater-Mutter-Differenzwerte in traditioneller Weise bestimmbar (Lienert & Raatz, 1994, bzw. Amelang & Zielinski, 1997). Eine direkte Schätzung der Differenzreliabilitäten erfolgte bisher aber nicht, obwohl sie angesichts der Itemidentitäten im Gegensatz zum FDTS möglich wäre. Der FIT-KIT erfüllt ebenfalls die Voraussetzungen zur Bearbeitung des Problems der korrelierten Fehler in der sorgerechlichen Familiendiagnostik. Jedoch fehlen auch hier die entsprechenden Erhebungen und Auswertungen. Korrelationen zwischen Vater- und Mutter-Bewertungen wurden z.B. nicht berichtet und die direkte Bestimmung von Differenzreliabilitäten wurde nicht vorgenommen, so dass möglicherweise kleinere kritische Differenzen aufgrund der exakten Differenzreliabilitäten nicht nutzbar sind.

## Fazit

Für die familienrechtliche Kinderdiagnostik wäre zu fordern, dass die direkte Bestimmung von Differenzreliabilitäten z.B. für ESI bzw. FIT-KIT erfolgt. Für die familienrechtliche Elterndiagnostik wären Verfahren wie der US-amerikanische PCRI unter Berücksichtigung der direkten Bestimmung von Differenzreliabilitäten zu entwickeln. Außerdem wären auch Neukonstruktionen denkbar, die dem Gleichbehandlungsprinzip genügen und die bezüglich der psychometrischen Eigenschaften umfassend untersucht worden sind.

## Literatur

- Amelang, M. & Zielinski, W. (1997). *Psychologische Diagnostik und Intervention*. Berlin: Springer.
- Bene, E. & Anthony, J. (1985). *Family Relations Test*. Windsor: NFER-Nelson Publishing Co.
- Davis, F.B. (1959). Interpretation of differences among averages and individual test scores. *Journal of Educational Psychology*, 40, 162-170.
- Ell, E. (1990). *Psychologische Kriterien bei der Sorgerechtsregelung und die Diagnostik der emotionalen Beziehungen*. Weinheim: Beltz.
- Hommers, W. (1993). Psychometrische Modelle für die Einzelfalldiagnostik in der Forensischen Psychologie. In L. Montada (Hrsg.), *Bericht über den 38. Kongress der Deutschen Gesellschaft für Psychologie in Trier 1992. Band II* (S. 206-213). Göttingen: Hogrefe.
- Huber, H. (1973). *Psychometrische Einzelfalldiagnostik*. Weinheim: Beltz.
- Irtel, H. (1996). *Entscheidungs- und testtheoretische Grundlagen der Psychologischen Diagnostik*. Frankfurt: Peter Lang.
- Krohne, H.W. & Pulsack, A.U. (1995). *Das Erziehungsstil-Inventar*. Göttingen: Beltz Test.
- Lienert, G.A. & Raatz, U. (1994). *Testaufbau und Testanalyse*. München: PsychologieVerlagsUnion.
- Mayer-Baumgärtel, B. (1996). Eine Auswahl von Testverfahren und Fragebogen zur Diagnostik in familienrechtlichen Fragestellungen. *Familie, Partnerschaft, Recht*, 2, 176-178.
- Meehl, P.E. (1978). Theoretical risks and tabular asterix: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 46, 805-818.
- Remschmidt, H. & Matzejat, F. (1999). *Familien-Identifikations-Test (FIT)*. Göttingen: Testzentrale.
- Salzgeber, J. (1992). *Der psychologische Sachverständige im Familiengerichtsverfahren*. München: Beck.
- Schneewind et al. (1985). *Familiendiagnostisches Testsystem*. München: Eigenvertrieb.
- Schwabe-Höllerlein, M. (2001). *Qualitätssicherung bei der Bindungsdiagnostik im familiengerichtlichen Verfahren*. Unveröffentlichter Vortrag: 2. Tage der Rechtspsychologie Leipzig.
- Sturzbecher, D. & Freytag, R. (2000). *Familien- und Kindergarten-Interaktionstest (FIT-KIT)*. Göttingen: Testzentrale.
- Wegener, H. (1981). *Einführung in die Forensische Psychologie*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Williams, R. H. & Zimmermann, D. W. (1977). The reliability of difference scores when errors are correlated. *Educational and Psychological Measurement*, 37, 679-689.

## **Tabelle 1: Maßgeschneiderte Ansätze in den USA**

### **Das Bricklin-System:**

Village Publishing, Furlong, PA 18925

- PORT, Perception-of-Relationships-Test

Seit 1962, 7 Items projektiver Art für Kinder ab 3 Jahre für Vater und Mutter gleich. Kritik: Keine Nachweise für Reliabilität oder Validität, für letzteres allerdings Eigenkontrolle hinsichtlich der Übereinstimmung von gutachterlichem Vorschlag mit Richter-Entscheid.

- BPS, Bricklin-Perceptual-Scales

Seit 1962, 33 Fragen für die Beurteilung seitens des Kindes von "Kompetenzen" der Eltern in Fähigkeiten, Unterstützung, konsistentem Erziehungsverhalten und bewunderten Eigenschaften, ab 6 Jahre, je für Mutter und Vater gleich.

Kritik: Keine Nachweise für Reliabilität und Validität, jedoch Idee der elterlichen Vergleichsdiagnostik. Da es für Vater und Mutter die gleichen Fragen gibt, wäre die Bestimmung der Fehlerkorreliertheit durchführbar.

- PASS, Parent Awareness Skills Survey

Seit 1987, explorativer Elternvergleich mit schwierigen Erziehungssituationen.

- PPCP, Parent Perception of Child Profile

Seit 1991, Versuch einer Gesamtintegration von explorativen Daten über Wissen vom und Verstehen eines Kindes durch den Gutachter.

### **The Custody Quotient (Joint Custody)**

R. Gordon & L. A. Peek (1987). Dallas: The Wilmington Institute.

Es handelt sich um eine Zusammenstellung von gewichteten und standardisierten Beurteilungen beider Eltern seitens des Gutachters mit der Zielsetzung einer Gesamtbewertung. Es gibt eine separate Skala für eine Joint-Custody-Fragestellung. Die psychometrischen Angaben stammen aus einer Stichprobe von 125 Elternpaaren, die sich in einem Sorgestreit befanden. Es gibt 12 Skalen mit Werten für Cronbachs Alpha, die zwischen den beiden kleinsten von 0.30 (PL, Planning for the child's needs), 0.58 (PA, Parent Assistance) und dem Maximum von 0.94 (AO, Misconducts, Acts and Omissions of Parents) liegen.

### **PCRI, Parent-Child-Relationship-Inventory**

Gerard (1994). Los Angeles: Western Psychological Services.

Das PCRI besteht aus 7 inhaltlich begründeten Skalen über Einstellungen der Eltern zu ihrer Rolle, zum Verhältnis zum Kind und zur Erziehung. Außerdem gibt es zwei Kontrollskalen zur Inkonsistenz und zur Sozialen Erwünschtheit der Antworten. Mütter und Väter beantworten die gleichen Fragen.

## **Tabelle 2: Maßgeschneiderte Ansätze in den USA**

### **ASPECT**

#### **(Ackermann-Schoendorf Scales for Parent Evaluation of Custody)**

M. J. Ackerman, & K. Schoendorf, 1992, Los Angeles, Western Psychological Services

Das Inventar ASPECT stellt Beurteilungen durch den Gutachter und Einzelerhebungen aus unterschiedlichen Quellen zusammen: Aktuarische (... history of psychiatric hospitalization? ... of taking psychiatric medication? ... of substance abuse?), explorative (Has the parent resolved his or her own conflicts regarding the divorce? .. shown dependence on the child?) und Test-Daten (Rorschach, 4 Abfragen; MMPI, 7 Abfragen, IQ-Differenz zum Kind  $< 5$ , HAWIE-AV-WP  $> 9$  und Schulbildung der 9. Klasse). Diese erfolgen jedoch getrennt und in gleicher Weise sowohl für Vater als auch für Mutter. Derartige 9, 28 und 19 Abfragen bilden drei Skalen: Observational Scale, Social Scale und Cognitive-Emotional Scale. Die Abfragen letzterer sind zuvor aufgeführt. Als Rohwerte gelten die erreichten Punkte in Prozent, die in T-Werte oder Prozentrangnormen transformiert werden können. Für diese drei Skalen und den Gesamtwert (Parental Custody Index) werden Reliabilitäten von  $\text{Alpha}=0.50$ ,  $0.72$  (Social Scale),  $0.50$  und  $0.76$  (Gesamtwert), sowie deren Standardabweichungen und Mittelwerte angegeben. Als interpretierbar gelten bei jeder Skala und dem Gesamtwert T-Wert-Differenzen von 10 und mehr (das wäre eine Irrtumswahrscheinlichkeit von ca. 10% beim Gesamtwert und bei der Social Scale und eine höhere bei den beiden anderen Skalen!). Die Objektivitätsindizes der Beurteilungen (interrater reliability) liegen zwischen  $0.92$  (Observational Scale) und  $0.96$  (Gesamtwert).

Es gibt 1, 7 und 4 sogenannte Critical Items, um “more serious deficits in parenting” (p. 35) anzuzeigen (Appreciation of the effects of the divorce for the child? Arrest record?

Alcoholic? Charged with sexual abuse? Charged with physical abuse? Convicted of physical abuse? Cooperation with previous court orders regarding custody, visitation or placement? All clinical MMPI-scales below  $T=80$ ? Thought disorder in Rorschach? Psychiatric medication? History of substance abuse?). Wie man mit den Critical Items allerdings psychometrisch oder interpretativ umgehen soll, blieb offen. Denkbar wäre hier, einen Chi-Quadrat-Test für die Mutter-Vater-Differenz durchzuführen: Erfüllt ein Elternteil signifikant mehr kritische Items als der andere, würde das gegen ihn sprechen.

### **Tabelle 3: FDTS ein maßgeschneiderter Ansatz?**

Vier Skalen zu den Erziehungspraktiken: Die Familienmitglieder beantworten Items zu 7 Skalen. Die Items waren fünfstufig von "nie" bis "immer".

#### **Mutter oder Vater als Proband**

(MEPS-M oder VEPS-V mit jeweils 53 Items)

##### **Skala A: Belohnung durch liebevolle Zuwendung**

(10 oder 11 Items, Rel: 0.93 oder 0.95,  $Dif_{krit,5\%}=1.36$ )

##### **Skala B: Materielle Belohnung und Verstärkung durch angenehme Aktivitäten**

(11 oder 7 Items, Rel: 0.95 oder 0.94,  $Dif_{krit,5\%}=1.30$ )

##### **Skala C: Eingeschränktes Lob**

(4 oder 6 Items, Rel: 0.94 oder 0.95,  $Dif_{krit,5\%}=1.30$ )

##### **Skala D: Bestrafung durch Liebesentzug**

(6 oder 5 Items, Rel: 0.90 oder 0.86,  $Dif_{krit,5\%}=1.92$ )

##### **Skala E: Bestrafung durch Entzug materieller Verstärker, Privilegien oder durch Anordnung unangenehmer Aktivitäten**

(11 Items, Rel: 0.95 oder 0.94,  $Dif_{krit,5\%}=1.30$ )

##### **Skala F: Verbaler Ärgerausdruck und Geringschätzung**

(8 oder 7 Items, Rel: 0.92 oder 0.85,  $Dif_{krit,5\%}=1.88$ )

##### **Skala G: Körperliche Bestrafung**

(3 Items, Rel: 0.87)

##### **oder Geringschätzung**

(6 Items, Rel: 0.91,  $Dif_{krit,5\%}=1.84$ )

#### **Sohn als Proband über Mutter oder Vater**

(MEPS-S oder VEPS-S mit 44 oder 46 Items insgesamt)

##### **Skala A: Belohnung durch liebevolle Zuwendung**

(11 oder 9 Items, Rel: 0.95 oder 0.96,  $Dif_{krit,5\%}=1.18$ )

##### **Skala B: Materielle Belohnung und Verstärkung durch angenehme Aktivitäten**

(6 oder 9 Items, Rel: 0.95 oder 0.87,  $Dif_{krit,5\%}=1.66$ )

##### **Skala C: Eingeschränktes Lob**

(4 oder 4 Items, Rel: 0.94 oder 0.94,  $Dif_{krit,5\%}=1.36$ )

##### **Skala D: Bestrafung durch Liebesentzug**

(3 oder 4 Items, Rel: 0.90 oder 0.89,  $Dif_{krit,5\%}=1.80$ )

##### **Skala E: Bestrafung durch Entzug materieller Verstärker, Privilegien oder durch Anordnung unangenehmer Aktivitäten**

(9 oder 13 Items, Rel: 0.94 oder ???,  $Dif_{krit,5\%}=???$ )

##### **Skala F: Verbaler Ärgerausdruck und Geringschätzung**

(9 Items, Rel: 0.93)

##### **oder Bestrafung durch Geringschätzung, langanhaltendem Ärger, Anordnung unangenehmer Aktivitäten**

(13 Items, Rel: 0.96,  $Dif_{krit,5\%}=1.30$ )

##### **Skala G: Körperliche Bestrafung**

(2 Items, Rel: 0.87)

##### **oder Verbaler Ärgerausdruck und körperliche Bestrafung**

(7 Items, Rel: 0.94,  $Dif_{krit,5\%}=1.72$ )



## **Tabelle 4: Zwei maßgeschneiderte Verfahren**

### **Erziehungsstilinventar ESI**

Beim "Erziehungsstilinventar", ESI, von Krohne & Pulsack (2. Auflage, 1995) liegen die Verhältnisse anders als beim FDTs von Schneewind et al. (1985). Das ESI ist für den Altersbereich 8 bis 16 normiert und präsentiert 60 Items gleichen Wortlauts und 5 weitere Items gleichen Wortlauts (für die Straftintensität) zur Beschreibung von Vater und Mutter. Daher ist eine im familienrechtlichen Sinne maßgeschneiderte Vergleichsdiagnostik der Eltern aus der Perspektive des Kindes auf den Erziehungsstil-Dimensionen Unterstützung (US), Einschränkung (ES), Lob (LS), Tadel (TS), Inkonsistenz (IK) und Straftintensität (SI) formal möglich.

### **Familien- und Kindergarten-Interaktionstest FIT-KIT**

Der FIT-KIT von Sturzbecher & Freytag (2000) erfasst Interaktionen zwischen 4- bis 8-jährigen Kindern und Müttern sowie Vätern aus Kindersicht, ohne dass dabei die Items für die Eltern variieren. Im Unterschied zum FRT (Bene & Anthony, 1985) werden nicht gefühlsbezogene Aussagen Personen zugeordnet, sondern Verhalten des Kindes bzw. der Personen 3 (oder 4) Häufigkeitsklassen zugeordnet. Die 63 Fragen teilen sich auf in 14 Fragen nach dem Kindesverhalten und in 49 Fragen nach dem Eltern- oder Erzieherverhalten. Die 14 Fragen nach dem Kindesverhalten bestimmen die Werte dreier Skalen (Itemanzahlen) Hilfesuche (5), Diplomatie (4) und Renitenz (5), jeweils des Kindes aus Kindersicht und die anderen 49 Fragen bestimmen die Werte der Skalen Kooperation (8), Hilfe (6), Abweisung (8), Restriktion (4), Bekräftigung (5), Trösten (5), Emotionale Abwehr (6) und Faxen/Toben (5) des Erwachsenen mit dem Kind. Es werden, aufgrund bestehender Schiefe bei 7 Skalen teils problematische, Stanine-Normen für die Rohwertbereiche der 11 Skalen angegeben, wobei die Rohwertbereiche für die 3-kategoriale Version zwischen 12 und 24 liegen. Diese sind separat für Mütter, Väter und Erzieher ausgeführt. Allerdings unterscheiden sich die Vater- und Mutter-Normen nur geringfügig, was man den teststatistischen Befunden über die Items (Kapitel 4, S. 68 ff.) entnehmen kann. Die Reliabilitäten der Skalen lagen zwischen  $\text{Alpha}=.39$  und  $\text{Alpha}=.80$  bei Berücksichtigung der Altersgruppierungen. Der Standardmessfehler für die Differenz von Stanine-Werten variiert daher so, dass die kritischen Differenzen (bei  $p<.05$ ) für die Vergleiche von Vater und Mutter zwischen 3.86 (Restriktion) und 2.51 (Faxen/Toben) auf der Stanine-Skala liegen (Reliabilitäten unter Zusammenfassung der Altersgruppen). In dem Anwendungsbeispiel des Manuals trifft das auf die Skalen Diplomatie, Renitenz, Kooperation, Abweisung, Trösten, Faxen/Toben zu.

## **Tabelle 5: Korrelierte Fehler und Reliabilität von Differenzen**

Für den allgemeineren Fall ungleicher Varianzen und nicht erfülltem vierten Axiom der Klassischen Testtheorie (KTT), d.h.  $\text{KOR}(E_x, E_y) \neq 0$ , bzw. nicht vorliegender lokaler Unkorreliertheit (Irtel, 1996, S. 98), bestimmt sich die Reliabilität von Differenzen zwischen  $x$  und  $y$  zu

$$\text{REL}(D) = \frac{\frac{1}{2} \left( \frac{\text{VAR}_x}{\text{VAR}_y} \text{REL}_x + \frac{\text{VAR}_y}{\text{VAR}_x} \text{REL}_y \right) - \text{KOR}(x, y)}{\frac{1}{2} \left( \frac{\text{VAR}_x}{\text{VAR}_y} + \frac{\text{VAR}_y}{\text{VAR}_x} \right) - \text{KOR}(x, y)}$$

$$+ \frac{\text{KOR}(E_x, E_y) \sqrt{(1 - \text{REL}_x)(1 - \text{REL}_y)}}{\frac{1}{2} \left( \frac{\text{VAR}_x}{\text{VAR}_y} + \frac{\text{VAR}_y}{\text{VAR}_x} \right) - \text{KOR}(x, y)}$$

Man sieht, dass bei korrelierten Fehlern mathematisch gesehen die Reliabilität der Differenzen größer werden kann als bei unkorrelierten Fehlern. Vorstellungsmäßig lässt sich das wie folgt erläutern: Die Fehlerwerte heben sich bei positiver Korrelation durch die Differenzenbildung auf. Zwei hohe positive Fehler in gleicher Richtung ergeben subtrahiert Null: Positiv minus Positiv gleich Null, Negativ minus Negativ gleich Null. Also „stören“ sie in der Differenz der beiden Messungen nicht mehr.

<sup>1</sup> Widmung: Prof. Dr. Dr. Hermann Wegener zum 80. Geburtstag.

# Zum Einfluss des Tatmotivs auf eine Voraussetzung der strafrechtlichen Verantwortungsreife

*Wilfried Hommers & Martin Lewand*  
*Lehrstuhl für Psychologie I der Universität Würzburg*

Ein neuer psychometrischer Ansatz, das Gedanken-Paradigma, wird zur empirischen Fundierung der Altersgrenze der strafrechtlichen Verantwortlichkeit von 14 Jahren vorgestellt. Ausgehend von dem Straftatbestand eines Einbruchsdiebstahls und des variierten Tatmotivs (altruistisch bzw. egoistisch) beurteilten 263 Probanden zwischen 7 und 16 Jahren Gedanken eines Straftäters, die sich inhaltlich auf die ersten vier Stufen der Moralentwicklungstheorie Kohlbergs und die moralische Pflicht der Wiedergutmachung bezogen, im Hinblick auf dessen Strafzumessung. Bereits 10-Jährige waren unabhängig vom Tatmotiv zur Beachtung der Ordnungsgefährdung bei der Strafzumessung fähig (Inhalt der Kohlbergstufe IV). Schon ab einem Alter von 10 Jahren wies unter allen Kohlbergstufeninhalten die Ordnungsgefährdung den größten Strafzumessungseffekt auf. Bei 14-Jährigen hatte zudem die Ordnungsgefährdung auch im Vergleich zur Wiedergutmachungsabsicht den größten Strafzumessungseffekt. Die Ergebnisse sprechen für eine Stützung der Altersgrenze der strafrechtlichen Verantwortlichkeit von 14 Jahren.

Mit einem neuen experimentellen Ansatz zur Bestimmung der Wirkung von moralischen Informationen, dem Gedanken-Paradigma (Hommers, 1997), wird versucht, einen Beitrag zur empirischen Fundierung der Altersgrenze der strafrechtlichen Verantwortlichkeit nach § 19 Strafgesetzbuch (StGB) von 14 Jahren zu leisten. Die Betrachtungsebene des Beitrags beschränkt sich hierbei auf die Entwicklung der strafrechtlichen Unrechtserkenntnisfähigkeit und damit auf den kognitiven Teilaspekt der strafrechtlichen Verantwortlichkeitsreife Jugendlicher nach § 3 Jugendgerichtsgesetz (JGG). Der Einbezug des Tatmotivs dient der Validierung des Ansatzes.

## Schnittpunkte zwischen Psychologie und Jurisprudenz

Die empirisch bearbeitbaren Fragestellungen zur strafrechtlichen Verantwortungsreife sind abhängig von den beiden grundsätzlichen Zielrichtungen zwischen Psychologie und Jurisprudenz. Während es *de lege lata* darum ginge, einzelne Täter nach Vollendung des 14. Lebensjahres hinsichtlich ihrer Tat zu beurteilen, müsste *de lege ferenda* geprüft werden, ob Entwicklungsverläufe bestimmter, zur strafrechtlichen Verantwortlichkeitsreife notwendiger Fähigkeiten mit der Altersgrenze in Einklang zu bringen sind. Die Aufgabenstellung aus *de lege lata* wird im Folgenden wegen deren doppelter Spezifität (Täter und Tat) ausgeklammert. Die Aufgabenstellung *de lege ferenda* lässt sich jedoch anhand der durchschnittlichen Ausprägungen der in Betracht kommenden Fähigkeiten über eine geeignete Altersspanne hinweg untersuchen. Gleichzeitig ergeben sich hieraus jedoch auch Ansatzpunkte *de lege lata*, indem das im folgenden vorgeschlagene Verfahren inhaltlich für den nach § 3 JGG zu beurteilenden Täter auf dessen Tat maßgeschneidert werden würde.

Welche Fähigkeiten der Gesetzgeber als hinreichend für die strafrechtliche Verantwortlichkeitsreife Jugendlicher bis zur Volljährigkeit nach § 2 BGB voraussetzt, sind § 3 JGG zu entnehmen. Danach sind Jugendliche für die von ihnen begangenen strafbaren Handlungen nur dann strafrechtlich verantwortlich, wenn sie zum Zeitpunkt der Tat nach ihrer seelischen und geistigen Entwicklung reif genug waren, das Unrecht der Tat einzusehen und nach dieser Einsicht zu handeln. Peters (1967) führte hierfür den Begriff der Sozialreife ein, worunter er die Fähigkeit verstand, das Unrecht der Tat aus der sozialen Bindung heraus zu begreifen und um dieser Sozialbindung willen sein Handeln rechtmäßig gestalten zu können.

Aus der Reihenfolge der Kriterien der strafrechtlichen Verantwortlichkeitsreife in § 3 JGG und der dazugehörigen Kommentierung von Peters geht hervor, dass es sich bei der kognitiven Fähigkeit, das Unrecht der Tat aus der sozialen Bindung heraus zu erkennen, um den grundlegenden Fähigkeitsbereich der strafrechtlichen Verantwortlichkeitsreife handelt. Deshalb beschränken sich die folgenden Betrachtungen auf diesen Teilbereich der als strafrechtliche Unrechtserkenntnisfähigkeit bezeichnet wird. Für die Betrachtungsebene der Steuerungsfähigkeit wäre jedoch zu vermuten, dass diese auch mit der zunehmenden strafrechtlichen Unrechtserkenntnisfähigkeit ansteigen könnte.

Ein Modell der empirischen Bestätigung der strafrechtlichen Altersgrenze von 14 Jahren für den Bereich der strafrechtlichen Unrechtserkenntnisfähigkeit wäre, dass vor dem vollendeten vierzehnten Lebensjahr diese Fähigkeit im Sinne der absoluten Schuldunfähigkeitsvermutung des § 19 StGB überwiegend fehlen und danach zumindest mehr und mehr sowie ab der Volljährigkeit bis auf Ausnahmen gänzlich nachweisbar sein würde.

Besondere Anforderungen dahin, wie die zur strafrechtlichen Verantwortlichkeitsreife notwendige kognitive Fähigkeit erkannt und bestimmt werden soll, werden vom Gesetz nicht vorgegeben. Eine Erfassung der strafrechtlichen Unrechtserkenntnisfähigkeit durch sprachliche Äußerungen erscheint allerdings nicht notwendig, da z.B. *de lege lata* Taubstumme oder minderbegabte Jugendliche auch hinsichtlich ihrer strafrechtlichen Verantwortlichkeitsreife nach § 3 JGG zu untersuchen wären. Deshalb müsste auf Verfahren zurückgegriffen werden, die keine Verbalisation erfordern.

Diese non-verbale Verfahren müssten zudem im Sinne einer diagnostischen Hypothesenbildung Aussagen ermöglichen, ob die Nullhypothese „Fehlen der strafrechtlichen Unrechtserkenntnisfähigkeit beim Täter für die von ihm begangene Tat“ verworfen werden kann oder nicht, da die Unrechtserkenntnisfähigkeit täter- und tatspezifisch zu prüfen ist. In der Experimentalpsychologie erfolgt im Regelfall die Prüfung der Nullhypothese durch den Vergleich zweier oder mehrerer empirischen Statistiken (z.B.: t-Test, Varianzanalyse). Wegen der täter- und tatspezifischen Prüfung der strafrechtlichen Unrechtserkenntnisfähigkeit erscheint jedoch das übliche Vorgehen in der diagnostischen Psychologie, die Summierung von Testwerten und der anschließende Vergleich dieser mit einer Normstichprobe, ungeeignet. Vielmehr erfordert die Spezifität der zu beurteilenden Täter und Taten eine Methode, die eine individualstatistische Prüfung der Beibehaltung oder Ablehnung der Nullhypothese erlaubt.

### Moralstufentheorie von Kohlberg

Nach Hommers (1983, 2000) bestehen inhaltliche Parallelen zwischen den ersten vier der insgesamt sechs hierarchischen Stufen der psychologischen Theorie der Entwicklung des moralischen Urteils von Kohlberg (vgl. dazu auch deutschsprachige Darstellungen bei Eckensberger, 1998, Heidbrink, 1991, Montada, 1995, oder die Aufsatzsammlung von Kohlberg, 1995) und den sich aus den juristischen Anforderungen ergebenden rechtlich relevanten Fähigkeitsbereichen zur zivil- und strafrechtlichen Verantwortlichkeit Minderjähriger. Während die zweite Moralentwicklungsstufe Kohlbergs das für die zivilrechtliche Verantwortlichkeit Minderjähriger nach § 828 (2) BGB notwendige Kriterium der allgemeinen Unrechtserkenntnis gegenüber dem Mitmenschen beinhaltet, trifft dies für die Unrechtserkenntnis aus der Sozialbindung bei der vierten Stufe zu. Kohlberg gewann sein Stufenmodell mit Hilfe des "Moral Judgment Interview" (MJ; Colby, Kohlberg, Speicher, Hewer, Candee, Gibbs & Power, 1987), in dessen Verlauf die Probanden ihre Lösungen zu moralischen Dilemmata bezüglich der darin enthaltenen Wertklassen (z.B.: Leben vs. Eigentum) begründen mussten. Probanden der vierten Moralentwicklungsstufe rechtfertigen ihre Entscheidungen in moralischen Dilemmata aus der Verpflichtung gegenüber der gesellschaftlichen Ordnung und ihrer Einhaltung. Bei ihnen dominiert die "social-moral perspective" von "law and order". Ihre Handlungsbegründungen gehen über die Erfüllung von Rollenerwartungen anderer hinaus, wie sie beispielsweise bei Probanden der dritten moralischen

Entwicklungsstufe mit ihrer "good-boy"- Orientierung noch vorherrschen.

Das MJI wurde jedoch nicht unter dem Gesichtspunkt der Diagnose der strafrechtlichen Unrechtserkenntnisfähigkeit Jugendlicher konstruiert. Außerdem würden die mit dem MJI gewonnenen Befunde keine individualstatistische Prüfung der Nullhypothese ermöglichen. Gleiches gilt für bereits existierende non-verbale Verfahren (DIT: Rest, 1974, 1979; DIT2: Rest, Narvaez, Thoma & Bebeau, 1999; MUT: Lind, 1978), die auf der Moralstufentheorie von Kohlberg basieren, da sie auf die Berechnung von Summenscores bzw. Konsistenzscores abzielen, die nur im Gruppenvergleich zu interpretieren sind. Außerdem vernachlässigten diese Verfahren den für die empirische Begründbarkeit der Altersgrenze der strafrechtlichen Verantwortlichkeit nach § 19 StGB zu untersuchenden Alterszeitraum, was nicht zuletzt an der Schwierigkeit der dort verwendeten Aufgaben liegen dürfte. So ist beispielsweise für den DIT ein Leseniveau von 12-Jahren notwendig (Rest, Narvaez, Bebeau & Thoma, 1999).

### Eigener Ansatz

Der eigene Ansatz besteht darin, dass den Probanden, die in die Rolle eines Richters schlüpfen, nach der Schilderung einer Straftat und deren Umstände nur jeweils ein thematisch fixierter Inhalt der ersten vier Kohlbergstufen zur Beurteilung vorgegeben wird. Diese werden jedoch jeweils in ihrer Ausprägung systematisch variiert, im einfachsten Fall dichotom: Groß versus Klein, Ja versus Nein, Dafür versus Dagegen. Die Aufgabe der Probanden ist es, zu entscheiden, wieviel Strafe der Täter für seine Straftat bekommen soll, wenn er einen dieser variierten Inhalte während der Ausführung seiner Tat gedacht hatte. Es handelt sich also um ein experimentelles Vorgehen.

Da pro Kohlbergstufe zwei dichotome Ausprägungen eines fixierten Inhalts vorliegen, müssten diese sich auf die Strafzumessung unterschiedlich auswirken. Die Differenz der Strafzumessung zwischen den dichotomen Variationen eines fixierten thematischen Inhalts würde somit dessen Effekt auf die Strafzumessung widerspiegeln. Derjenige der auf die ersten vier Kohlbergstufen bezogenen fixierten Inhalte, der den größten Effekt (Differenz) bei der Strafzumessung aufweist, wäre dann für die beurteilende Person bei der Strafzumessung am einflussreichsten. Wenn dies für den fixierten Inhalt der vierte Kohlbergstufe zutreffen würde, dann würde dies für das Vorhandensein der strafrechtlichen Unrechtserkenntnisfähigkeit für diesen Straftatbestand bei der beurteilenden Person sprechen.

Allerdings muss für die Ablehnung der Nullhypothese noch auf einen individualstatistischen Maßstab zurückgegriffen werden, um sicherzustellen, dass die von dem jeweiligen Probanden getroffene Differenzierung in der Strafzumessung nicht zufällig ist. Hierzu kann der individuell bestimmte Messfehler herangezogen werden, mit dessen Hilfe die Differenzen statistisch gesichert werden können. Die Bestimmung des individuellen Messfehlers erfolgt über die zweimalige Beurteilung der variierten Gedankeninhalte.

### Fragestellungen

Für die Beurteilung der diagnostischen Brauchbarkeit des Gedanken-Paradigmas ist als erster Anhaltspunkt für die Validität zu prüfen, wie viele Probanden der einzubeziehenden Altersspanne individuell gesicherte Effekte in der Strafzumessung für die variierten Inhalte der ersten vier Kohlbergstufen aufweisen. Die altersabhängige Veränderung der Häufigkeiten der individuell gesicherten Strafzumessungseffekte müsste analog zu der von Kohlberg beobachteten Stufenabfolge verlaufen. Mit zunehmendem Alter müsste die Variation des Inhalts der vierten Kohlbergstufe und damit der Sozialbindung häufiger einen individualdiagnostisch gesicherten Strafzumessungseffekt aufweisen. Gleichzeitig müssten die individualdiagnostisch gesicherten Strafzumessungseffekte der niedrigeren Kohlbergstufen zurückgehen. Außerdem muss überprüft werden, ob die Zu- und Abnahme von Strafzumessungseffekten unabhängig von dem Tatmotiv des Täters erfolgt und somit rein tatspezifischen Charakter besitzt. Diesbezüglich wird das

Gedanken-Paradigma mit zwei Einbruchsdiebstahlfällen mit unterschiedlichem Tatmotiv durchgeführt.

In einem zweiten Schritt kann die Frage der intraindividuellen Relationen der individualdiagnostisch gesicherten Strafzumessungseffekte untersucht werden. Hierbei geht es um die altersabhängigen Veränderungen der individuell stärksten individualdiagnostisch gesicherten Strafzumessungseffekte. Die Überlegung dabei ist, dass der moralische Inhalt mit dem stärksten individualdiagnostisch gesicherten Strafzumessungseffekt bei realen moralischen Entscheidungen die größte Chance hätte, im Sinne der Steuerungsfähigkeit handlungsrelevant zu werden. Diesbezüglich wird geprüft, ob die Zu- und Abnahme der intraindividuellen größten individualdiagnostisch gesicherten Strafzumessungseffekte unabhängig von dem Tatmotiv des Täters erfolgt und somit rein tatspezifischen Charakter besitzt.

Abschließend müsste zur Einordnung der auf die ersten vier Kohlbergstufen bezogenen individualdiagnostisch gesicherten Strafzumessungseffekte bzw. der individuell größten individualdiagnostisch gesicherten Strafzumessungseffekte ein Vergleich zur Entwicklung des Einflusses einer anderen moralischen Norm auf die Strafzumessung untersucht werden. Hierfür bietet sich der Rückgriff auf die Wiedergutmachungsabsicht an. In früheren Untersuchungen (Hommers, 1991) zeigte sich, dass deren Einfluss mit zunehmendem Alter abnahm, obwohl ihre moralische Bedeutung als *prima facie duty* (Ross, 1930) kaum in Frage steht. Daher würde ein größerer Strafzumessungseffekt von anderen moralischen Inhalten deren moralische Wertigkeit belegen, so daß die Effekte von unterschiedlichen Inhaltsvariationen mit der Variation der Wiedergutmachungsabsicht zu prüfen war.

## **Methode**

### **Aufgabe**

Zwei Einbruchsdiebstahlszenarien, die sich hinsichtlich des Tatmotivs unterschieden, wurden unterschiedlichen Probanden vorgegeben. Das erste Szenario mit einem altruistischen Tatmotiv basierte auf der bekannten und von Kohlberg entwickelten Geschichte von Heinz und dem Apotheker:

*Eine Frau liegt im Sterben, sie leidet an einer besonderen Art von Krebs. Ein Medikament könnte ihr aber vielleicht helfen. Es handelt sich um eine Art Radium, das ein Apotheker entdeckt hat. Doch der Apotheker verlangt eine unglaubliche Menge Geld dafür, etwa zehnmal mehr als ihn die Herstellung des Medikaments gekostet hat. Heinz, der Ehemann der kranken Frau, geht zu allen Bekannten, um sich das Geld auszuleihen; aber er kann nur ungefähr die Hälfte der notwendigen Summe zusammenbringen. Er erzählt dem Apotheker, dass seine Frau im Sterben liegt und bittet ihn, das Medikament billiger zu verkaufen. Doch der Apotheker sagt: "Nein, ich kann das Medikament nicht billiger verkaufen. Ich habe viele Jahre daran gearbeitet, ich muss meine Mitarbeiter und meine Schulden bezahlen und will mit dem Medikament auch Geld verdienen". Heinz weiß sich keinen Rat, bricht in die Apotheke ein und stiehlt das Medikament.*

Dem eher ungewöhnlichen, altruistischen Tatmotiv der Heinz-Geschichte wurde im zweiten Szenario ein realistischer Einbruchsdiebstahl mit einem egoistischen Tatmotiv gegenübergestellt:

*Heinz liebt teure Uhren. Eigentlich kann er sich solche teuren Uhren aber gar nicht leisten. Dennoch gelingt es ihm durch eisernes Sparen das Geld für die teure Uhr zusammenzubekommen. An einem Nachmittage steckt Heinz seinen Geldbeutel mit seinem gesamten gesparten Geld in seine Jackentasche und geht zu dem Uhrengeschäft, in dessen Schaufenster er genau die Uhr gesehen hatte, die er haben möchte. Als Heinz vor dem Schaufenster des Uhrengeschäfts steht, greift er in seine Jackentasche, um seinen Geldbeutel*

*zu holen. Aber was ist das? In seiner Jackentasche ist ein riesiges Loch. Der Geldbeutel mit seinem gesamten gesparten Geld ist verschwunden. Wahrscheinlich hat Heinz seinen Geldbeutel auf dem Weg zum Uhrengeschäft verloren. Heinz sucht sofort auf dem ganzen Weg nach seinem Geldbeutel. Aber er kann seinen Geldbeutel nicht mehr finden. Sein Geldbeutel und damit sein gesamtes gespartes Geld bleiben für immer verschwunden. In der Nacht kommt Heinz zu dem Uhrengeschäft zurück. Er schlägt mit einem großen Stein die Schaufensterscheibe des Uhrengeschäfts ein und stiehlt die Uhr, die er haben wollte.*

Beide Szenarien wurde als "Richter-Spiel" eingeführt. Nach der Vorgabe dieses Null-Stimulus verhängten die Probanden auf einer 13-stufigen Strafe-Skala, die in Abstufungen von 0 bis 12 Monaten reichte, eine Gefängnisstrafe. Anschließend an das Urteil über den Null-Stimulus wurden mehrere mögliche Gedanken von Heinz, die er während seines Einbruchs gehabt haben könnte, vorgestellt. Die Wortlaute der vier auf die ursprünglichen Kohlbergstufen bezogenen Gedanken wurden in zwei Abstufungen variiert und lauteten im einzelnen (abweichende Formulierungen im egoistischen Einbruchsdiebstahl werden in Klammern dargestellt):

a) passend zu Stufe I, Gehorsams- bzw. Strafvermeidungs-Orientierung: *"Die Gefahr entdeckt und bestraft zu werden, ist klein." bzw. "... ist groß."*

b) passend zu Stufe II, Orientierung des instrumentellen Austausches: *"Meine Frau findet es bestimmt toll, dass ich in die Apotheke einbreche und das Medikament (... in den Uhrenladen einbreche und die Uhr ...) stehle." bzw. "... ist bestimmt dagegen, dass ... ."*

c) passend zur Stufe III "good-boy"-Orientierung: *"Mein Freund würde für seine Frau in so einem Fall sicher auch das Medikament stehlen (Mein Freund würde in so einem Fall die Uhr sicher auch stehlen)." bzw. "... nicht das Medikament stehlen (... sicher nicht stehlen)."*

d) passend zu Stufe IV, "law-and-order"-Orientierung: *"Wenn jeder so handeln würde, dann bleiben Recht und Ordnung immer erhalten." bzw. "... , geht es in der Gesellschaft bald drunter und drüber."*

Darüber hinaus wurde ein Gedankenpaar über die Wiedergutmachungsabsicht von Heinz eingesetzt: *"Ich werde später dem Apotheker für die Medizin und die Fensterscheibe (... dem Uhrenhändler für die Uhr und die Fensterscheibe ...) Geld mit der Post schicken." bzw. "... kein Geld mit der Post schicken."* Da der Gedanke über die Wiedergutmachungsabsicht nicht in den Kohlbergstufen erscheint, ließ sich damit der Stellenwert der auf Kohlbergstufen bezogenen Gedanken beurteilen.

## Durchführung

An der Untersuchung nahmen 65 8-jährige Zweitklässler, 65 10-jährige Viertklässler, 70 12-jährige Sechstklässler und 63 14-jährige Achtklässler teil, wobei zum Testzeitpunkt die Zweit- und Viertklässler bzw. die Sechst- und Achtklässler jeweils dieselbe Grund- bzw. Hauptschule besuchten. Die Probanden wurden im Klassenverband untersucht. Pro Klassenstufe bearbeiteten insgesamt 50, 44, 47 bzw. 40 Probanden den altruistischen und 15, 21, 23 bzw. 23 Probanden den egoistischen Einbruchsdiebstahl. Das Geschlecht war jeweils annähernd gleichverteilt und es zeigten sich keine Unterschiede im Ergebnis hinsichtlich dieser Variablen.

Beide Szenarien wurde als "Richter-Spiel" eingeführt. Die Szenarien wurden jeweils zweimal vorgetragen und anschließend von den Kindern mehrmals wiedergegeben. Nach Beurteilung des Nullstimulus wurden in einem ersten Durchgang den Probanden die zehn variierten Gedanken jeweils kontrastierend aufeinander und zwar in der Reihenfolge Entdeckungsfahr, Meinung der Frau, Vermutung über den Freund, Ordnungsgefahr und Wiedergutmachung zur Beurteilung vorgegeben. Die Vorgabe der einzelnen Gedanken-Stimuli erfolgte immer zweimal hintereinander verbal ohne Wiederholung durch die Kinder. Für die Beurteilung jedes einzelnen Gedankens stand eine separate Antwortseite im Fragebogen zur Verfügung. Es wurde

stichprobenartig kontrolliert, ob die Urteile auf die richtige Seite des Antwortbogens abgegeben wurden.

Nach der erstmaligen Vorgabe der zehn einzelnen Gedanken folgte eine Zwischenphase, in der die Probanden Kombinationen aus zwei einzelnen Gedanken zu beurteilen hatten. Dieses Vorgehen diente zur Prüfung der Stabilität der Strafzumessungsschemata der Probanden. Über die diesbezüglichen Ergebnisse wird hier nicht berichtet. Abschließend wurden nochmals die zehn einzelnen Gedanken jedoch in einer gemischten Reihenfolge vorgegeben.

## Auswertung

Die individuelle Stärke der Beachtung eines fixierten moralischen Inhalts bei der Strafzumessung drückt sich im Gedanken-Paradigma durch die Höhe der Differenz, d.h. dem Effekt, der variierten Gedanken eines fixierten moralischen Inhalts aus. Durch die zweimalige Beurteilung jedes Einzelgedanken am Anfang und am Ende ließ sich der individuelle Messfehler eines Probanden aus den 10 zweimaligen Beurteilungen bestimmen. Damit ließen sich dann die Strafzumessungseffekte eines Probanden relativieren, so dass bei der Bestimmung der individuellen Strafzumessungseffekte die Ausprägung der zufälligen Urteilsvariation eines Probanden berücksichtigt werden konnte. Ein durch diese Relativierung entstandener kritischer Bruch wurde mit 10 Freiheitsgraden unter der t-Verteilung beurteilt. Dies erfolgte unter zweiseitiger Fragestellung mit 10%-Irrtumsniveau gegen  $t=1.812$ . Die Aussage, ob ein Proband einen fixierten moralischen Inhalt signifikant bei der Strafzumessung beachtet hatte, wurde ergänzt durch die Feststellung, welcher davon der individuell größte Strafzumessungseffekt war.

## Ergebnisse

Die Ergebnisse betreffen die mittleren Strafhöhen für Heinz in den beiden Szenarien, wenn sie ohne die Berücksichtigung weiterer moralischer Informationen beurteilt wurden (Tabelle 1). Weiterhin werden die Wirkungen der Vorgabe der zehn Einzelgedanken (Tabelle 2) angesichts der individuell auf dem 10%-Niveau gesicherten Strafzumessungseffekte der fixierten moralischen Inhalte untersucht. Schließlich werden die individuell größten auf dem 10%-Niveau gesicherten Strafzumessungseffekte dargestellt (Tabelle 3). Da zwischen erster und zweiter Vorgabe nur geringe Unterschiede bestanden, wurden die Auswertungen zur Wirkung von Einzelgedanken durch Mittelung über die Messwiederholung vorgenommen.

Tabelle 1: Mittlere Bestrafung in Monaten je Einbruchdiebstahlszenario ohne die Berücksichtigung weiterer moralischer Inhalte

Altersgruppe:	Altruistisches Tatmotiv				Egoistisches Tatmotiv			
	8	10	12	14	8	10	12	14
N =	50	44	47	40	15	21	23	23
Mittlere Bestrafung in Monaten (Standardabweichung)	5,7 (3,7)	3,7 (3,2)	3,4 (2,9)	2,5 (2,9)	8,9 (2,9)	8,3 (2,0)	6,1 (3,1)	6,2 (3,8)

Aus der Tabelle 1 geht hervor, dass ältere Probanden in beiden Szenarien weniger bestraft wurden. Dabei war generell die Bestrafung beim egoistischen Einbruch höher als beim altruistischen. Hervorzuheben ist, dass bereits 8-Jährige fähig waren, den moralischen Unterschied zwischen einem altruistischen und einem egoistischen Tatmotiv bei der Strafzumessung zu berücksichtigen.

Tabelle 2: Prozentuale Häufigkeiten von individuell gesicherten Strafzumessungseffekten von Gedankeninhalten

Altersgruppe:	Altruistisches Tatmotiv				Egoistisches Tatmotiv			
	8	10	12	14	8	10	12	14
N =	50	44	47	40	15	21	23	23



Maximaler Gedankeninhalt:								
Entdeckungsgefahr	16	32	36	28	27	38	30	35
Meinung der Frau	56	55	60	55	47	57	57	22
Handeln des Freundes	46	36	17	18	53	71	65	35
Gefahr für die Ordnung	52	61	72	80	27	67	65	52
Wiedergutmachung	92	95	94	88	87	95	78	65
Mittlerer Meßfehler für ein Urteil	2,0	1,7	1,9	1,8	2,0	1,4	1,6	1,4

Gemäß Tabelle 2 ergab sich hinsichtlich der Strafzumessung ein Anstieg der individuell gesicherten Effekte des Inhalts über die Ordnungsgefährdung (Kohlbergstufe IV) bei beiden Szenarien. Ausgehend von den 8-Jährigen stieg bis zum Alter von 14 Jahren die Häufigkeit der individuell gesicherten Strafzumessungseffekte der Ordnungsgefährdung bei dem altruistischen Einbruch von 52% um etwas weniger als das anderthalbfache auf 76%. Bei dem egoistischen Einbruch war ein etwas stärkerer Anstieg allerdings auf niedrigerem Niveau von 27% bei den 8-Jährigen auf 52% bei den 14-Jährigen zu verzeichnen. Bei dem altruistischen Einbruch wurde zudem ab einem Alter von 10 Jahren der Effekt der Ordnungsgefährdung bei der Strafzumessung im Vergleich zu den übrigen Kohlbergstufeninhalten individuell am häufigsten gesichert. Vorher wurde bei den 8-Jährigen der Effekt der Information über die Meinung der Frau bei der Strafzumessung am häufigsten individuell gesichert. Bei dem egoistischen Einbruch bis einschließlich zum Alter von 10 Jahren wurde der Effekt der Information über das Handeln des Freundes am häufigsten unter den übrigen Kohlbergstufeninhalten bei der Strafzumessung individuell gesichert. Erst bei den 12-Jährigen zog die Häufigkeit der individuell gesicherten Effekte der Ordnungsgefährdung mit denen der Information über das Handeln des Freundes bei der Strafzumessung gleich und übertraf diese schließlich bei den 14-Jährigen. Somit ergaben sich Übereinstimmungen zwischen dem Entwicklungsverlauf der Strafzumessungseffekte und dem hierarchischen Stufenmodell der Moralentwicklung von Kohlberg. Eine Ausnahme bildet der Effekt der Information über die Entdeckungsgefahr, bei der sowohl im altruistischen als auch im egoistischen Szenario kein Entwicklungsverlauf bei der Strafzumessung zu beobachten war. Vielmehr lag die Häufigkeit der individuell gesicherten Effekte dieser Information bei allen Altersgruppen auf relativ niedrigem Niveau.

Im Vergleich zu den sich auf die ersten vier Kohlbergstufen beziehenden Informationen wurde jedoch sowohl im altruistischen als auch im egoistischen Einbruchsdiebstahlszenario der Effekt der Wiedergutmachungsabsicht bei der Strafzumessung am häufigsten individuell gesichert. Während im egoistischen Einbruchsdiebstahlszenario die Häufigkeit der individuell gesicherten Strafzumessungseffekte der Wiedergutmachungsabsicht mit zunehmenden Alter von 87% bei den 8-Jährigen auf 65% bei den 14-Jährigen abnahm, waren im altruistischen Einbruch die Häufigkeiten annähernd über alle Altersgruppen gleichverteilt.

Tabelle 3: Prozentuale Häufigkeiten der größten individuell gesicherten Strafzumessungseffekte von Gedankeninhalten

Altersgruppe:	Altruistisches Tatmotiv				Egoistisches Tatmotiv			
	8	10	12	14	8	10	12	14
N =	50	44	47	40	15	21	23	23
Maximaler Gedankeninhalt:								
Entdeckungsgefahr	0	2	0	0	0	0	0	4
Meinung der Frau	4	2	4	8	0	5	9	0
Handeln des Freundes	8	5	0	0	20	5	9	4
Gefahr für die Ordnung	10	11	36	43	0	14	17	44
Wiedergutmachung	64	75	53	33	53	67	44	22
Mehrere Maxima	8	5	2	10	13	10	13	9
Keine Signifikanz	6	0	4	8	13	0	9	17
$\Sigma =$	100	100	99	102	99	101	101	100

Bei Auswertung nach dem größten individuell gesicherten Strafzumessungseffekt pro Proband traten nach Tabelle 3 sowohl im altruistischen als auch im egoistischen Szenario zwei entgegengesetzte Alterstrends auf. Während mit zunehmendem Alter die Häufigkeiten der Probanden mit einem größten Strafzumessungseffekt bei der Wiedergutmachungsabsicht abnahm, stieg die Häufigkeiten der Probanden mit einem größten Strafzumessungseffekt bei der Ordnungsgefährdung. Letztere Probanden bildeten ab einem Alter von 14 Jahren sogar die Mehrzahl. Allerdings fiel der Anstieg der Häufigkeit der Probanden mit einem größten Strafzumessungseffekt bei der Ordnungsgefährdung in beiden Szenarien unterschiedlich aus. Im altruistischen Fall war diesbezüglich der größte Anstieg zwischen den 10- und den 12-Jährigen zu beobachten (von 11% auf 36%). Im egoistischen Einbruch trat ein vergleichbarer Anstieg erst zwei Jahre später zwischen den 12- und 14-Jährigen Probanden auf (von 17% auf 44%). Die mit zunehmenden Alter beobachtete Abnahme der Beachtung und der größten Beachtung der Wiedergutmachungsabsicht bei der Strafzumessung wurde bereits in früheren Untersuchungen mit einer Brandstiftungsthematik (vgl. Hommers, 1991) vorgefunden. Im Fall des egoistischen Einbruchdiebstahls ist hervorzuheben, dass 20% der 8-Jährigen den größten Strafzumessungseffekt bei der Information über das Handeln des Freundes aufwiesen und sonst nur noch die Wiedergutmachungsabsicht, wenn auch am häufigsten, den größten individuell gesicherten Effekt hatte. Da dieser Befund nur beim egoistischen Einbruch auftrat, scheint bei einem realitätsnäheren Straftatgeschehen zumindest bei 8-jährigen Kindern die peer group einen im Vergleich zur Ordnungsgefährdung größeren Einfluss zu besitzen.

## Diskussion

Ein neuer experimenteller Ansatz zur Bestimmung der Wirkung von moralischen Informationen, das Gedanken-Paradigma (Hommers, 1997), wurde vorgestellt. Beim Gedanken-Paradigma handelte es sich um ein psychometrisches einzelfalldiagnostisches Verfahren. Die Fragestellungen bezogen sich auf die empirische Fundierung der Altersgrenze der strafrechtlichen Verantwortlichkeit nach § 19 Strafgesetzbuch (StGB) von 14 Jahren. Die Betrachtungsebene beschränkte sich auf die Entwicklung der strafrechtlichen Unrechtserkenntnisfähigkeit und damit auf den kognitiven Teilaspekt der strafrechtlichen Verantwortlichkeitsreife Jugendlicher nach § 3 Jugendgerichtsgesetz (JGG). Zur Operationalisierung der strafrechtlichen Unrechtserkenntnisfähigkeit wurden moralische Inhalte der ersten vier Stufen der Moralentwicklungstheorie Kohlbergs aufgegriffen, da nach Hommers (1983, 2000) zwischen der vierten Kohlbergstufe und dem von Peters (1967) in seiner Kommentierung zum § 3 JGG geprägten Begriff der Sozialreife inhaltliche Übereinstimmungen bestehen.

Ein Vorteil des Gedanken-Paradigmas war, dass im Vergleich zu bisherigen auf der Kohlbergstufentheorie basierenden Verfahren jüngere Altersgruppen untersucht werden konnten.

Zudem besaß das psychometrische Vorgehen im Gedanken-Paradigma den Vorteil, dass im Sinne einer diagnostischen Hypothesenbildung individuell gesicherte Aussagen darüber getroffen werden können, ob die Nullhypothese „Fehlen der strafrechtlichen Unrechtserkenntnisfähigkeit beim Täter für die von ihm begangene Tat“ verworfen werden kann oder nicht. Damit versuchte das Gedanken-Paradigma Wege für eine maßgeschneiderte Diagnostik im Bereich der Diagnose der strafrechtlichen Verantwortungsreife Jugendlicher aufzuzeigen, die tat- und täterspezifisch erfolgen müsste.

Die vorgestellten Befunde sprachen für die Validität des diagnostischen Ansatzes des Gedanken-Paradigmas, da der Entwicklungsverlauf der individuell gesicherten Strafzumessungseffekte von Informationen über einen Täter, die sich auf die ersten vier Kohlbergstufen bezogen, im Wesentlichen mit dem von Kohlberg postulierten Entwicklungsverlauf der einzelnen Stufen übereinstimmte, wenn auch hier bei erheblich jüngeren Altersgruppen. Jedoch wird mit dem Gedanken-Paradigma nicht der Anspruch erhoben, das Verständnis von oder die Urteilsfähigkeit auf bestimmten Stufen im Sinne Kohlbergs zu erfassen. Es geht im Gedanken-Paradigma allein um die Fähigkeit, bei der Strafzumessung bestimmte moralische Inhalte, die Aussagen über die strafrechtliche Unrechtserkenntnisfähigkeit des Beurteilers erlauben, zu berücksichtigen.

Bezüglich der empirischen Fundierung der Altersgrenze der strafrechtlichen Verantwortlichkeit zeigte sich, dass bereits die Mehrheit der 10-Jährigen unabhängig vom Tatmotiv einen individuell gesicherten Strafzumessungseffekt bei der Information über die Ordnungsgefährdung aufwies (Tabelle 2). Bis zum Alter von 12-Jahren besaß jedoch bei den meisten Probanden die Information über die Wiedergutmachungsabsicht den größten individuell gesicherten Effekt auf die Strafzumessung. Erst im Alter von 14-Jahren kehrte sich das Verhältnis um, so dass bei der Mehrzahl der Jugendlichen nun die Information über die Ordnungsgefährdung den größten individuellen Strafzumessungseffekt hatte (Tabelle 3). Diese Befunde stützten die Altersgrenze der strafrechtlichen Verantwortlichkeit empirisch. Denn selbst wenn noch nicht bei allen 14-Jährigen die strafrechtliche Unrechtserkenntnisfähigkeit, wie in dem vorliegenden Fall (Einbruchsdiebstahl), nach psychometrischen Kriterien vorliegt, würde § 3 JGG diesem Umstand Rechnung tragen.

Der in den erlangten Befunden enthaltene Validitätsnachweis ermutigt dazu, das Gedanken-Paradigma, neben der hier verfolgten Intention *de lege ferenda*, auch *de lege lata* einzusetzen, da dessen experimenteller und psychometrischer Aufbau eine maßgeschneiderte täter- und tatspezifische Diagnostik ermöglicht.

## Literatur

- Colby, A., Kohlberg, L., Speicher, B., Hewer, A., Candee, D., Gibbs, J., & Power, C. (1987). *The measurement of moral judgment*. (Vol. I, Vol. II). New York: Cambridge University Press.
- Eckensberger, L. H. (1998). Die Entwicklung des moralischen Urteils. In H. Keller (Hrsg.), *Lehrbuch Entwicklungspsychologie* (S. 475-516). Bern: Huber.
- Heidbrink, H. (1991). *Stufen der Moral: Zur Gültigkeit der kognitiven Entwicklungstheorie Lawrence Kohlbergs*. München: Quintessenz.
- Hommers, W. (1983). *Die Entwicklungspsychologie der Delikts- und Geschäftsfähigkeit*. Göttingen: Hogrefe.
- Hommers, W. (1991). Das „Zündeln“ im Urteil: Alterstrends und psychometrische Diagnostizierbarkeit der zivilrechtlichen Verantwortlichkeit nach § 828 BGB. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 12, 163-175.
- Hommers, W. (1997). Integration of Kohlbergian information in punishment. *European Journal of Applied Psychology* (Revue Européenne de Psychologie Appliquée), 47, 31-37.
- Hommers, W. (2000). Begründen Kultur und Natur die Altersgrenze der Strafmündigkeit? Eine vorläufige Antwort mit dem Gedanken-Scenario. In M. Usteri (Hrsg.), *Gene, Kultur und Recht. Schriften zur Rechtspsychologie*, Band 5 (S. 91-112). Bern: Stämpfli.

- Kohlberg, L. (1995). *Die Psychologie der Moralentwicklung*. Frankfurt a. M.: Suhrkamp.
- Lind, G. (1978). Wie mißt man das moralische Urteil? Probleme und Möglichkeiten der Messung eines komplexen Konstrukts. In G. Portele (Hrsg.), *Sozialisation und Moral* (S. 171-201). Weinheim: Beltz.
- Montada, L. (1995). Moralische Entwicklung und moralische Sozialisation. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie*. Ein Lehrbuch (3. Aufl., S. 862- 894). Weinheim: Beltz und PVU.
- Rest, J. (1974). *Manual for the Defining Issue Test: An objective Test for moral judgment development*. Minnesota: University of Minnesota.
- Rest, J. (1979). *Revised manual for the Defining Issue Test: An objective Test for moral judgment development*. Minneapolis: Minnesota Moral Research Projects.
- Rest, J., Narvaez, D., Thoma, S. J., & Bebeau, M. J. (1999). DIT2: Devising and testing a revised instrument of moral judgment. *Journal of Educational Psychology*, *91*, 644-659.
- Rest, J., Narvaez, D., Bebeau, M. J., & Thoma, S. J. (1999b). *Postconventional Moral Thinking. A Neo-Kohlbergian Approach*. Mahwah, NJ: Erlbaum.
- Ross, W. D. (1930). *The right and the good*. London: Oxford University Press.

# Soziale Unterstützung im Strafvollzug: Der Einfluss sozialer Beziehungen auf das Befinden und die Normorientierung junger Männer

*Daniela Hosser*

*Kriminologisches Forschungsinstitut Niedersachsen (KFN)*

Die Jugendstrafe soll die Verurteilten dazu befähigen, künftig ein Leben in sozialer Verantwortung führen zu können. Im Vollzug sollen den Jugendlichen einerseits die Konsequenzen ihres Handelns aufgezeigt andererseits aber auch Unterstützungsressourcen und Handlungsalternativen vermittelt werden, um künftigen delinquenten Handlungen vorzubeugen und deviante Einstellungen abzubauen. Die relativ hohen Rückfallraten lassen jedoch gewisse Zweifel am Erfolg der Resozialisierungsbemühungen aufkommen. Zudem deuten empirische Studien auf ein erhöhtes Stresserleben und Anpassungsprobleme der Inhaftierten im Strafvollzug hin (Greve & Hosser, 1998). Beide Faktoren, erhöhte Rückfallraten und vermehrte psychische Belastungen, lassen sich in einen Zusammenhang stellen.

Vermutlich dürften langfristige und stabile Einstellungs- und Verhaltensänderung der Inhaftierten vor allem dann zu erreichen sein, wenn sich bei den Verurteilten im Verlauf der Haft die Überzeugung einstellt, dass eine Hinwendung zur Gesellschaft und ihren Regeln für sie von Vorteil sein kann. Die Inhaftierten müssen eine Veränderung ihrer devianten Einstellungen dabei nicht nur selbst anstreben, sondern auch an einen Erfolg dieser Bemühungen glauben. Sie müssen sich über die **Unterstützungsentention** und die **potentielle Unterstützungsfunktion** ihrer Umwelt klar werden. Einstellungs- und Verhaltensveränderungen setzen dementsprechend ein entwicklungsförderliches soziales Klima und Umfeld voraus, wobei eine stabile, ausgeglichene psychische Befindlichkeit eine Voraussetzung für dauerhafte Veränderungs- und Lernbereitschaft sowie Lernerfolge bietet. Die vorgegebenen Strukturen des Vollzuges, die mit dem Gefängnis verbundenen Feindbilder, das Dilemma zwischen Straf- und Erziehungsfunktion der Jugendstrafe stehen der Bereitstellung und Schaffung von Unterstützungsressourcen jedoch im Wege (Haney, Banks & Zimbardo, 1973).

Vor dem Hintergrund dieser Überlegungen erscheinen Art und Ausmaß der sozialen Unterstützung der Inhaftierten im Vollzug in mehrfacher Hinsicht bedeutsam. Sie dürfte einen **direkten** Einfluss auf die sozialen Einstellungen der Inhaftierten nehmen, wobei die Richtung dieses Einflusses zunächst offen ist. Sie wird von der Art und Quelle der Unterstützung sowie den Unterstützungserwartungen und -bewertungen abhängen. Sie kann zudem aber auch **indirekt** durch die Reduzierung der mit der Haft verbundenen psychischen Belastungen zu einem entwicklungsförderlicheren Anstaltsklima und einer daraus resultierenden größeren Chance auf positive Einstellungsveränderungen beitragen (vgl. Hosser, 2001a). Soziale Unterstützung im Vollzug lässt sich in diesem Sinne als ein Schlüsselfaktor ansehen, der einerseits zur Vermeidung negativer Haftfolgen beizutragen und die Tür für eine künftige erfolgreiche Integration in die Gesellschaft zu öffnen vermag, andererseits aber auch im ungünstigen Falle, z. B. wenn enge soziale Bindungen einseitig nur zu Mitinsassen bestehen, die Verstrickung in deviante Lebenswelten und Subkulturen im Vollzug erschließen bzw. fördern kann und damit einer im Sinne des Vollzugszieles erfolgreichen Wiedereingliederung in die Gesellschaft im Wege steht.

Empirische Befunde zur Rolle sozialer Unterstützung und zur Bedeutung von Beziehungsqualitäten im Vollzug liegen bislang nur vereinzelt vor. Sowohl theoretische

Ansätze wie die soziale Kontrolltheorie (Gottfredson & Hirschi, 1990; Hirschi 1969; Sampson & Laub, 1993) als auch empirische Studien (Quinton, Pickles, Maughan & Rutter, 1993; Werner, 1993) sprechen aber dafür, dass kontinuierliche, stabile und möglichst vielfältige soziale Bindungen dazu beitragen können, delinquentes Handeln zu reduzieren und Entwicklungs- und Bewältigungsdefizite zu kompensieren. Eine aktuelle Untersuchung von Biggam und Power (1997) aus dem schottischen Jugendstrafvollzug weist darauf hin, dass eine als befriedigend erlebte soziale Unterstützung sowie positive Beziehungen zu den Anstaltsmitarbeitern psychischen Störungen bei den Inhaftierten vorbeugen können. Nach den Befunden von Lieblich (1992) kommt den Anstaltsmitarbeitern insbesondere im Hinblick auf die Vermeidung depressiver Symptomatiken eine wichtige Rolle zu. Wright (1993) konnte feststellen, dass positive Beziehungen zum Anstaltspersonal mit einer geringeren Anzahl von Krankheitsfällen und Disziplinarverstößen bei den Inhaftierten in Verbindung stehen. Dahle und Steller (1990) zeigten, dass ein positives soziales Anstaltsklima (im Sinne einer höheren wahrgenommenen Behandlungsorientierung) mit geringerer Aggression und Feindseligkeit bei den Inhaftierten verbunden war.

## **Methode**

Verschiedene Formen sozialer Unterstützung (Netzwerkkontakte, Unterstützungserhalt, wahrgenommene Unterstützung) werden in ihrer Funktion als *protektive Faktoren* des psychischen Befindens (Selbstwert, Depressivität) und hinsichtlich ihres Einflusses auf die Normorientierung der Inhaftierten untersucht. Abgesehen von den o.g. generellen Unterstützungsaspekten werden als situationsspezifische Komponenten zusätzlich die Qualität der Beziehungen zu den Mitinsassen und zum Anstaltspersonal sowie das soziale Anstaltsklima berücksichtigt. Die Erhebung erfolgte einheitlich anhand standardisierter vierstufiger Skalen, lediglich die Netzwerkkontakte wurden fünfstufig erfasst (eine ausführliche Beschreibung der Erhebungsinstrumente findet sich bei Hosser & Greve, 1999).

Die Untersuchung selbst ist Teil des umfangreichen Längsschnittprojektes "Gefängnis und die Folgen", das seit 1997 vom kriminologischen Forschungsinstitut Niedersachsen durchgeführt wird (Greve, Hosser & Pfeiffer, 1997). Die hier vorgestellten Daten beruhen auf einer Querschnitterhebung, die im Zeitraum von April bis August 1998 im Rahmen des Projektes durchgeführt wurde. Die Datenerhebung fand zeitgleich in fünf Jugendanstalten in Norddeutschland statt. Alle zu einem Stichtag (01.04.1998) inhaftierten Personen wurden zur Teilnahme an der Studie gebeten. Mit den Freiwilligen wurden dann persönlich-mündliche Interviews durchgeführt, die ca. 1 1/2 Stunden dauerten und mit 20 DM vergütet wurden. Kriterien für die Teilnahme waren deutsche Staatsangehörigkeit, Alter zwischen 14 und 24 Jahren sowie eine erstmalige Verurteilung zu Straftat.

Insgesamt nahmen 299 Inhaftierte an der Untersuchung teil; die Verweigerungsquote lag bei 31%. Das mittlere Durchschnittsalter der Befragten ist mit 20,9 Jahren (SD: 2,11; Range: 16-28) sehr hoch. Jugendliche bis 17 Jahren sind nur zu 2,7 % in der Stichprobe vertreten, das Gros stellen Heranwachsende im Alter von 18-21 Jahren. Die Inhaftierungszeit zum Interviewzeitpunkt beträgt im Mittel 13,5 Monate (SD: 9,5). Die Dauer des Einweisungsurteils liegt bei durchschnittlich 2,8 Jahren (SD: 1,7). Sehr hoch ist mit 63 % der Anteil der Gewalttäter in der Stichprobe. Für die folgenden Auswertungen wurden 288 Interviews als Datenbasis genutzt. Die Abbruchquote der Interviews lag bei 1,4 %, aufgrund multivariat extrem abweichender Werte wurden 2,4 % der Stichprobe ausgeschlossen.

## Ergebnisse

Auf deskriptiver Ebene ist erkennbar, dass Netzwerkkontakte sehr selten sind und der Unterstützungserhalt von den Inhaftierten als gering eingeschätzt wird (siehe Tabellen 1, 2). Im Gegensatz dazu ist die wahrgenommene Unterstützung als subjektive Einschätzung eher hoch ausgeprägt, was in Einklang mit der hier zugrundegelegten Konzeption der Variable als generalisiertes und damit weitgehend situationsunabhängiges Maß steht. In Bezug auf die Depressivität weisen die Inhaftierten im Vergleich zur Allgemeinbevölkerung signifikant höhere Werte auf. Anhand des aus der Normstichprobe abgeleiteten kritischen Cut-Offs lässt sich ein Anteil von 24 % der Inhaftierten als depressiv auffällig bezeichnen; während der Anteil in der männlichen Normalbevölkerung mit 14 % angegeben wird (Hautzinger & Bailer, 1993). Die hier vorgefundenen Depressivitätswerte stimmen mit denen von Biggam und Power (1997) aus dem schottischen Jugendstrafvollzug überein. Auffallend ist der im Gegensatz zur Depressivität hohe Selbstwert der Inhaftierten, der auf die Rolle von Bewältigungsprozessen verweist.

Tabelle 1: Statistische Kennwerte der Variablen

N = 287	M	SD	Alpha
Depression [ADS-K: Hautzinger & Bailer, 1993]	1.88	.54	.85
Selbstwert [Rosenberg, 1965]	3.14	.50	.81
Normorientierung (Gesamt)* [PF: Ortmann, 1987; PFI: Seitz, 1983]	2.54	.52	.86
Konformismus (F1)	2.64	.66	.75
Legalismus (F2)	2.39	.67	.83
Normakzeptanz (F3)	2.54	.57	.72
Wahrgenommene Unterstützung [Eigenentwicklung]	2.89	.61	.91
Unterstützungserhalt (Gesamt)* [angelehnt an Bliesener, 1988]	.91	.65	-
Emotionale Unterstützung (F1)	1.05	.84	-
Praktische Unterstützung (F2)	.83	.73	-
Informative Unterstützung (F3)	.83	.73	-
Netzwerkkontakte (Gesamt)* [angelehnt an Filipp, 1990]	.83	.40	-
persönliche Kontakte (F1)	.71	.54	-
telefonische Kontakte (F2)	.97	.65	-
briefliche Kontakte (F3)	.80	.54	-
Emotionale Bindung an die Mitinsassen [PF: Ortmann, 1987]	2.93	.64	.74
Erwartung von Feindseligkeit seitens des Vollzugspersonals [PFI: Seitz, 1983]	2.72	.64	.91
Behandlungs- und Rehaorientierung [CIES: Dahle & Steller, 1990]	2.65	.50	.68

Anmerkungen: \* Als Gesamtwert wurde das arithmetische Mittel über die drei Subskalen berechnet.

Tabelle 2: Bivariate Korrelationen der Untersuchungsvariablen

N=287	Selbstwert	Depressivität	Normorient.
Selbstwert	1.00		
Depressivität	-.53***	1.00	
Normorientierung	.05	-.12*	1.00
Netzwerkkontakte	.18**	-.08	.05
Unterstützungserhalt	-.08	.11	-.03
Wahrg. Unterstützung	.25***	-.14*	-.12*
Bindung zu den Mitinsassen	.26***	-.30***	-.19***
Feindseligkeit seitens des Personals	-.10	.22***	-.49***
Rehaorientierung	.07	-.06	.35***

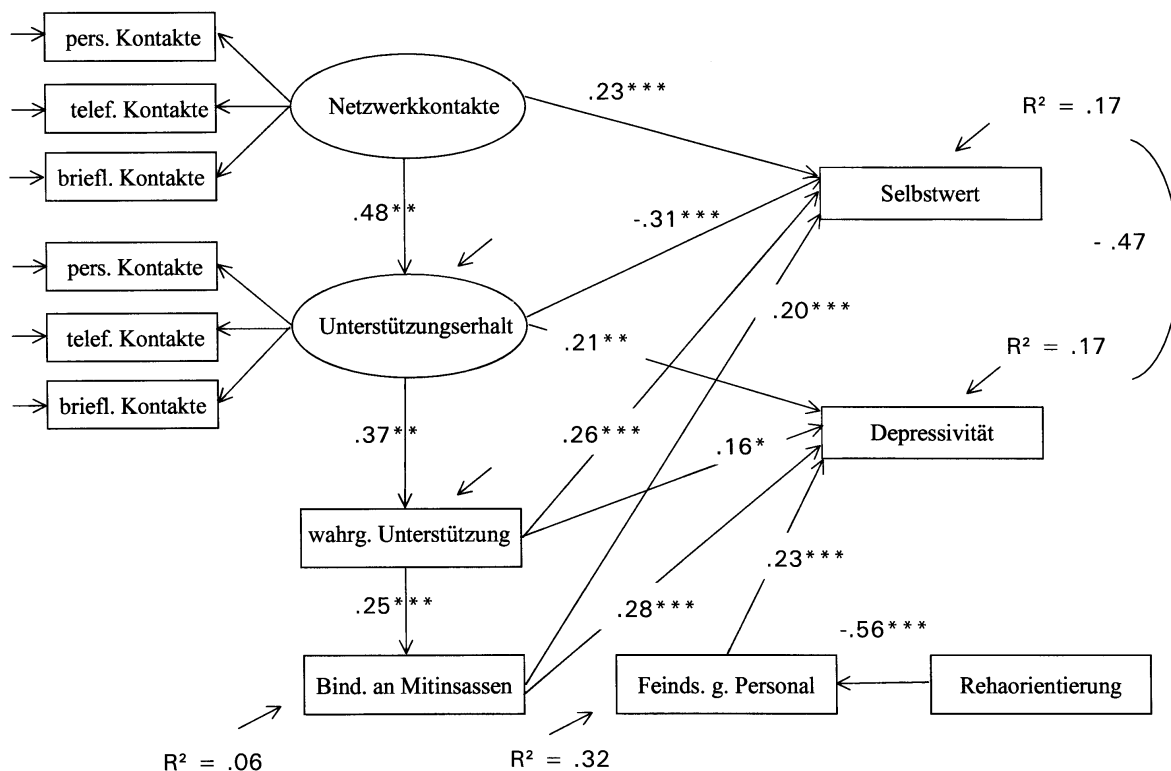
Anmerkungen: \* =  $p < .05$ ; \*\* =  $p < .01$ ; \*\*\* =  $p < .001$

Zur Prüfung der Effekte sozialer Unterstützung auf die psychische Befindlichkeit wird ein Strukturgleichungsmodell formuliert. Die Netzwerkkontakte und die Häufigkeit des Unterstützungserhalts werden als latente Faktoren spezifiziert. Die Gesamtheit der Netzwerkkontakte setzt sich aus der Häufigkeit persönlicher, telefonischer und brieflicher Kontakte der Inhaftierten zusammen. Die Häufigkeit des Unterstützungserhalt wurde jeweils hinsichtlich praktischer, emotionaler und informativer Unterstützungsinhalte erfasst. Als Vorannahmen wurden in dem Modell direkte Pfade zwischen allen Unterstützungsmaßen und den beiden Befindlichkeitsindikatoren angenommen. Ferner wurde davon ausgegangen, dass die Netzwerkkontakte für den Erhalt der Unterstützung maßgeblich sind, während der Unterstützungserhalt zur Wahrnehmung der Unterstützung beiträgt. Zwischen den beiden Befindlichkeitsvariablen wurde die Korrelation der Fehlerterme zugelassen.

Das empirische Modell bestätigt die Vorannahmen weitgehend. Der Modellfit kann als gut bezeichnet werden ( $\chi^2(21)=31.958$ ,  $p > .05$ ; CFI=.980; AGFI=.950; RMSEA=.043; CO(90)=.000-.071), das Ausmaß der Varianzaufklärung beträgt 15 % für den Selbstwert und 6 % für Depressivität. Bemerkenswert ist, dass ein häufiger Unterstützungserhalt im Vollzug mit einem niedrigen Selbstwert ( $r=-.33$ ,  $p < .001$ ) und höherer Depressivität ( $r=.26$ ,  $p < .01$ ) einhergeht (vgl. auch Aymanns, 1992). Die wahrgenommene Unterstützung wirkt ebenso wie die Netzwerkkontakte als protektiver Faktor der Befindlichkeit ( $r_{\text{sozu/dep}} = -.20$ ,  $p < .01$ ;  $r_{\text{sozu/selbst}} = .30$ ,  $p < .001$ ;  $r_{\text{kontakt/selbst}} = .29$ ,  $p < .001$ ).

In einem erweiterten Modell wird zusätzlich der Einfluss der sozialen Beziehungen im Vollzug berücksichtigt (siehe Abbildung 1). Es wird davon ausgegangen, dass enge Beziehungen zu den Mitinsassen zu einer besseren Befindlichkeit beitragen. Schlechte Beziehungen zum Anstaltspersonal und ein schlechtes Anstaltsklima sollten sich in höherer Depressivität niederschlagen. Eine hohe wahrgenommene Unterstützung (z. B. seitens der Mitinsassen) soll das Eingehen enger emotionaler Beziehungen zu den Mitgefangenen begünstigen.





$\chi^2 (45) = 68.806$ ; CFI = .968; AGFI = .935; RMSEA = .043; CO (90) = .020-.063

Abbildung 1: Der Einfluss sozialer Unterstützung und sozialer Beziehungen auf die psychische Befindlichkeit

Die Auswertungen zeigen, dass für das psychische Befinden der Insassen vor allem die emotionalen Bindungen zu den Mitinsassen ausschlaggebend sind. Schlechte Beziehungen zum Personal stehen mit höherer Depressivität in Zusammenhang. Sie machen sich damit zwar in negativem Affekt bemerkbar, führen aber offensichtlich nicht zu ernsthaften Zweifeln an der eigenen Person. Ein positiv bewertetes, transparentes und auf die Resozialisierung ausgerichtetes Anstaltsklima trägt in hohem Maße zu einem besseren Verhältnis zwischen Inhaftierten und Anstaltspersonal bei. Die Daten legen nahe, dass die Unterstützung durch Personen außerhalb der Anstalt, die sich in erster Linie in den Netzwerkkontakten und dem Unterstützungserhalt niederschlägt, vor allem selbstwertrelevant ist, während sich die Unterstützung durch Personen innerhalb des Vollzuges in erster Linie affektiv niederschlägt.

Wird anstelle der beiden Befindlichkeitsvariablen die Normorientierung, als latenter Faktor, in das bereits vorgestellte Strukturgleichungsmodell eingesetzt, zeigt sich, dass die Normorientierung erheblich von der Qualität der Beziehungen im Vollzug beeinflusst wird. Das empirische Modell stimmt mit dem theoretischen Modell weitgehend überein ( $\chi^2(58)=99.593$ ,  $p < .001$ ; CFI=.952; AGFI=.922; RMSEA=.051; CO(90)=.033-.067), welches einen direkten Einfluss aller Beziehungsaspekte sowie der wahrgenommenen Unterstützung und der Netzwerkkontakte auf die Normorientierung zugrundegelegt hat. Die Varianzaufklärung fällt mit 42 % hoch aus. Nehmen die Inhaftierten das Anstaltspersonal als feindselig wahr, ist bei den Gefangenen eine geringere Orientierung an gesellschaftlichen Normen und Regeln feststellbar ( $r = -.39$ ,  $p < .001$ ). Enge Bindungen an die Mitinsassen gehen, übereinstimmend mit den Theorien und Befunden zur Subkulturbildung, mit einer geringen Normorientierung einher ( $r = -.30$ ,  $p < .001$ ). Ein positiv eingeschätztes Anstaltsklima fördert hingegen eine prosoziale Orientierung ( $r = .19$ ,  $p < .01$ ). Ebenso wirken sich Netzwerkkontakte förderlich auf die Normorientierung aus ( $r = .20$ ,  $p < .01$ ). Die wahrgenommene Unterstützung

steht in einem indirekten Zusammenhang ( $r = -.10$ ,  $p > .05$ ) zur Normorientierung, der über die Bindung an die Mitinsassen vermittelt wird.

## **Diskussion**

Die vorgestellten Befunde zeigen in Einklang mit anderen Untersuchungen (Biggam & Power, 1997; Wright, 1993), dass soziale Unterstützung und die Qualität der Beziehungen zu den Mitinsassen und dem Anstaltspersonal die psychische Befindlichkeit und die soziale Einstellung der Inhaftierten signifikant beeinflussen. Sowohl die Höhe als auch die Richtung des Zusammenhangs schwanken in Abhängigkeit vom jeweils fokussierten Unterstützungsaspekt und dem Kriterium. Obwohl die Effekte statistisch gesehen relativ gering ausfallen, kann ihnen für die Praxis eine hohe Bedeutung zugemessen werden. Der Einfluss sozialer Beziehungen ist dabei auch vor dem Hintergrund zu bewerten, dass soziodemographische Faktoren wie Alter und sozialer Status oder haftbezogenen Merkmale wie die Inhaftierungszeit oder das Delikt keinen nennenswerten Beitrag zur Vorhersage der psychischen Befindlichkeit der Inhaftierten leisten.

Ansatzpunkte für die Praxis im Jugendstrafvollzug bieten vor allem die Befunde zur Bedeutung der Beziehungen zum Anstaltspersonal und den Mitinsassen. Enge emotionale Beziehungen zu den Mitinsassen sind für das Wohlbefinden der Inhaftierten von hoher Bedeutung, verstärken aber die Gefahr eines Abgleitens in deviante Subkulturen. Statt den Einfluss der Mitinsassen zu unterbinden, was nicht zuletzt aufgrund seiner befindlichkeitsstabilisierenden Funktion schwer sein dürfte, sollte vermehrt versucht werden, den Einfluss der Inhaftierten positiv zu nutzen und einen Gruppendruck hin zu normkonformem Verhalten zu entfalten. Kontinuierliche Ansprechpersonen und persönliche Zuwendung durch die Mitarbeiter des Vollzugs können dem Abbau bestehender Feindbilder und Feindseligkeiten dienen. Transparente Abläufe und Entscheidungsstrukturen im Vollzugssystem tragen zu einer höheren Akzeptanz durch die Inhaftierten und damit zu einem positiveren Anstaltsklima bei, welches sich wiederum positiv auf die Beziehungen zu den Mitarbeitern auswirkt. Regelmäßige Kontakte zu den Angehörigen, die im Haftverlauf sowohl selbstwert- als auch normfördernd wirken, sollten von den Anstalten unterstützt und möglichst weiter ausgebaut werden.

Die dargestellten Ergebnisse und die daraus abgeleiteten Schlussfolgerungen sind allerdings unter dem generellen Vorbehalt zu betrachten, dass es sich hier um Querschnittbefunde handelt, die einer weiteren längsschnittlichen Untersuchung bedürfen, um Aussagen über Kausaleffekte zu gestatten. Hinsichtlich des Zusammenhangs von wahrgenommener Unterstützung und Depressivität (Rook, 1984, 1992; Shinn, Lehmann & Wong, 1984) oder des Zusammenhangs zwischen Normorientierung und den Beziehungen zu den Anstaltsmitarbeitern ist sicherlich nicht von einfachen Kausaleffekten, sondern von Wechselwirkungen auszugehen. Eine Einschränkung der Befunde ergibt sich zudem daraus, dass der Anteil von Straftätern, die zu längeren Haftstrafen und damit zumeist aufgrund schwerwiegenderer Delikte verurteilt wurden, in der Querschnittstichprobe überrepräsentiert ist. Eine Generalisierung der Befunde auf die Population im Jugendstrafvollzug insgesamt ist dadurch nur mit Einschränkung möglich. Allerdings konnten die hier vorgestellten Befunde an einer fortlaufend rekrutierten Stichprobe ( $N = 698$ ), mit heterogen zusammengesetzter Klientel im Kern repliziert werden (Hosser, 2001b). Dabei zeigte sich, dass die Effekte der wahrgenommenen Unterstützung in Abhängigkeit vom Alter, der Inhaftierungszeit und der Deliktart variieren, was die weiterführende Frage nach intervenierenden Variablen im

Unterstützungsprozess wie z. B. Persönlichkeitseigenschaften oder Bewältigungsfähigkeiten aufwirft.

## Literatur

- Aymanns, P. (1992). *Krebserkrankung und Familie. Zur Rolle familialer Unterstützung im Prozeß der Krankheitsbewältigung*. Göttingen: Huber.
- Biggam, F. H., Power, K. G. (1997). Social support and psychological distress in a group of incarcerated young offenders. *International Journal of Offender Therapy and Comparative Criminology*, 41, 213-230.
- Dahle, K.-P. & Steller, M. (1990). Coping im Strafvollzug: Eine Untersuchung zu Haftfolgen bei Jugendlichen. *Zeitschrift für experimentelle und angewandte Psychologie*, 37, 31-51.
- Gottfredson, M. R. & Hirschi, T. (1990). *A General Theory of Crime*. Stanford: Stanford University Press.
- Greve, W. & Hosser, D. (1998). Psychische und soziale Folgen einer Jugendstrafe. Forschungsstand und Desiderate. *Monatsschrift für Kriminologie und Strafrechtsreform*, 81, 83-103.
- Greve, W., Hosser, D. & Pfeiffer, C. (1997). Gefängnis und die Folgen. Identitätsentwicklung und kriminelles Handeln während und nach Verbüßung einer Jugendstrafe. JuSt-Bericht Nr. 1. *KFN Forschungsberichte Nr. 64*. Hannover: KFN.
- Haney, C., Banks, C. & Zimbardo, P. (1973). Interpersonal Dynamics in a Simulated Prison. *International Journal of Criminology and Penology*, 1, 69-97.
- Hautzinger, M. & Bailer, M. (1993). *Allgemeine Depressions Skala (ADS). Manual*. Weinheim: Beltz-Test.
- Hirschi, T. (1969). *Causes of Delinquency*. Berkeley: University of California Press.
- Hosser, D. (2001a). *Soziale Unterstützung im Strafvollzug. Hafterleben und protektive Faktoren bei jungen Männern*. Baden-Baden: Nomos.
- Hosser, D. (2001b). Jugendstrafe im Spannungsfeld zwischen Integration und Desintegration. Soziale Beziehungen und Haftfolgen im Jugendstrafvollzug. In M. Bereswill & W. Greve (Hrsg.), *Forschungsthema "Strafvollzug"*. Baden-Baden: Nomos (im Druck).
- Hosser, D. & Greve, W. (1999). Gefängnis und die Folgen. Identitätsentwicklung und kriminelles Handeln während und nach Verbüßung einer Jugendstrafe. Das Erhebungsinstrument der standardisierten Befragung. JuSt-Bericht Nr. 3. *KFN Forschungsberichte Nr. 77*. Hannover: KFN.
- Liebling, A. (1992). *Suicides in prison*. London: Routledge.
- Quinton, D., Pickles, A., Maughan, B. & Rutter, M. (1993). Partners, peers, and pathways: Assortative pairing and continuities in conduct disorder. *Development and Psychopathology*, 5, 763-783.
- Rook, K. S. (1984). The negative side of social interaction: Impact on psychological well-being. *Journal of Personality and Social Psychology*, 46, 1097-1108.
- Rook, K. S. (1992). Detrimental aspects of social relationships: Taking stock of an emerging literature. In H. O. F. Veiel & U. Baumann (Eds.), *The meaning and measurement of social support* (pp. 157- 169). New York: Hemisphere.
- Sampson, R. J. & Laub, J. H. (1993). *Crime in the Making. Pathways and Turning Points through Life*. Cambridge, MA: Harvard University Press.
- Shinn, M., Lehmann, S. & Wong, N. W. (1984). Social interaction and social support. *Journal of Social Issues*, 40 (4), 55-76.
- Werner, E. E. (1993). Risk, resilience, and recovery: Perspectives from the Kauai Longitudinal Study. *Development and Psychopathology*, 5, 503-515.
- Wright, K. N. (1993). Prison environment and behavioral outcomes. *Journal of Offender Rehabilitation*, 20, 93-113.



# Dimensionen retributiver Gerechtigkeit: Moralische Standards und ihre Auswirkungen auf Sanktionswünsche<sup>1</sup>

*Jörg Hupfeld<sup>2</sup>*

*Institut für Psychologie der Universität Bern*

Während zur Frage der Dimensionalität distributiver Gerechtigkeitsüberlegungen empirische Untersuchungen vorliegen, fehlen bislang vergleichbare Strukturanalysen für den Bereich der retributiven Gerechtigkeit. Bei der Bewertung einer Straftat können z.B. relativistische oder universalistische Standards angewendet werden, die Überlegungen können primär fürsorge- oder gerechtigkeitsorientiert sein und sich auf die Absichten des Straftäters oder die Konsequenzen der Straftat beziehen. Berichtet werden die empirischen Befunde einer schriftlichen Befragung zu Reaktionen der Bevölkerung auf Rechtsverstöße. Jede der 691 teilnehmenden Personen beurteilte anhand eines standardisierten Fragebogens eine von insgesamt drei Fallgeschichten (Raubüberfall, Körperverletzung, Betrug). Im Rahmen der Untersuchung wurde zum einen der Frage nachgegangen, inwieweit die in der Forschung diskutierten Gerechtigkeits- und Moralorientierungen tatsächlich empirisch voneinander abgrenzbar sind, bzw. welche allgemeine Struktur ihnen zugrunde liegt. Zum anderen wurden Zusammenhänge zwischen den grundlegenden Gerechtigkeitsdimensionen und Sanktionspräferenzen (Sanktionshärte, Akzeptanz alternativer Massnahmen) analysiert.

## **Das Nachdenken über Moral und Gerechtigkeit**

Tagtäglich werden wir durch die Massenmedien mit Nachrichten über Straftaten und Gerichtsurteile konfrontiert und nicht nur an Stammtischen entbrennen heftige Diskussionen, wie die berichteten Taten zu bewerten sind und was mit den Tätern geschehen soll. Angesichts der grossen gesellschaftspolitischen Bedeutung dieses Themenbereichs erscheint es verständlich, dass sich auch die empirische Sozialforschung seit geraumer Zeit damit beschäftigt, welche Attributionsprozesse bei der Beurteilung von normabweichendem Verhalten eigentlich ablaufen (vgl. Heider, 1977; Shaver, 1985; Weiner, 1995), ob und warum die Bevölkerung nach härteren staatlichen Strafen ruft (Sessar, 1992; Zamble & Kalm, 1990), ob bestimmte Persönlichkeitsmerkmale mit Strafhärte assoziiert sind (vgl. Feather, 1999) und welche Ziele Menschen mit ihren Strafwünschen verfolgen (vgl. Hupfeld, 1996; Oswald, 1994; Miller & Vidmar, 1981). Zusammenfassend lässt sich feststellen, dass sich ein grosser Teil dieser Forschung direkt oder indirekt mit der Frage der Quantifizierung von Schuldzuschreibungen und Strafhärte beschäftigt. Relativ gering ist jedoch das Forschungsinteresse in Hinblick auf die grundsätzliche Frage, in welcher Weise Menschen über die Frage der Schuldhaftigkeit, Moral und Gerechtigkeit nachdenken, welche ethischen Standards sie hierbei anlegen. Während im Bereich der distributiven Gerechtigkeit, d.h. der Zuteilung positiver Güter, eine Vielzahl theoretischer und empirischer Studien über unterschiedliche Gerechtigkeitsstandards sowie ihre Geltungsbereiche und Auswirkungen vorliegt (vgl. Deutsch, 1975, 1985; Reis, 1984), sind vergleichbare Studien im Bereich der retributiven Gerechtigkeit eher selten. Zwar könnte die Psychologie bei der empirischen Klärung dieser Frage auch auf die umfangreiche Diskussion über Moral und Ethik in der praktischen Philosophie zurückgreifen. Erschwert wird dies jedoch durch die dort vorliegende Vielfalt unterschiedlichster Ansätze und Differenzierungen (vgl. Witte & Doll, 1995). Aufgrund der Vielzahl der diskutierten ethischen Grundpositionen sind der empirischen Forschung zudem aus pragmatischen Gründen Beschränkungen auferlegt. Nachfolgend sollen daher nur diejenigen ethischen Grundpositionen kurz skizziert werden, die Gegenstand der vorliegenden Untersuchung waren.

Eine Differenzierung, die sich z.B. sowohl bei Forsyth (1980) als auch bei Witte und Doll (1995) findet, bezieht sich auf die Allgemeingültigkeit der Bewertungsstandards. Die

Annahme, dass es sich bei den anzulegenden Standards um solche handelt, die eine universelle Gültigkeit aufweisen, wird als Universalismus bzw. *Absolutismus* bezeichnet. Im Gegensatz hierzu stehen verschiedene Formen des moralischen *Relativismus*. Hierbei wird anerkannt, dass moralische Standards in Abhängigkeit von Zeit, Kultur, Person und/oder Situation variieren können.

Eine weitere häufig genannte Differenzierung bezieht sich auf die Frage, ob zur Bewertung einer Tat eher die Gesinnung des Täters, d.h. seine Motive und Absichten heranzuziehen sind, oder ob es eher die Konsequenzen der Tat sind, die über die Frage der Bewertung als „gut oder schlecht“ entscheiden. Die erstgenannte Auffassung findet beispielsweise ihren Ausdruck im Kategorischen Imperativ Immanuel Kants. Die Fokussierung der Konsequenzen einer Handlung findet sich hingegen in der Vielzahl neuerer teleologischer bzw. utilitaristischer Ethiken (einführend hierzu z.B. Anzenbacher, 1992; Höffe, 1975).

Nachfolgend soll die Bewertung der Motive und Absichten des Akteurs als *Intentionalismus* und die Bewertung der Handlungsfolgen als *Konsequenzialismus* bezeichnet werden.

In ihrem vielbeachteten Buch „Die andere Stimme“ wies die Entwicklungspsychologin Carol Gilligan (1982) darauf hin, dass in der bisherigen (psychologischen) Moraltheorie möglicherweise eine bestimmte Moralperspektive systematisch unberücksichtigt geblieben sei. Nach Gilligans Meinung steht Kohlbergs Moralprinzip der Gerechtigkeit (*Justice*) ein Moralprinzip der Fürsorge (*Care*) entgegen. Folgt man den Arbeiten von Gilligan (1982), Lyons (1988) und Benhabib (1995), so richtet sich eine Justice-Orientierung eher auf die Anwendung situationsübergreifender Regeln und Prinzipien und die allgemeinen Rechte und Pflichten des verallgemeinerten autonomen Anderen. Hingegen fokussiert eine Care-Orientierung den konkreten Anderen, d.h. das Interaktionsgefüge und die Bedürfnisse der real Beteiligten und die Aufrechterhaltung ihrer Beziehungen (vgl. Hupfeld, 2000).

Insbesondere zur Differenzierung zwischen Care und Justice sowie den postulierten Beziehungen zum Geschlecht des Urteilers liegen bereits empirische Untersuchungen vor (vgl. z.B. Nunner-Winkler, 1995). Vieles deutet darauf hin, dass diese beiden ethischen Grundpositionen zwar voneinander abgrenzbar sind, sich jedoch nicht gegenseitig ausschließen. Der empirische Zusammenhang von Care und Justice mit den übrigen aufgeführten Ethiken wurde hingegen bislang kaum untersucht. Theoretisch könnte man davon ausgehen, dass z.B. eine Care-Orientierung positiv mit einem gewissen Ausmaß an Relativismus einhergeht. In einer Untersuchung von Forsyth, Nye und Kelley (1988) fand sich jedoch zumindest auf Trait-Ebene ein signifikanter, leicht negativer Zusammenhang ( $r = -.13$ ). Es bleibt abzuwarten, inwieweit dieser Befund replizierbar ist. Zur empirischen Abgrenzbarkeit und Struktur der übrigen ethischen Grundpositionen sind dem Autor keine weiteren Untersuchungen bekannt. Zu vermuten ist jedoch, dass zwischen Intentionalismus und Care gewisse empirische Zusammenhänge bestehen. So setzt die Auseinandersetzung mit einem konkreten Täter und Opfer sowie der zwischen ihnen bestehenden Beziehung voraus, dass man sich auch mit den Beweggründen auseinandersetzt, die zur Tat geführt haben. Hierbei könnte zugleich eine gewisse Bereitschaft zur Relativierung moralischer Standards (z.B. Situations- und Beziehungsangemessenheit) bestehen. Andererseits sind jedoch auch Zusammenhänge zwischen Intentionalismus und Absolutismus durchaus plausibel, sofern es Menschen gibt, die z.B. entsprechend der deontologisch-universalistischen Moralvorstellung des kategorischen Imperativs urteilen.

Die empirische Forschung hat gezeigt, dass moralische Urteile in hohem Ausmaß situationsabhängig sind (vgl. Kurtines, 1986). Ob dies auch für die dimensionale Struktur moralischer Bewertungsstandards gilt, oder ob sich die Situationsabhängigkeit nur in unterschiedlichen Mittelwerten niederschlägt, ist weitgehend ungeklärt. Darüber hinaus stellt sich jedoch auch die Frage nach den verhaltensmässigen Implikationen unterschiedlicher ethischer Standards. Ob systematische Zusammenhänge mit der Sanktionshärte oder der Bevorzugung alternativer außergerichtlicher Reaktionen auf Straftaten bestehen, ist ebenfalls

weitgehend ungeklärt. So wird von manchen Autoren vermutet, dass z.B. eine Care-Orientierung mit mildereren Sanktionen assoziiert ist als eine Justice-Orientierung (vgl. Drewniak, 1994). Und auch die Akzeptanz alternativer Reaktionen dürfte eher positiv mit einer Care-Orientierung korrelieren. Hingegen deuten einige empirische Befunde darauf hin, dass Absolutismus eher positiv mit Strafhärte korreliert (vgl. McHoskey, 1996).

## **Methode**

Aus den Telefonverzeichnissen in Bern (Schweiz) und Erlangen (Deutschland) wurde jeweils eine Zufallsstichprobe von Teilnehmernummern gewählt. Nach telefonischer Einholung der Teilnahmebereitschaft wurden standardisierte Fragebögen sowie frankierte Rückumschläge zugesandt. Insgesamt wurden 691 verwertbare Fragebögen zurückgeschickt (Bern: N= 357; Erlangen: N= 333). Der Frauenanteil bei den Untersuchungsteilnehmern lag bei 54 Prozent, das Durchschnittsalter betrug 44 Jahre. Der jüngste Teilnehmer war 18 und der älteste 92 Jahre alt.

Der Fragebogen enthielt jeweils eine von drei zufällig ausgewählten Fallgeschichten. Es handelte sich hierbei entweder um einen Raubüberfall, eine Körperverletzung oder ein Betrugsdelikt. Die Deliktschwere der drei Fälle war gemäss der Ergebnisse von Voruntersuchungen vergleichbar. Von den insgesamt 691 Fragebögen entfielen 223 auf das Raubdelikt, 210 auf die Körperverletzung und 228 auf den Betrug. 30 Fragebögen waren leider nicht mehr eindeutig zuzuordnen.

Auf die Darstellung der Fallgeschichte folgte jeweils eine Reihe fallbezogener und allgemeiner Fragen, die nicht alle Gegenstand der vorliegenden Untersuchung sind. Die nachfolgende Beschreibung beschränkt sich daher auf die hier interessierenden fallbezogenen Variablen. Erfasst wurden u.a. die zuvor beschriebenen ethischen Standards, die fallbezogene Strafhärte und die Akzeptanz alternativer Reaktionen (vgl. Tabelle 1).

Tabelle 1: Konstrukte und Beispielitems

Konstrukt	Anzahl der Items	Beschreibung/Beispiel
Relativismus	3	Item: Je nach Situation wird man zur moralischen Bewertung der Handlung von Beat N. einen anderen Maßstab heranziehen. Es gibt keine situationsunabhängigen moralischen Prinzipien.
Absolutismus	3	Item: Zur Bewertung der Handlung von Beat N. muss ich mir einzig und allein anschauen, gegen welche allgemeingültigen moralischen Prinzipien sie verstösst oder mit welchen sie eventuell übereinstimmt.
Intentionalismus	3	Item: Bei der Beurteilung der Tat von Beat N. sollte man vor allem berücksichtigen, was ihn dazu bewogen hat, sich so zu verhalten.
Konsequenzialismus	3	Item: Um sagen zu können, wie das Verhalten von Beat N. moralisch zu bewerten ist, muss ich überwiegend die konkreten Folgen seiner Handlung, z.B. eventuelle Verletzungen, finanzielle Schäden usw., berücksichtigen.
Justice	3	Item: Bei der Beurteilung des Verhaltens von Beat N. sollte man eine übergeordnete Gerechtigkeit im Auge behalten, indem man Regeln zur Beurteilung heranzieht, die auf andere Fälle übertragbar sind.
Care	3	Item: Ohne bereit zu sein, die Beteiligten anzuhören, um mehr über ihre spezifischen Bedürfnisse und Interessen zu erfahren und darauf eingehen zu können, wird man im konkreten Fall kein gerechtes Urteil finden.
Strafhärte	1	Grafisches deservingness-Rating: Beurteilung einer vorgegebenen Gefängnisstrafe von 8 Monaten auf Bewährung auf einer 120 mm langen Skala mit den Endpunkten viel zu hart vs. viel zu milde
Akzeptanz alternativer Reaktionen	6	Fragen nach der Befürwortung alternativer Reaktionen anstelle einer herkömmlichen Bestrafung durch ein Gericht (siehe Tab. 9)

Zur Überprüfung der Struktur der fallbezogenen retributiven Überlegungen wurden die Daten mittels Multidimensionaler Skalierung (MDS) mit dem Programm SYSTAT 9 analysiert. Gewählt wurde hierbei die Skalierungsmethode nach Young. Grundlage der MDS waren jeweils die Produktmomentkorrelationen.

Zur Überprüfung der Stabilität der gefundenen generellen Lösung wurde ein mehrstufiges Verfahren gewählt. Berechnet wurde zum einen eine MDS auf der Basis des gesamten Datensatzes. Anschliessend wurden getrennte INDSCAL-Analysen (individual differences



scaling) durchgeführt. In der ersten INDSCAL-Analyse wurde eine gemeinsame Lösung auf der Basis der beiden Korrelationsmatrizen für Bern und Erlangen gesucht. In der zweiten INDSCAL-Analyse wurde eine gemeinsame Lösung auf der Basis der drei deliktspezifischen Korrelationsmatrizen (Raubüberfall, Körperverletzung, Betrug) gesucht. Jede INDSCAL-Lösung ist die an die jeweiligen Korrelationsmatrizen optimal angepasste gemeinsame mehrdimensionale Darstellung. Neben der Gesamtanpassung der INDSCAL-Modelle lassen sich zusätzlich Informationen entnehmen, wie gut die gemeinsamen Lösungen an die jeweiligen Einzelmatrizen angepasst sind und inwieweit stadt- bzw. deliktspezifische Unterschiede in der Gewichtung der gemeinsamen Dimensionen zu beobachten sind. Solche Gewichtungsunterschiede lassen darauf schließen, dass bestimmte Zusammenhänge zwischen den Items je nach Teilstichprobe systematischer akzentuiert oder verringert sind, d.h. dass z.B. die Endpole einer grundlegenden Ethik-Dimension bei einem Körperverletzungsdelikt extremer negativ miteinander korrelieren als bei einem Eigentumsdelikt. Hiervon zu unterscheiden sind stichprobenbezogene Unterschiede in den Mittelwerten von Items oder Skalen. Diese werden im Anschluss an die dimensionsanalytischen Befunde mittels t-Test oder Varianzanalyse überprüft.

## Ergebnisse

*Ergebnisse der MDS auf der Basis des Gesamtdatensatzes (N = 691):*

Der nachfolgende Scree-Plot verschiedener Lösungen zeigt keinen klaren Knick. Allein auf der Basis der Stress-Werte ist somit keine klare Entscheidung über die Dimensionalität der zu bevorzugenden Lösung möglich.

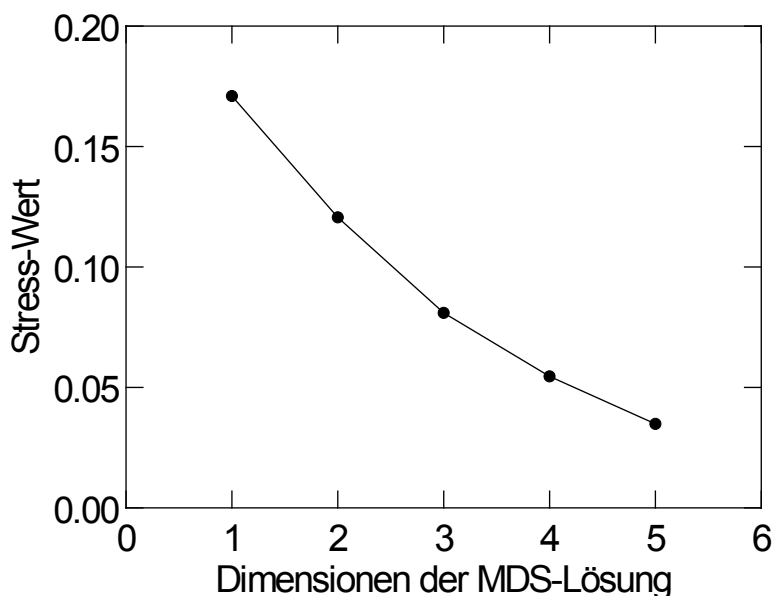


Abbildung 1: Scree-Plot für MDS-Lösungen auf der Basis des Gesamtdatensatzes

Eine genauere Inspektion der jeweiligen Punkte-Konfigurationen zeigt jedoch, dass bereits in einer zweidimensionalen Lösung bis auf Care und Intentionalismus alle moralischen Orientierungen auf Itemebene voneinander abgrenzbar sind (vgl. Abb. 2). In einer dreidimensionalen Lösung sind diese Abgrenzungen noch eindeutiger. Zwischen Care- und Intentionalismus-Items sind jedoch auch hier keine klaren räumlichen Abgrenzungen

möglich. Auf der Basis dieses Befundes beschränken sich die nachfolgenden Ergebnisdarstellungen auf die zweidimensionale MDS-Lösung.

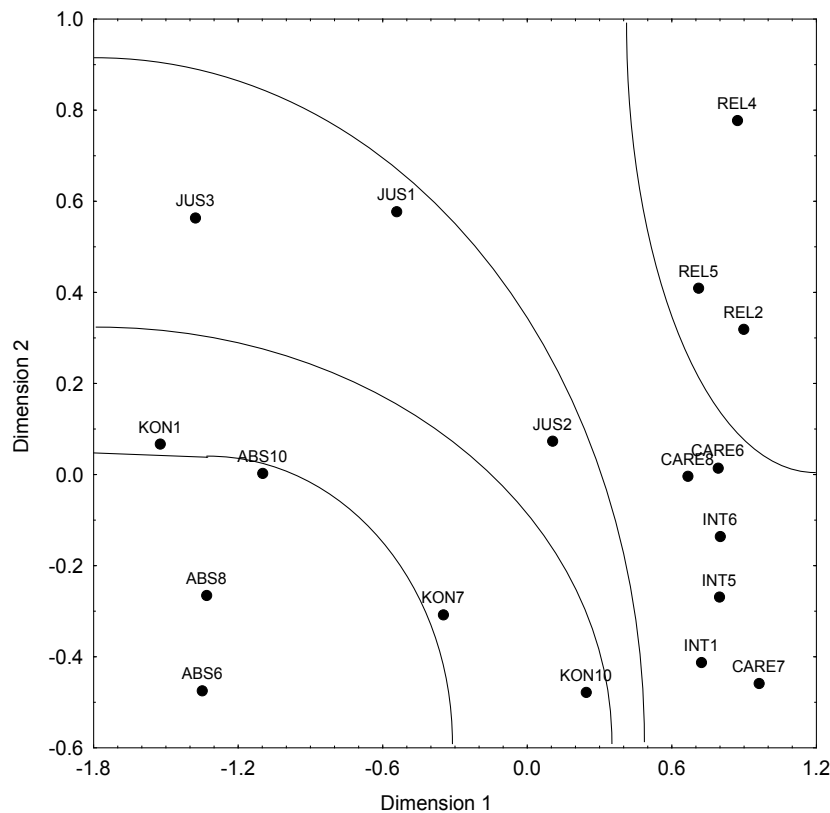


Abbildung 2: Zweidimensionale MDS-Lösung (Item-Ebene)

Innerhalb der zweidimensionalen Lösung (siehe Abb. 2) finden sich ausgehend von der linken unteren Ecke der Grafik jeweils voneinander abgrenzbare Regionen mit Absolutismus-Items (ABS), Konsequentialismus-Items (KON) und Justice-Items (JUS). In der rechten oberen Ecke finden sich räumlich klar abgegrenzt alle Relativismus-Items (REL). Die Care- (CARE) und Intentionalismus-Items (INT) finden sich hingegen in einer gemeinsamen Region im rechten unteren Bereich.

Interpretiert man die Konfiguration nicht regional sondern dimensional, so ergibt sich das folgende Bild: In den beiden Hauptdiagonalen liegen sich jeweils Care- und Justice- bzw. Absolutismus- und Relativismus-Items gegenüber, wohingegen Konsequentialismus- sowie Intentionalismus-Items einander im unteren Bereich der Konfiguration gegenüberstehen. Ein ähnliches Bild zeigt sich, wenn man eine MDS auf der Basis der Interkorrelationen der Skalen (Durchschnittswerte der entsprechenden Items; vgl. Borg & Mohler, 1993) durchführt (vgl. Abb. 3).

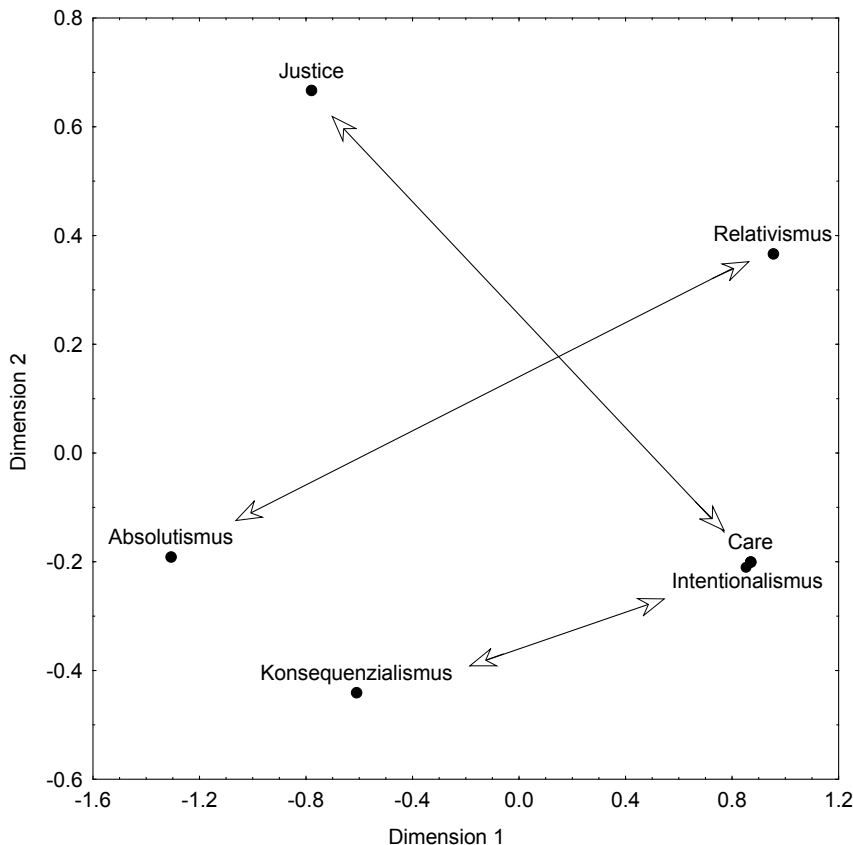


Abbildung 3: Zweidimensionale MDS-Lösung (Skalen-Ebene)

Obwohl Care und Intentionalismus innerhalb der ordinalen MDS-Lösungen kaum eindeutig voneinander abzugrenzen sind, besteht dennoch ein deutlicher konzeptioneller Unterschied. Die räumliche Nähe der Items bzw. Skalen kann zwar als empirischer Beleg für die Annahme dienen, dass für Care-Überlegungen die Beachtung der Schädigungsabsichten wesentlich ist. Jedoch muss umgekehrt nicht jede intentionalistische Betrachtungsweise auch zwangswise mit Care-Überlegungen einhergehen. Hierauf deutet ebenfalls die bivariate Korrelation in Höhe von 0.54 zwischen Care und Intentionalismus hin. Es handelt sich hierbei zwar um die zahlenmässig höchste Skaleninterkorrelation, weshalb die Skalenpunkte innerhalb der MDS-Lösungen nahezu aufeinander fallen. Die absolute Höhe der Korrelation deutet jedoch auch darauf hin, dass der empirische Zusammenhang nicht perfekt ist und andere Kombinationen zu beobachten sind. Kombiniert mit absolutistischen Überlegungen können intentionalistische Erwägungen beispielsweise Anzeichen einer deontologischen Moral im Sinne Kants sein. Die signifikante bivariate Korrelation zwischen Intentionalismus und Absolutismus in Höhe von -.16 deutet zwar darauf hin, dass diese Kombination seltener ist, jedoch ebenfalls vorkommt.

### *INDSCAL-Ergebnisse*

Eine zweidimensionale INDSCAL-Lösung auf der Basis der beiden Korrelationsmatrizen aus Bern und Erlangen bestätigt die zuvor gefundene Struktur. Die Anpassungsgüte der gemeinsamen Lösung ist mit einem Stresswert in Höhe von 0.23 auf akzeptablem Niveau. Die gemeinsame Lösung scheint zudem beide Datensätze ähnlich gut zu repräsentieren. Für die Daten aus Bern beträgt der Stresswert 0.15 und für Erlangen 0.18. Die Dimensionsgewichtungen sind für beide Datensätze nahezu identisch. Ähnliches gilt für eine zweidimensionale INDSCAL-Lösung auf der Basis der delikt-spezifischen Korrelationsmatrizen. Die Anpassungsgüte der gemeinsamen Lösung beträgt insgesamt 0.26. Speziell für die Raubüberfalldaten beträgt der Stresswert 0.16, für die Körperverletzungsdaten 0.21, und für die Betrugsdaten beträgt der Stresswert der

gemeinsamen Lösung 0.19. Auch finden sich für alle drei Delikte wiederum fast identische Dimensionsgewichtungen.

Die Konfigurationen der Items in beiden INDSCAL-Lösungen weisen keine wesentlichen Unterschiede zu derjenigen Konfiguration auf, die auf der Basis des Gesamtdatensatzes gefunden wurde (siehe Abb. 2). Die Koordinaten der Items auf den einander entsprechenden Dimensionen aller drei hier dargestellten Skalierungen (MDS-gesamt, INDSCAL-städtespezifisch, INDSCAL-deliktspezifisch) korrelieren in Höhe von 0.99 bis 1. Es handelt sich somit um nahezu identische räumliche Konfigurationen, so dass eine getrennte grafische Darstellung der INDSCAL-Lösungen entfallen kann.

### *Faktorenanalytische Befunde*

Durch eine faktorenanalytische Auswertung der Skaleninterkorrelationsmatrix sollten weitere Hinweise auf die den retributiven moralischen Überlegungen zugrunde liegende dimensionale Struktur gewonnen werden. Gewählt wurde das Verfahren der Hauptachsenanalyse mit anschließender schiefwinkliger Rotation. Eine dreidimensionale Lösung erwies sich sowohl für den Gesamtdatensatz, als auch für die städte- und deliktspezifischen Datensätze als weitgehend stabil und interpretierbar. Die nachfolgende Darstellung kann sich daher wiederum auf die Ergebnisse für den Gesamtdatensatz beschränken. Zur Interpretation der Ladungsmuster wird nachfolgend die Mustermatrix dargestellt (vgl. Tabelle 2).

Tabelle 2: Mustermatrix der dreifaktoriellen obliquen Hauptachsen-Lösung (Ladungen mit einem Absolutbetrag  $< .15$  werden nicht dargestellt)

	Faktor 1	Faktor 2	Faktor 3
Care	.90		
Intentionalismus	.57		
Justice		.41	
Konsequenzialismus		.37	
Absolutismus		.43	-.57
Relativismus	.26		.46

Tabelle 3: Interkorrelationen der drei Faktoren

	Faktor 2	Faktor 3
Faktor 1	.11	.39
Faktor 2		-.24

Faktor 1 belegt den engen Zusammenhang zwischen Care und Intentionalismus, zwei Moralvorstellungen, die sich jeweils an konkreten Individuen orientieren. Das Ladungsmuster auf Faktor 2 zeigt einen empirischen Zusammenhang zwischen moralischen Überlegungen, die jeweils auf die eine oder andere Weise von konkreten Personen abstrahieren. Hierbei handelt es sich um Konsequenzialismus, Justice und Absolutismus. Interessanterweise schliessen sich diese beiden Grunddimensionen moralischer Überlegungen nicht aus. Es handelt sich um zwei nahezu unabhängige unipolare Dimensionen ( $r = .11$ ). Der dritte Faktor beschreibt eine bipolare Relativismus-Absolutismus-Dimension. Diese Befürwortung relativistischer bzw. Ablehnung absoluter moralischer Überlegungen korreliert positiv ( $r = .39$ ) mit individuumszentrierten Überlegungen (Faktor 1) und negativ ( $r = -.24$ ) mit der Bereitschaft, von konkreten Personen zu abstrahieren (Faktor 2).

### *Mittelwertsunterschiede in den moralischen Orientierungen*

Bei vier der sechs moralischen Orientierungen zeigen sich signifikante Mittelwertsunterschiede zwischen Bern und Erlangen (siehe Tabelle 4). Die praktische Bedeutung dieser Unterschiede ist mit maximal drei Prozent erklärter Varianz zwar eher

gering, doch ist das Muster recht eindeutig. In Erlangen werden in signifikant stärkerem Ausmaß absolutistische und in schwächerem Ausmaß relativistische Standards angelegt und in Bern werden signifikant stärker auf konkrete Individuen bezogene Überlegungen vorgenommen (Intentionalismus, Care).

Tabelle 4: Stadtbezogene Mittelwertsunterschiede in den moralischen Orientierungen

	Mittelwert		t-Wert	Signifikanz	Erklärte Varianz (Eta <sup>2</sup> )
	Bern	Erlangen			
Absolutismus	.70	.88	2.78	.01	.01
Relativismus	-.21	-.54	-4.36	.00	.03
Intentionalismus	.38	.13	-2.83	.01	.01
Care	.58	.25	-4.16	.00	.03

Auch in Hinblick auf die drei Delikte finden sich geringfügige signifikante Unterschiede zwischen den bevorzugten moralischen Standards (vgl. Tabelle 5). Das Unterschiedsmuster ist auch hier relativ eindeutig. Insgesamt werden beim Körperverletzungsdelikt im Durchschnitt mehr absolutistische und weniger relativistische, care-bezogene und intentionalistische Standards angelegt als bei den beiden anderen Delikten.

Tabelle 5: Deliktbezogene Mittelwertsunterschiede in den moralischen Orientierungen

	Mittelwert			F-Wert	Signifikanz	Erklärte Varianz (Eta <sup>2</sup> )
	Raub	Körperverletzung	Betrug			
Absolutismus	-.32	-.29	-.52	3.25	.04	.01
Relativismus	.61	.24	.63	11.51	.00	.03
Intentionalismus	.47	.09	.23	6.30	.01	.02
Care	.54	.33	.49	3.41	.03	.01

### *Moralische Überlegungen und Strafhärte*

Bei der Analyse der Korrelationen zwischen moralischen Überlegungen und Strafhärte finden sich nur geringe Zusammenhänge, die sich jedoch aufgrund des grossen Stichprobenumfangs nahezu alle als signifikant erweisen (vgl. Tabelle 6).

Tabelle 6: Korrelative Zusammenhänge zwischen moralischen Überlegungen (Skalenwerte) und der Strafhärte (N= 691)

	Absolutismus	Justice	Konsequenzialismus	Relativismus	Intentionalismus	Care
Strafhärte	.11**	.02	.13**	-.09*	-.16**	-.15**

\*  $p \leq .05$  \*\*  $p \leq .01$

Insgesamt sind individuumszentrierte Überlegungen (Intentionalismus, Care) eher negativ mit Strafhärte assoziiert und vom konkreten Individuum abstrahierende Überlegungen (insbes. Absolutismus und Konsequenzialismus) eher mit härteren Strafen. Im Rahmen einer multiplen linearen Regression erweisen sich nur Konsequenzialismus und Intentionalismus als Strafhärteprädiktoren mit eigenständigem Erklärungsbeitrag. Care verfehlt als Prädiktor knapp die Signifikanzgrenze. Offensichtlich liefern Care-Überlegungen keinen wesentlichen eigenständigen Vorhersagebeitrag, der über denjenigen intentionalistischer Erwägungen hinausreicht (vgl. Tabelle 7). Allein aufgrund der Art der angestellten moralischen Überlegungen ist die Härte der bevorzugten rechtlichen Sanktion mit ca. 5 Prozent erklärter Varianz insgesamt jedoch nur sehr unzureichend vorhersagbar.

Tabelle 7: Ergebnisse der linearen Regression zur Vorhersage der Strafhärte (N=691, multiples R=.23, multiples R<sup>2</sup>=.05)

	Beta	t-Wert	Signifikanz
Absolutismus	0.04	1.05	0.29
Relativismus	-0.02	-0.41	0.68
Konsequenzialismus	0.13	3.44	0.00
Intentionalismus	-0.11	-2.49	0.01
Care	-0.08	-1.83	0.07
Justice	-0.01	-0.20	0.84

### *Moralische Überlegungen und Akzeptanz alternativer Reaktionen*

Neben Fragen zur gewünschten Härte einer gerichtlichen Sanktion beantworteten die Vpn auch Fragen nach eventuellen Reaktionsalternativen, die sie anstelle einer herkömmlichen Strafe befürworten würden. In Tabelle 8 sind die Zusammenhänge zwischen moralischen Überlegungen und der prinzipiellen Befürwortung solcher Alternativen aufgeführt.

Tabelle 8: Korrelative Zusammenhänge zwischen moralischen Überlegungen (Skalenwerte) und der grundsätzlichen Befürwortung alternativer Reaktionen (N= 678)

	Absolutismus	Justice	Konsequenzialismus	Relativismus	Intentionalismus	Care
Alternative Maßnahme statt gerichtl. Sanktion (Ja/Nein)	0.02	-0.05	0.04	0.10*	0.11**	0.16**

\*  $p \leq .05$  \*\*  $p \leq .01$

Die Zusammenhänge sind nicht sehr ausgeprägt, jedoch ist generell eine Bevorzugung relativistischer, an konkreten Individuen ausgerichteter Überlegungen überzufällig positiv mit der Akzeptanz alternativer Reaktionen assoziiert. Andererseits führen aber absolutistische, vom konkreten Individuum abstrahierende Überlegungen nicht zu einer Ablehnung alternativer Reaktionen auf die Straftat.

Auch in Hinblick auf die einzelnen alternativen Reaktionen sind die Zusammenhänge eher gering (vgl. Tab. 9).

Tabelle 9: Korrelative Zusammenhänge zwischen moralischen Überlegungen (Skalenwerte) und der Befürwortung einzelner alternativer Reaktionen (N= 678)

	Absolutismus	Justice	Konsequenzialismus	Relativismus	Intentionalismus	Care
Geld an Opfer	.06	.03	.01	.03	-.06	.07
Gemeinnützige Spende	-.07	-.01	.03	-.01	.01	.03
Gemeinnützige Arbeit	-.04	-.04	-.06	-.05	.04	.01
Arbeit für Opfer	-.09*	.01	-.04	.08*	.06	.08*
Gespräch mit Opfer	.03	-.03	-.01	.13**	.12**	.19**
Persönliche Entschuldigung	.09*	-.03	.10*	.12**	.09*	.18**

\*  $p \leq .05$  \*\*  $p \leq .01$

Insbesondere an konkreten Individuen ausgerichtete moralische Erwägungen korrelieren positiv mit Massnahmen, die als Grundlage zur Klärung der sozialen Situation und Wiederherstellung einer zufriedenstellenden Beziehung zwischen Täter und Opfer dienen können (die persönliche Entschuldigung und das Gespräch mit dem Opfer). Dies gilt erwartungsgemäss insbesondere für care-orientierte moralische Erwägungen. Wiederum auffällig ist, dass diese Massnahmen auch bei der Bevorzugung absolutistischer, abstrakterer moralischer Standards nicht abgelehnt werden.

## Diskussion

Die empirischen Befunde belegen eine stabile Struktur voneinander abgrenzbarer moralischer Bewertungsstandards. Die Struktur verändert sich weder in Abhängigkeit vom Erhebungsort, noch in Abhängigkeit vom Delikt. Hierbei ist jedoch zu bedenken, dass sich beide Erhebungsorte in wesentlichen kulturellen Dimensionen stark gleichen. Und auch die in die Untersuchung einbezogenen Delikte wurden in Hinblick auf die subjektiv wahrgenommene Deliktschwere parallelisiert. Somit bleibt letztlich die Frage ungeklärt, ob sich ähnliche Strukturen auch bei grösserer kulturell- und deliktbedingter Vielfalt zeigen. Wünschenswert wäre hierbei sicherlich der Einsatz umfangreicherer Itembatterien als dies in der vorliegenden Untersuchung möglich war.

Generell jedoch stellt das Faktum der empirischen Differenzierbarkeit unterschiedlicher Arten moralischer Standards eine wesentliche Voraussetzung für weitere Untersuchungen dar. Auf der Basis dieser – oder ähnlicher – Befunde lassen sich u.a. gezielte Vorhersagen ableiten, welche Fallmerkmale in den Fokus der Aufmerksamkeit einer urteilenden Person gelangen. Sollten sich auch stabile personenbezogene Präferenzen für bestimmte moralische Orientierungen feststellen lassen, könnte so ein wesentlicher Beitrag zur Beantwortung der Frage geleistet werden, welche Personen auf welche Fallmerkmale reagieren (vgl. Hupfeld, 1999). Auch dem Problem, welche Auswirkungen dies auf die jeweiligen Sanktionswünsche hat, lässt sich auf diese Weise näherkommen. Die Zusammenhänge zwischen moralischen Orientierungen und Sanktionswünschen könnten jedoch auch deutlich anders ausfallen als in der vorliegenden Untersuchung. Letztendlich dürfte die jeweilige Kombination der Fallmerkmale (z.B. Absichten des Täters, Höhe des Schadens, die Art der Täter-Opfer-

Beziehung etc.) ausschlaggebend dafür sein, welche ethischen Standards mit milden oder harten Sanktionswünschen assoziiert sind.

## Literatur

- Anzenbacher, A. (1992). *Einführung in die Ethik*. Düsseldorf: Patmos.
- Benhabib, S. (1995). *Selbst im Kontext. Kommunikative Ethik im Spannungsfeld von Feminismus, Kommunitarismus und Postmoderne*. Frankfurt a.M.: Suhrkamp.
- Borg, I., & Mohler, P. P. (1993). Zur Indexbildung in der Facettentheorie. *ZUMA Nachrichten*, 33, 10-24.
- Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues*, 31, 137-149.
- Deutsch, M. (1985). *Distributive justice: A social psychological perspective*. New Haven, C.T.: Yale University Press.
- Forsyth, D. R. (1980). A taxonomy of ethical ideologies. *Journal of Personality and Social Psychology*, 39 (1), 175-184.
- Forsyth, D. R., Nye, J. L. & Kelley, K. (1988). Idealism, relativism, and ethic of caring. *The Journal of Psychology*, 122 (3), 243-248.
- Gilligan, C. (1982). *In a different voice*. Cambridge, MA: Harvard University Press.
- Heider, F. (1977). *Psychologie der interpersonalen Beziehungen*. Stuttgart: Klett.
- Höffe, O. (1975). *Einführung in die utilitaristische Ethik*. München: Beck.
- Hupfeld, J. (1996). *Jugendrichterliches Handeln*. Baden-Baden: Nomos.
- Hupfeld, J. (1999). Richter- und gerichtsbezogene Sanktionsdisparitäten in der deutschen Jugendstrafrechtspraxis. *Monatsschrift für Kriminologie und Strafrechtsreform*, 82 (5), 342-358.
- Hupfeld, J. (2000). Frauen, eine Gefahr für die Justiz? Geschlecht und Reaktionen auf Kriminalität. In Abteilung für die Gleichstellung von Frauen und Männern der Universität Bern (Hrsg.), *Frauen im Recht. Kindsmörderinnen und Richterinnen – Quoten und soziale Sicherheit* (S. 43-59). Bern: eFeF-Verlag.
- Kurtines, W. M. (1986). Moral behavior as rule governed behavior: Person and situation effects on moral decision making. *Journal of Personality and Social Psychology*, 50 (4), 784-791.
- Lyons, N. P. (1988). Two perspectives: On self, relationship, and morality. In C. Gilligan, J. V. Ward & J. McLean Taylor (eds.), *Mapping the moral domain* (pp. 21-48). Cambridge, MA: Harvard University Press.
- McHoskey, J. W. (1996). Authoritarianism and ethical ideology. *The Journal of Social Psychology*, 136 (6), 709-717.
- Miller, D. T. & Vidmar, N. (1981). The social psychology of punishment reactions. In M. J. Lerner & S. C. Lerner (eds.), *The justice motive in social behavior* (pp. 145-172). New York: Plenum Press.
- Nunner-Winkler, G. (1995). *Weibliche Moral*. München: DTV.
- Oswald, M. E. (1994). *Psychologie des richterlichen Strafens*. Stuttgart: Enke.
- Reis, H. T. (1984). The multidimensionality of justice. In R. Folger (ed.), *The sense of injustice: Social psychological perspectives* (pp. 25-61). New York: Plenum Press.
- Sessar, K. (1992). *Wiedergutmachung oder Strafen. Einstellungen in der Bevölkerung und der Justiz*. Pfaffenweiler: Centaurus-Verlag.
- Shaver, K. G. (1985). *The attribution of blame*. New York: Springer.
- Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. New York: The Guilford Press.



Witte, E. H. & Doll, J. (1995). Soziale Kognition und empirische Ethikforschung. Zur Rechtfertigung von Handlungen. In E. H. Witte (Hrsg.), *Soziale Kognition und empirische Ethikforschung* (S. 97-115). Lengerich: Pabst.

Zamble, E. & Kalm, K. L. (1990). General and specific measures of public attitudes toward sentencing. *Canadian Journal of Behavioural Science*, 22, 327-337.

<sup>1</sup> Das Projekt wird mit Mitteln der Volkswagenstiftung finanziert; Projektleitung: Prof. Dr. M. E. Oswald

<sup>2</sup> Dank für die Vorbereitung und Durchführung der Untersuchung gilt insbesondere Stefan Klug



# Die Verwendung ereigniskorrelierter Potentiale im Rahmen der Psychophysiologischen Aussagebegutachtung

*Prisca Jager & O. Berndt Scholz*

*Psychologisches Institut der Universität Bonn*

Die P300 Komponente tritt im Oddball-Paradigma auf bedeutsame, selten gezeigte Reize mit größerer Amplitude auf als auf häufige, neutrale Reize. Zur Täuschungsaufdeckung wird eine dritte Klasse von Reizen eingeführt, die ebenfalls aufgrund ihrer Tatspezifität für einen zu Recht beschuldigten Täter bedeutsam ist (siehe Tabelle 1). Die tatspezifischen aber bedeutsamen Reize erhalten ihre besondere Bedeutsamkeit durch die Aufgabe, die während der EEG-Ableitung zu absolvieren ist (z.B. abweichenden Knopfdruck bei Erscheinen der relevanten Reize). Die Formulierung bezieht sich auf die Deliktgruppe, der die Tat zugehört. Die P300-Amplituden für tatspezifische und tatunspezifische Reize sollten sich bei einer zu Recht beschuldigten Person gleichen und die der neutralen Reize übertreffen. Für einen zu Unrecht beschuldigten Täter, so die Annahme, sollte der Bedeutungsgehalt dieser tatspezifischen Reize eher dem der neutralen Items gleichen. Die Amplitude der P300-Komponenten sollte in diesem Fall nicht viel ausgeprägter sein als die bezüglich der neutralen Items (Scholz, 1999).

Tabelle 1: Unterschiedliche Itemtypen im Drei-Kategorien-Oddball-Paradigma

Zielreiz Tatunspezifisch (= U)	Relevant für die Deliktgruppe	1/6 der Items
Neutraler Reiz (= N)	Irrelevant für die Deliktgruppe und für die vorgeworfene Tat	2/3 der Items
Testreiz (= S) tatspezifisch	Relevant bezüglich der vorgeworfenen Tat	1/6 der Items

Farwell und Donchin haben in einer Untersuchung von 1991 vorgestellt, wie man ein Drei-Kategorien-Oddball-Paradigma zur Täuschungsaufdeckung nutzbar machen kann. Sie ließen 20 Probanden ein Mock Crime aus dem Themenbereich „Spionage“ begehen. Diese Spionagehandlung umfasste das Auswendiglernen bestimmter relevanter Informationen per Computerprogramm und die Durchführung der Handlung, die im Überreichen eines Umschlags an eine Person bestand. Einen Tag später sollten alle Probanden ihre simulierten Spionagehandlungen während einer EEG-Ableitung verheimlichen. Bezüglich eines weiteren Spionageszenarios hatten die Probanden keine tatspezifischen Informationen erhalten. Auf diese Weise wurde jeder Proband einmal zu Recht beschuldigt und einmal zu Unrecht beschuldigt. Dabei stellten die gelernten Spionageinformationen die tatspezifischen Items dar. Die tatunspezifischen Items wurden per Zufall aus den neutralen ausgewählt und erst durch die Aufgabenstellung bedeutsam. Abgeleitet wurde von drei Ableitorten frontal (Fz), zentral (Cz) und parietal (Pz) entlang der Mittellinie. Tatsächlich ähnelten die EKP-Verläufe bezüglich der tatspezifischen Reize denen bezüglich der tatunspezifischen Reize im Fall der zu Recht beschuldigten Probanden (Abbildung 1 A). Im Falle der zu Unrecht beschuldigten Probanden ähnelten die Verläufe eher denen der neutralen Reize (Abbildung 1 B). Mit Hilfe eines korrelativen Bootstrapverfahrens konnten die Autoren auch statistisch belegen, dass mit Hilfe der P300-Komponenten sicher zwischen den Beschuldigungsbedingungen unterschieden werden konnte. Die Treffsicherheit betrug 100%, bei einer Rate von 12,5% unentscheidbaren Fällen.

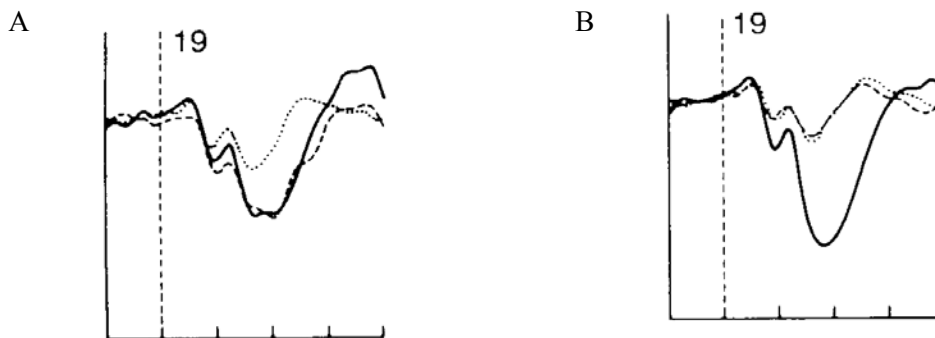


Abbildung 1: Ergebnisse der Untersuchung von Farwell und Donchin (1991). A: Die gemittelten EEG-Kurven eines zu Recht beschuldigten Probanden. B: Die gemittelten EEG-Kurven desselben Probanden in der Bedingung „zu Unrecht beschuldigt“ (gepunktete Linie = neutral, gestrichelte Linie = tatspezifisch, durchgezogene Linie = tatunspezifisch).

Relevant sind für die Psychophysiologische Aussagebegutachtung (PAG) zur Zeit in Deutschland solche Delikte, bei denen ein Verdächtiger einer Sexualstraftat, bzw. des Kindesmissbrauchs beschuldigt wird. Da dieser Bereich als Anwendungsfeld für die PAG diskutiert wird, werden auch experimentelle Untersuchungen in Analogie zu dieser Straftat benötigt. Forschungsparadigmen, die in Analogie zu einem Sexualstraftäter stehen, sollten folgende Charakteristika haben:

- Das Untersuchungsgeschehen konzentriert sich auf eine sexuelle Handlung;
- Der Tatvorwurf, sei es, dass er zu Recht oder Unrecht besteht, wirkt moralisch aversiv;
- Der Proband hat den Eindruck, während des Tatszenarios aus eigener Veranlassung gehandelt zu haben. Der anschließende Tatvorwurf veranlasst den Probanden, dass er sein eigenes Handeln leugnet, also Täuschungsverhalten zeigt.
- Wie einer Studie von Deegener (1995) zu entnehmen ist, verwenden Beschuldigte, bei denen der Tatvorwurf gerechtfertigt ist, häufig Neutralisierungstechniken, um sich vor sich selbst und anderen zu rechtfertigen. Diese Möglichkeit sollte auch den Probanden gegeben sein.
- Dem zu Unrecht Beschuldigten wird die gleiche Beschuldigung entgegengebracht wie dem zu Recht Beschuldigten und er versucht aus eigenen Stücken möglichst unschuldig zu erscheinen, wenn er beschuldigt wird.
- Das Verhalten des zu Unrecht Beschuldigten ist durch starke Besorgnis gekennzeichnet. Auch er möchte aversive Konsequenzen vermeiden, die sich für ihn ergeben können. Möglicherweise entstehen bei zu Unrecht Beschuldigten Zweifel im Hinblick auf den Tatvorwurf, so dass auch dieser sein eigenes Handeln nicht klar bewerten kann.

Eine Durchsicht der Literatur erwies sich in dieser Hinsicht als wenig hilfreich, da ein Großteil der Untersuchungen mit dem sogenannten „Mock Crime“ operiert, was einen eher spielerischen Kontext schafft, und die Szenarien auch thematisch nicht für den intendierten Einsatzbereich relevant sind. Es wurde daher ein eigenes Szenario konzipiert, das die reale Situation hinsichtlich der genannten Aspekte besser simulieren sollte.

Die Verwendung des Oddball-Paradigmas wurde also in der vorliegenden Arbeit eines ökologisch valideren Untersuchungsszenarios angewendet. In bisherigen Studien (Farwell & Donchin, 1991; Allen & Iacono, 1997; Zhou, Yang, Liao & Zou, 1999) wurde meist der Umstand nutzbar gemacht, dass sich ein Täter von einem Nicht-Täter durch sein Tatwissen unterschied (abweichend: Rosenfeld, Angell, Johnson & Quian, 1991). Tatspezifisches Wissen war also nur von zu Recht beschuldigten Probanden als solches überhaupt identifizierbar. In der Anwendung wird es jedoch schwierig sein, entsprechende Items zu finden, die auch zum Zeitpunkt der Hauptverhandlung noch anwendbar sind. In der

vorliegenden Untersuchung variiert ähnlich wie im Kontrollfragentest lediglich der Bedeutungsgehalt der tatspezifischen Items. Das heißt, sowohl der zu Recht als auch der zu Unrecht Beschuldigte haben relevantes Tatwissen.

## **Fragestellung**

Unter Bezugnahme auf die Symbolik von Tabelle 1 soll gelten: Wenn zu Recht beschuldigte Personen, die also Anlass zur Täuschung haben, sich dem Oddball-Paradigma unterziehen und dabei die P300 als ereigniskorreliertes Potential frontal, zentral und parietal abgeleitet wird, soll gelten:  $(S \sim U) > N$ . Analog soll für zu Unrecht beschuldigte Personen gelten:  $(S \sim N) \neq U$ . Beide Hypothesen können damit begründet werden, dass a) Täuschung eine kognitive Leistung ist, die mittels P300 zuverlässig abgebildet werden kann (z. B. Farwell & Donchin, 1991) und dass b) das Oddball-Paradigma den Bedeutungsgehalt der S in Abhängigkeit von der „Täterschaft“ unterschiedlich festlegt.

## **Design**

Es ergeben sich zwei unabhängige Variablen aus der Fragestellung: Die Schuldbedingung als Between-Faktor und der Itemtyp (neutral, tatspezifisch, tatspezifisch) als Within-Faktor. Als abhängige Variable dienen die Amplitudenmaße der P300-Komponenten.

## **Stichprobe**

Die Probanden wurden durch Zeitungsartikel, Aushänge und persönliche Ansprache rekrutiert. Zur Vermeidung geschlechtsspezifischer Effekte wurden ausschließlich Männer in die Stichprobe aufgenommen. Gegenüber den Probanden wurde das Untersuchungsanliegen als Möglichkeit dargestellt, die eigenen detektivischen Fähigkeiten zu trainieren. Als Ausschlusskriterium galt eine klinisch relevante psychische Störung. Darüber hinaus wurde auf die Abwesenheit allgemeinmedizinischer und neurologischer Erkrankungen, sowie akuter und chronischer Schmerzzustände geachtet. Auch die Einnahme von Medikamenten führte zum Ausschluss der Probanden.

An der Untersuchung nahmen 45 Männer im Alter zwischen 20 und 43 Jahren (Mittelwert = 27,16; Standardabweichung = 5,4) teil. Vierzig (89%) der Probanden hatten eine akademische Ausbildung der unterschiedlichsten Fachrichtungen (Jura, Chemie, Agrarwissenschaften, Physik, BWL, Informatik) angefangen oder beendet. Fünf (11%) Probanden hatten eine Ausbildung als Dachdecker, Steuerfachgehilfe, Polizist, Grenzschutzbeamter oder Sozialarbeiter abgeschlossen. Alle Probanden hatten normale oder zu normal korrigierte Sehstärken.

Ein Proband erschien nicht zum zweiten Termin, ein weiterer Proband musste aufgrund der Ausschlusskriterien aus der Stichprobe ausgesondert werden. Es gingen damit Daten von 43 Probanden in die Auswertung ein.

## **Durchführung**

Die Untersuchung wurde für die Probanden als ein Wahrnehmungsexperiment deklariert. Die Probanden sollten lernen, bestimmte Beobachtungsregeln in simulierten Ausnahmesituationen (z.B. Straftaten) zu befolgen. Ihnen wurde mitgeteilt, dass Sie einen emotional unangenehmen

Film sehen würden, auf den sie die gelernten Regeln anwenden sollten. Diese Regeln schrieben unter anderem vor, dass die Aufmerksamkeit nicht auf sogenannte reißerische Aspekte, sondern auf zeugenrelevante Details gelenkt werden sollte. Nachdem die Probanden den Film gesehen hatten, bei dem es sich um die Darstellung einer Sexualstraftat handelte, erhielten Sie einen Feedback-Bogen, der zur eigenen Kontrolle des von ihnen Gelernten diente.

Darin enthalten waren sowohl Items, die für die Zeugenleistung wichtig sein sollten, als auch reißerische Details des Films. Die Schwierigkeit der Fragen war so konstruiert, dass der Untersuchungsleiter jeden Probanden mit Hilfe der Bögen der Bedingung „zu Recht beschuldigt“ oder „zu Unrecht beschuldigt“ zuweisen konnte. Kriterium war dabei die vom Probanden selbst dokumentierte Nichtbeachtung von Wahrnehmungsregeln, die sich darin zeigte, dass er zu sehr auf die „reißerischen“ Details geachtet hatte. Es war also wichtig, dass die Probanden ihr „scoring“ als eine substantielle Untersuchungsbedingung vergegenwärtigten. Denn je nach den Untersuchungsanforderungen erhielten sie anhand ihrer eigenen Festlegung die Rückmeldung, dass sie gut beobachtet und die Regeln gut befolgt hätten oder dass ihre Leistungen schlecht gewesen seien.

Zusätzlich wurden vor der Filmdarbietung Elektroden zur Registrierung eines Elektrokulogramms (EOG) appliziert. Im Anschluss bearbeiteten die Probanden den Suggestibilitätstest von Gudjonsson (1984).

Ein bis zwei Wochen später erfuhren die Probanden, dass ihre Werte in dem Suggestibilitätstest sehr gering gewesen seien – sie also wenig suggestibel seien. Dieser Befund stünde in starkem Kontrast zu den gemessenen Augenbewegungen. Der Untersuchungsleiter konfrontierte den Probanden damit, dass er vermutlich auf die sexuellen, reißerischen Details geachtet habe, statt die Beobachtungsregeln zu befolgen. Das Begehen oder Nichtbegehen der Tat – hier: das Vernachlässigen oder Einhalten der Beobachtungsregeln – war den Probanden durch die Feedback-Bögen rückgemeldet worden. Im Zuge der Ergebnisbesprechung wurde ihm vorgehalten, sich nicht an Beobachtungsregeln gehalten zu haben und sich darüber hinaus auch normabweichend verhalten zu haben. Dadurch wurde bei den Probanden eine kognitive, emotionale und sicherlich auch valutive Dissonanz erzeugt, die zu reduzieren nur sie vermochten.

Um ihnen dabei entgegenzukommen, d.h. den Vorwurf zu entkräften, wurde den Probanden als „Ausweg“ nahegelegt, sich der EEG-Ableitung zu unterziehen, denn damit würde klar, ob der Vorwurf zu Recht bestünde. Dieser letzte Untersuchungsteil eröffnete ihnen sozusagen den Weg zurück in die Konformität. Für die zu Recht beschuldigten Probanden bedeutete dies, dass sie Täuschungsverhalten zeigen und ihre Beeinflussung durch die reißerischen Details während der Filmdarbietung verheimlichen mussten. Die zu Unrecht beschuldigten hingegen waren gefordert, ihre gute Beobachtungsleistung nachzuweisen.

Das gesamte Itemset bestand aus sechs tatspezifischen und sechs tatunspezifischen Items, sowie 24 neutralen Items, die einer gemeinsamen semantischen Kategorie angehörten und durch eine gleiche Silben- und Wortzahl gekennzeichnet waren (siehe Tabelle 2). Alle Items wurden innerhalb eines Blocks viermal dargeboten (144 Darbietungen insgesamt). Es wurden drei Blöcke durchgeführt, zu deren Beginn jeweils die tatunspezifischen Items gelernt wurden. Es ergaben sich also insgesamt 432 Darbietungen. Jedes Item erschien für 300 ms bei einem Interitemintervall von 3000 ms auf dem Bildschirm. Ableitorte waren Fz, Cz und Pz. Der Übergangswiderstand betrug weniger als 10 kΩ.

Tabelle 2: Itemset, geordnet nach den drei Kategorien „tatspezifisch“, „tatunspezifisch“ und „neutral“

Tatspezifisch	Tatunspezifisch	Neutral
Kurzer Mini	Langer Mantel	Enge Hotpants Lange Strümpfe Knappes Höschen Push-up BH

Bauchkette	Ohrstecker	Fußkette Intimschmuck Bauchpiercing Sonnenhut
über Zaun hängen	in Auto steigen	An Brücke lehnen Im Wasser liegen Unter Deck gehen Durch Scheibe fliegen
Blonde Haare	Schlechte Zähne	Volle Wangen Schwarze Locken Brauner Rücken Kleiner Busen
Biß in Busen	Tritt in Hoden	Kniff in Hintern Hieb in Gesicht Stich in Schenkel Kick in Gesicht
Geile Pussi	Dumme Nutte	Schnieke Lady Scharfe Ische Super Schmitte Heißer Feger

Im Anschluss wurden die Probanden ausführlich über das Anliegen der Untersuchung aufgeklärt und das Vorgehen mit ihnen diskutiert.

### Auswertung

Die gefilterten EEG-Epochen wurden bezüglich der Augenbewegungen korrigiert. Um trotz unbekannter Verteilung der Amplitudenmaße einen statistischen Wert für die Klassifikation der Probanden bezüglich der Schuldbedingung zu erhalten, wurden die Amplitudenmaxima mit Hilfe eines Bootstrapverfahrens verglichen (Allen & Iacono, 1997). Dieses Vorgehen lässt sich in vier Schritten darstellen: 1. Mitteln der neutralen Items: Für jeden Probanden resultierten insgesamt 432 Epochen, davon je 72 Epochen vom Testreiztyp (tatspezifisch) und Zielreiztyp (tatunspezifisch), sowie 288 vom neutralen Typ. Jeweils vier Epochen des irrelevanten Typs wurden zu einer gemittelt, so dass von jedem Typ schließlich 72 Epochen vorhanden waren. 2. Ziehen und Mitteln der Kurven: Aus jedem Pool (tatspezifisch, tatunspezifisch, neutral) wurden pro Iteration 72 Epochen gezogen, wobei nach jedem Ziehen die ausgewählte Epoche zurück in den Pool gegeben wurde. Die 72 Epochen wurden jeweils gemittelt, so dass pro Iteration drei gemittelte Kurven, eine für jede Bedingung, gewonnen wurden. 3. Vergleich der gemittelten Kurven: Bei jeder Iteration wurde aus dem Zeitfenster von 350 – 1000 ms das Maximum bestimmt und in einem Zeitfenster von 100 ms um diesen Datenpunkt gemittelt. Die so entstandenen gemittelten Maxima wurden verglichen. 4. Berechnung des Bootstrapindex pro Person: Der Bootstrapindex ergab sich aus dem Vergleich der Kurvenmaxima bezüglich N und S über 100 Iterationen. Bei jeder Iteration, bei der das Maximum bezüglich S das bezüglich N übertraf, wurde eine 0 addiert, ansonsten eine 1. Der Bootstrapindex konnte also zwischen 0 und 100 variieren und gibt an, wie häufig S N nicht übertrifft. Das heißt, ein niedriger Index spricht für die Klassifikation eines Probanden als „zu Recht beschuldigt“ und ein hoher für eine Klassifikation als „zu Unrecht beschuldigt“.

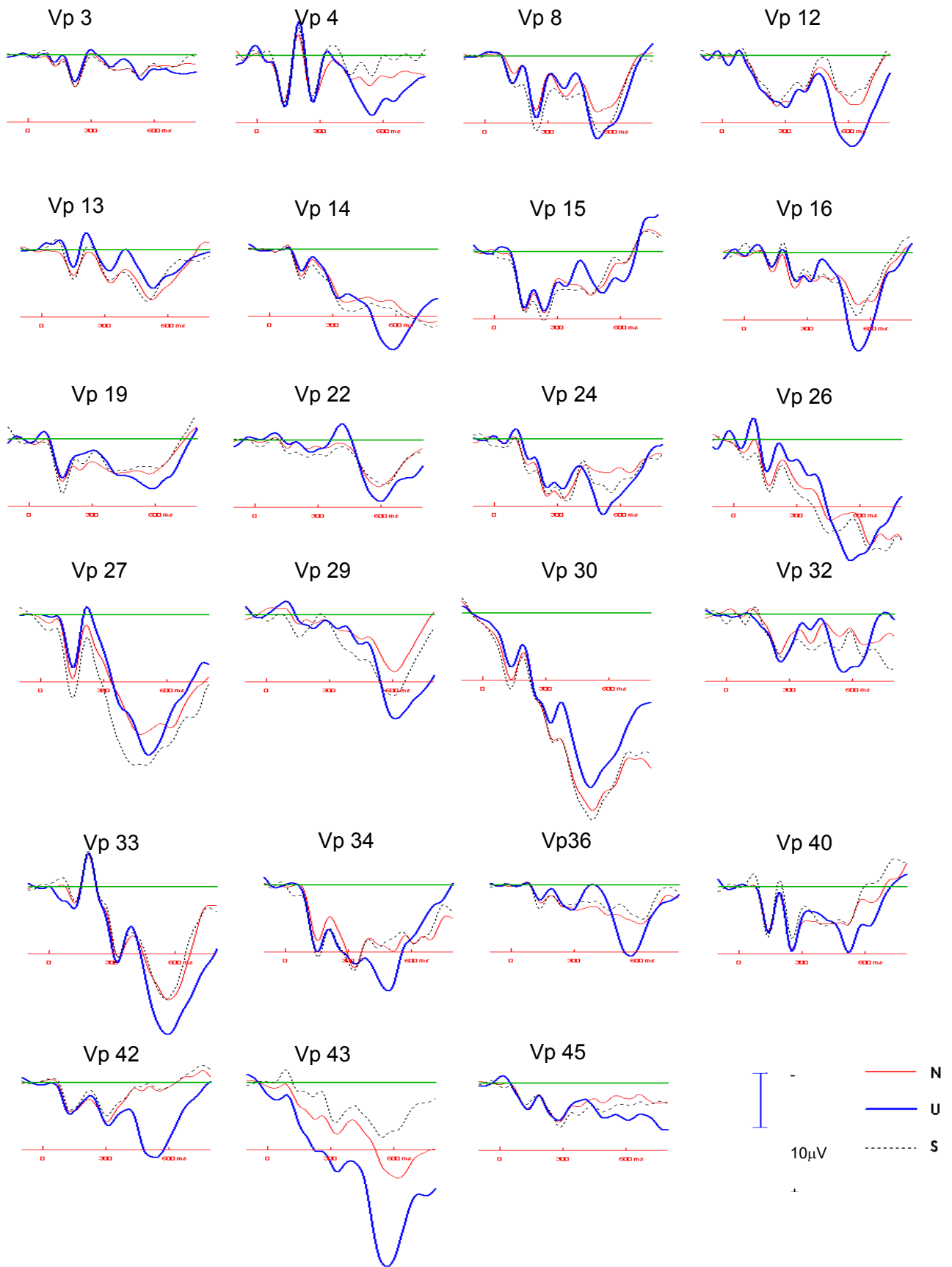


Abbildung 2: Ereigniskorrelierte Potentiale für 20 ~~zu~~ nicht beschuldigte Probanden



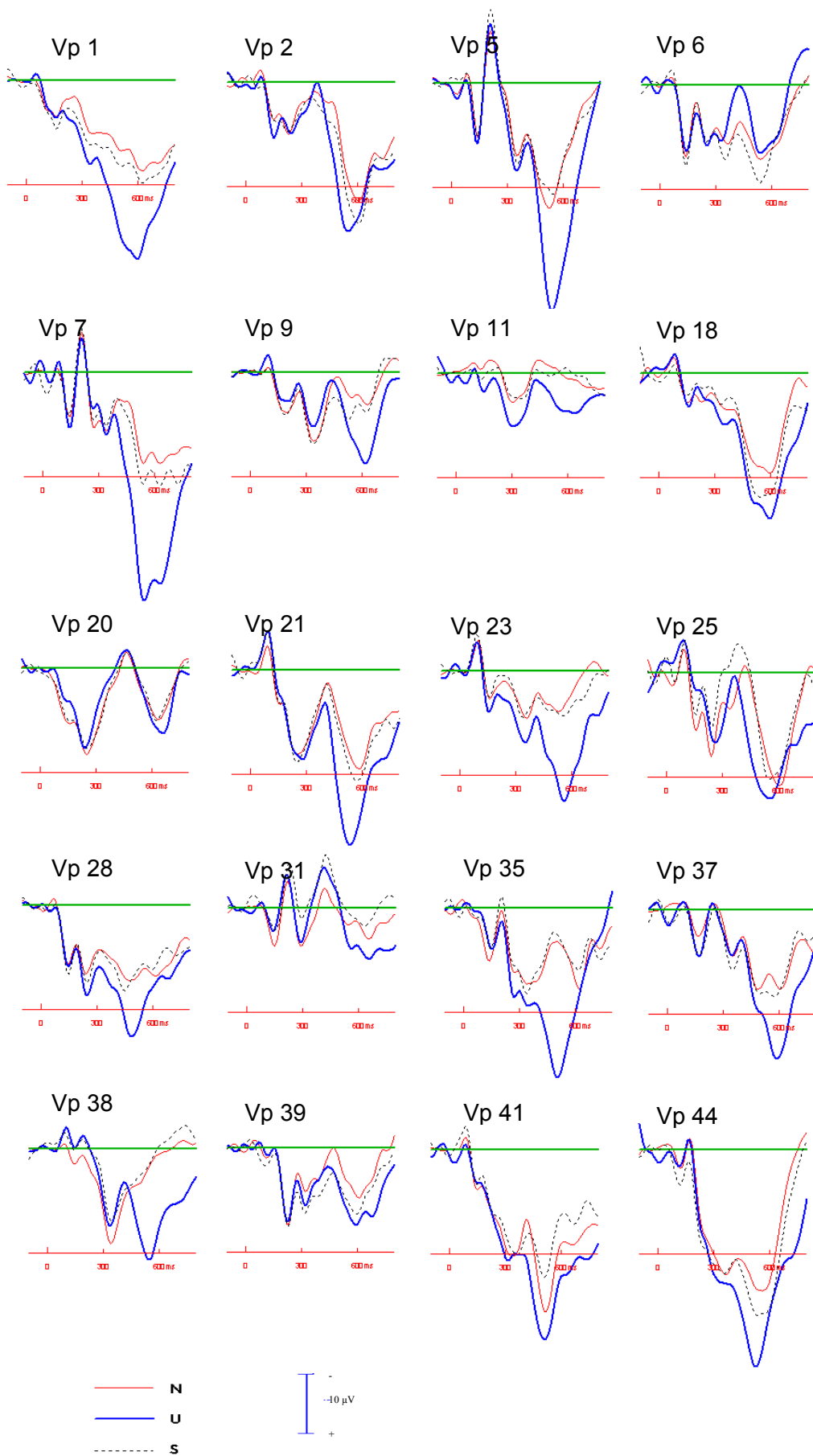


Abbildung 3: Ereigniskorrelierte Potentiale für 20 zu Unrecht beschuldigte Probanden  
219

## Ergebnisse

Die hier dargestellten Ergebnisse beziehen sich auf die Ableitung von Pz. Die gemittelten und bezüglich der Augenbewegung korrigierten Potentiale sind in Abbildung 2 und 3 wiedergegeben. Setzt man für die Bootstrapindizes Grenzwerte bei 10 und 70 fest (Farwell & Donchin, 1991), erhält man 28% korrekte Entscheidungen, 12% falsch negative, 19% falsch positive und 42% unentscheidbare Fälle (Tabelle 3). Anhand der entsprechenden Receiver Operating Characteristic (ROC) - Kurve lässt sich erkennen, dass die Methode nicht zwischen den Bedingungen differenziert ( $\text{Area under Curve (AUC)} \pm \text{SE} = .538 \pm .096$ ).

Tabelle 3: Klassifizierung der Probanden anhand der Bootstrapindizes (B). Grenzwerte bei 10 und 70. Grau unterlegte Felder beinhalten hypothesenkonforme Fälle

	(S ~ U) > N bzw. B < 10	(S ~ N) ≠ U bzw. B > 70
Zu Recht beschuldigt	4	5
Zu Unrecht beschuldigt	8	8

## Diskussion

Es zeigt sich, dass unter den ökologisch valideren Bedingungen eine eindeutige Klassifikation nicht möglich ist. Verschiedene Aspekte mindern die Wahrscheinlichkeit als zu Recht beschuldigt klassifiziert zu werden. Da die Probanden die Möglichkeit hatten, Neutralisierungstechniken anzuwenden und dies offensichtlich auch taten, was Bemerkungen im Nachgespräch zu entnehmen ist, besteht die Möglichkeit, dass diese Probanden sich nicht zu Recht beschuldigt fühlten und dass dies zu falsch negativen oder unentscheidbaren Ergebnissen führte. Da die tatspezifischen Items nicht auswendig gelernt wurden und der Zeitraum zwischen den Terminen im Vergleich zu bisherigen Untersuchungen lang war, bestand die Möglichkeit, dass Probanden Items vergaßen. Obwohl sich diesbezüglich kein Zusammenhang zu den Ergebnissen zeigte und alle Probanden wenigstens 4 der 6 relevanten Items in einem Rekognitionstest erinnerten, mindert jedes vergessene tatspezifische Item die Wahrscheinlichkeit als zu Recht beschuldigt klassifiziert zu werden. Andererseits sorgt möglicherweise die sexuelle Thematik und mit ihr auch die Formulierung der Items für unerwünschte Aufmerksamkeitsprozesse bei zu Unrecht beschuldigten Probanden und begünstigt damit falsch-positive Ergebnisse. Da diese Aspekte auch für die Anwendung der Methode in der Praxis relevant sind, liegt eine fruchtbare Fortführung dieses Forschungsansatzes sicherlich in einer valideren Kontrolle dieser Faktoren und ihrer Wirkung auf die Bedeutsamkeit der Reizworte.

## Literatur

- Allen, J. J. B. & Iacono, W. G. (1997). A comparison of methods for the analysis of event-related potentials in deception detection. *Psychophysiology*, 34, 234-240.
- Deegener, G. (1995). *Sexueller Mißbrauch: Die Täter*. Weinheim: Psychologie Verlags Union.
- Farwell, L. A. & Donchin, E. (1991). The truth will out: Interrogative Polygraphy („Lie Detection“) with event-related brain potentials. *Psychophysiology*, 28, 531-547.
- Gudjonsson, G. H. (1984). A new scale of interrogative suggestibility. *Journal of Personality and Individual Differences*, 5, 303-314.
- Rosenfeld, J. P., Angell, A., Johnson, M. & Qian, J.-H. (1991): An ERP-based, control-question lie detector analog: Algorithms for discriminating effects within individuals' average waveforms. *Psychophysiology*, 28, 319-335.

Scholz, O. B. (1999): Die physiologischen Parameter in der psychophysiologischen Aussagebegutachtung (PAG). In U. Willutzki, J. Salzgeber & M. Stadler (Hrsg.), *Polygraphie. Möglichkeiten und Grenzen der PAG* (S. 45-67). Köln: Bundesanzeiger-Verlag.

Zhou, L., Yang, W., Liao, S. & Zou, H. (1999). Experimental study of lie detection with P300 in simulated crime. *Chinese Journal of Clinical Psychology*, 7, 31-33.



# Vermutungen von Juristen und Rechtslaien zur Wirkung von Gerichtsfernsehen<sup>1</sup>

*Edgar W. Klinger & Heiner Gehring*

*Fachbereich Psychologie und Gesundheitswissenschaften der Universität  
Osnabrück*

Das Für und Wider einer Fernsehberichterstattung aus Gerichtsverhandlungen wird seit etwa 50 Jahren in rechtswissenschaftlichen Fachpublikationen im deutschen Sprachraum diskutiert. Die Entscheidung des Bundesverfassungsgerichts zur Zulassung von Gerichtsfernsehen hat zu Beginn dieses Jahres die Diskussion neu entfacht und in die Tagespresse getragen. In der Diskussion überwiegen Stellungnahmen von Juristen, in denen das Gerichtsfernsehen abgelehnt wird. Begründet wird diese Position zumeist mit der Vermutung, dass das Wissen, gefilmt zu werden, das Verhalten der Beteiligten im Gerichtssaal beeinflusse und damit die Wahrheitsfindung erschwere. Die Gerichtsverhandlung verkomme zu einer Show-Veranstaltung, in der Beschuldigte und Zeugen sich gehemmt fühlen müssten, über Einzelheiten ihres Privatlebens zu berichten, und in der sich die Beschuldigten an den Pranger gestellt sähen. Diese Vermutungen klingen plausibel, sind indes nicht durch empirische Studien belegt. In welchem Maße geben nun diese öffentlich vorgetragenen Stellungnahmen das Meinungsbild von Richtern, Rechtsanwälten und Rechtslaien zu Wirkungen von Gerichtsfernsehen wieder? Der Beitrag stellt die Ergebnisse von Umfragen vor, in denen nach solchen Einschätzungen gefragt wurde. Sie zeigen, dass Juristen negative Effekte eines Gerichtsfernsehens in stärkerem Maße vermuten als Rechtslaien. Ferner erwarten die befragten Juristen deutlichere Auswirkungen auf das Verhalten anderer Verhandlungsbeteiligter als auf das eigene Verhalten. Die Einschätzungen deutscher Juristen werden jenen amerikanischer Juristen gegenübergestellt, die seit etwa zehn Jahren Erfahrungen mit der Arbeit des Gerichtsfernsehensenders „Court-TV“ sammeln konnten. Abschließend werden Forschungsperspektiven skizziert, deren Umsetzung zur Gewinnung gesicherter Erkenntnisse über die Wirkungen von Gerichtsfernsehen beitragen kann.

## **1 Einführung**

Der Begriff Gerichtsfernsehen steht im deutschen rechtswissenschaftlichen Sprachgebrauch für Fernsehübertragungen und –aufzeichnungen von Teilen oder der Gesamtheit einer Gerichtsverhandlung (Braun, 1998). In Deutschland steht dem Gerichtsfernsehen der im Jahre 1964 eingeführte Wortlaut des § 169 Satz 2 des Gerichtsverfassungsgesetzes (GVG) entgegen, in dem es heißt: „Ton- und Fernseh-Rundfunkaufnahmen sowie Ton- und Filmaufnahmen zum Zwecke der öffentlichen Vorführung oder Veröffentlichung ihres Inhalts sind unzulässig.“ Das Bundesverfassungsgericht hat jüngst die Verfassungsmäßigkeit dieser Vorschrift bestätigt und eine Fernsehberichterstattung aus laufenden Prozessen mit Hinweis auf die Gefahr der Beeinträchtigung sowohl des Persönlichkeitsschutzes der Prozeßbeteiligten als auch der Wahrheitsfindung des Gerichts untersagt (Braun, 1998). Zulässig sind indes infolge einer Änderung des Bundesverfassungsgerichtsgesetzes (BverfGG) vom 16.07.1998 (BGBl. I, 1823) Ton- und Filmaufnahmen von Entscheidungsverkündungen des Bundesverfassungsgerichts. Damit ist erstmals in der Geschichte der deutschen Gesetzgebung die Möglichkeit gesetzlich verankert, einen Teil einer Verhandlung vor dem BverfG zum Zweck der Aussendung im Fernsehen zu filmen. Überdies bestehen Überlegungen, auch die Direktübertragung der Verhandlung samt Urteilsbegründung zuzulassen (Braun, 1998). Seit Jahrzehnten wird in der deutschen rechtswissenschaftlichen Fachliteratur eine Debatte um das Für und Wider des Gerichtsfernsehens ausgetragen (Braun, 1998), die während des letzten Jahrzehnts durch spektakuläre Einzelfälle von Fernsehgerichtsberichterstattung in den USA und nicht zuletzt durch die Ablehnung einer Verfassungsbeschwerde des Fernsehensenders n-tv durch das BverfG neu entfacht wurde. Nur vereinzelt finden sich in dieser Debatte Stimmen, die auf die Möglichkeit der Aufklärung über Recht und Gesetz mittels

Gerichtsfernsehen hinweisen und dadurch positive Effekte auf die öffentliche Anerkennung des Rechts erwarten (z.B. Gündisch & Dany, 1999). So spricht Merten (1997) den Medien eine Mittlerrolle bei der Vermittlung von Rechtskenntnis zu, auf die das Rechtssystem achten sollte. Immerhin ist zu erwarten, daß umfangreiche Rechtskenntnisse der Legitimation des Rechtssystems förderlich sind. Da nach Ansicht vieler Juristen die Gerichtswirklichkeit in den Printmedien jedoch oft unangemessen wiedergegeben wird (Eberle, 1996), stellt sich auch die Frage, ob im Medienzeitalter die Gerichtsöffentlichkeit auf die gelegentlich anwesenden Zuschauer im Gerichtssaal beschränkt bleiben kann und auch künftig kein Filmmaterial aus laufenden Gerichtsverhandlungen gezeigt werden darf (Walther, 1998). Demgegenüber herrscht in der juristischen Diskussion die u.a. von Kortz (1997) zusammengefaßte Meinung vor, Gerichtsfernsehen werde insbesondere die Wahrheitsfindung und die Unabhängigkeit der Rechtsprechung beeinträchtigen. Dieser Ansicht folgt auch der Deutsche Richterbund (1996), der im übrigen Beeinflussungen des Verhaltens der Prozeßbeteiligten befürchtet und davor warnt, im Fernsehen gezeigte Gerichtsprozesse könnten zu Unterhaltungssendungen verkommen, in denen Verurteilte an den Pranger gestellt werden.

Die in der deutschen Kontroverse um das Gerichtsfernsehen behaupteten psychischen Wirkungen beruhen weder auf eigenen Erfahrungen noch auf Ergebnissen einschlägiger empirischer Studien – beides liegt nicht vor. Die in der juristischen Debatte ausgetauschten Argumente gründen auf Wirkungsvermutungen der einzelnen Verfasser. Es stellt sich die Frage, ob diese Vermutungen von einer größeren Anzahl von Juristen geteilt werden. Zudem ist zu fragen, ob Juristen aus unterschiedlichen Tätigkeitsbereichen übereinstimmende Vermutungen über die Wirkungen von Gerichtsfernsehen hegen. Schließlich stellt sich die Frage, welche Wirkungen Rechtslaien infolge einer Einführung von Gerichtsfernsehen vermuten.

In den USA gehören Erfahrungen mit Gerichtsfernsehen seit über 20 Jahren zum Gerichtsalltag (Cohn & Dow, 1998; Machura, 1998): Fernsehaufnahmen sind in 48 Bundesstaaten zugelassen. Seit Juli 1991 berichtet der Privatsender Courtroom Television Network („Court TV“) mit einem 24-Stunden-Programm an sieben Tagen der Woche teils per Direktübertragung, teils per Aufzeichnung aus Gerichtsverhandlungen. Das Filmmaterial wird durch Informationen über die rechtlichen Hintergründe der verhandelten Fälle sowie durch Interviews mit Richtern, Anwälten und anderen Prozeßbeteiligten ergänzt (Braun, 1998; Thaler, 1994). Die Erfahrungen mit diesem US-Gerichtsfernsehen wurden in mehreren Umfragen unter Juristen und Rechtslaien dokumentiert. Im folgenden Abschnitt werden Ergebnisse dieser Umfragen präsentiert. Sie bilden den Ausgangspunkt für Befragungen deutscher Juristen und Rechtslaien, deren Ergebnisse im Anschluss vorgestellt werden.

## **2 US-Erfahrungen mit Gerichtsfernsehen**

In den USA werden seit den 70er Jahren Teilnehmer einzelner Gerichtsverhandlungen, die zwecks Ausstrahlung im Fernsehprogramm gefilmt wurden, von verschiedenen Institutionen nach ihren persönlichen Erfahrungen in dieser Situation befragt. Eine von Barber (1987) vorgenommene Auswertung der Ergebnisse von 19 solcher Befragungen zeigt, daß die Mehrheit der Befragten kaum unerwünschte Wirkungen der Fernsehaufnahmen auf das Verhalten der Prozeßbeteiligten sowie auf Verlauf und Ergebnis der Verhandlungen feststellen konnten. So äußerte jeweils eine Mehrheit der befragten Richter, Anwälte, Zeugen und Angeklagten die Ansicht, durch die Filmaufnahmen werde weder die Würde des Gerichts beeinträchtigt noch werde das eigene Verhalten oder das der übrigen Prozeßbeteiligten in beachtenswertem Umfang beeinflußt. Indes wurde gleichermaßen einheitlich befürchtet, daß (1) die Wichtigkeit einzelner Prozesse durch eine Fernsehberichterstattung überzogen dargestellt werden könnte, (2) das Sicherheitsempfinden von Zeugen beeinträchtigt werden

könnte und (3) die Geschworenen, die im Fernsehen nicht zu sehen sein dürfen, durch die Filmaufnahmen beunruhigt und deshalb in ihrer Wahrheitsfindung behindert werden könnten. Relativ häufig wiesen die befragten Anwälte und Gerichtsbediensteten auf Verhaltensänderungen bei Richtern, Zeugen und auch Anwälten hin, doch seien diese Verhaltensänderungen in ihren Konsequenzen zu vernachlässigen. Seitens der Richter und der Geschworenen wurden besonders selten Auswirkungen der Filmaufnahmen auf den Prozeßverlauf und das Verhalten der Beteiligten erwähnt. Angesichts dieser Ergebnisse warnt Barber (1987) davor, erstens einzelne spektakuläre Fälle von Fernsehgerichtsberichterstattung in der Kontroverse um das Gerichtsfernsehen zwecks Illustration ihrer mutmaßlichen negativen Auswirkungen zu verwenden und zweitens die guten Erfahrungen vieler Prozeßbeteiligter in der Kontroverse nicht zu vergessen. Gleichwohl werden gerade in der in den USA geführten Diskussion um Gerichtsfernsehen spektakuläre Einzelfälle immer wieder als Negativbeispiele in den Vordergrund gerückt (z.B. Thaler, 1994; Westfeldt & Wicker, 1998).

Auch von US-Justizbehörden wurden Prozeßbeteiligte zu ihren Erfahrungen mit dem Gerichtsfernsehen befragt. In der neuesten dieser Untersuchungen erhob das New York State Committee (1997) zehn Jahre nach der Zulassung von Gerichtsfernsehen in diesem Bundesstaat die Erfahrungen mit und Einschätzungen über den Einfluß der Fernsehaufnahmen auf Verlauf und Ergebnis der Verhandlungen von insgesamt 350 Juristen, ferner von 50 als Zeugen in Gerichtsprozessen erschienenen Personen, von den Mitgliedern zweier Jurys sowie von 50 mit der Thematik vertrauten Fachleuten. Alle Befragten hatten bereits an zum Zweck der Fernsehberichterstattung gefilmten Verhandlungen teilgenommen. Die Hälfte aller Befragten (52%) gab an, Unterschiede in der Prozeßführung der Richter zwischen gefilmten und nicht gefilmten Verhandlungen wahrgenommen zu haben. Jeweils etwa 40% der Befragten sahen die Geschworenen bei ihrer Wahrheitsfindung behindert und vermuteten eine Erhöhung der Nervosität der Zeugen durch Filmaufnahmen. Etwa ein Drittel der Befragten erlebte die Anwälte in gefilmten Verhandlungen besser vorbereitet. Trotz dieser Ergebnisse gelangte das New York State Committee (1997) zu der Auffassung, daß es keine Belege für Beeinträchtigungen der Prozeßfairneß oder für Behinderungen der Urteilsfindung durch Filmaufnahmen gebe. Vielmehr sei durch das Gerichtsfernsehen die Förderung der Akzeptanz des Rechtssystems und ein verbessertes öffentliches Verständnis des Rechts zu erwarten.

Thaler (1994) führte im Verlaufe eines Sensationsmordprozesses, der 1989 in New York stattfand, zahlreiche Interviews mit den Prozeßbeteiligten über die empfundenen Auswirkungen der Filmaufnahmen. Seine Auswertungen zeigen generell eine hohe positive Korrelation zwischen den Einstellungen zum Medium Fernsehen und den Einstellungen zum Gerichtsfernsehen: Wer viel Fernsehen schaut, hat eine positivere Einstellung zum Gerichtsfernsehen. Die befragten Strafverfolger wiesen auf den hohen psychischen Druck hin, der infolge Fernsehberichterstattung vermutlich auf den Zeugen laste, ihr Verhalten verändere und so die Wahrheitsfindung erschwere. Einige der Zeugen bestätigten diesen Verdacht. Der Richter indes befürwortete die Fernsehübertragungen mit dem Hinweis, die Beteiligten dieses gefilmten Prozesses würden ihr bestes Benehmen zeigen (vgl. auch Braun, 1998).

### **3 Vermutungen deutscher Juristen und Rechtslaien über Wirkungen von Gerichtsfernsehen**

#### **3.1 Überblick**

Wie schätzen deutsche Juristen und Rechtslaien die Wirkungen von Gerichtsfernsehen auf das Verhalten der Prozeßbeteiligten ein? In zwei Befragungen wurden diese Einschätzungen mit

Blick auf das Verfahrensrecht, das den Verlauf einer Verhandlung regelt, und das materielle Recht, das die Beurteilung des verhandelten Sachverhaltes regelt, erhoben. Die Ergebnisse sollen Auskunft darüber geben, ob und wie Gerichtsfernsehen die wahrgenommene prozedurale und distributive Gerechtigkeit von Gerichtsprozessen beeinflusst (Lind & Tyler, 1988; Bierbrauer & Klinger, im Druck). Vor diesem Hintergrund soll auch die Frage beantwortet werden, wie verbreitet die in der juristischen Kontroverse dominierende Meinung unter deutschen Juristen und Rechtslaien ist, wonach infolge einer Zulassung von Gerichtsfernsehen negative Wirkungen auf Verlauf und Ergebnis von Gerichtsprozessen zu erwarten seien.

Ein weiterer Aspekt soll anhand der Befragung unter Juristen untersucht werden: Vermuten jene Personen, die Fernsehaufnahmen aus Gerichtsverhandlungen ablehnen, Auswirkungen dieser Aufnahmen eher auf das Verhalten anderer Prozeßbeteiligter als auf das eigene Verhalten? Alltagspsychologische Überlegungen könnten Personen zu der Einschätzung verleiten, das eigene Verhalten sei durch Fernsehaufnahmen nicht oder in nur geringem Ausmaß beeinflussbar, wohingegen das Verhalten anderer Personen sehr wohl durch Fernsehaufnahmen beeinflusst werden könne. Hier läßt sich eine Parallele zu dem in einer Vielzahl empirischer Studien nachgewiesenen Phänomen des *Third Person Effect* (Gibbon & Durkin, 1995; Gunther, 1991; Innes & Zeitz, 1988) erkennen: Viele Menschen sind der Ansicht, hauptsächlich andere Personen würden in ihrer Weltsicht durch Medien beeinflusst; die eigene Weltsicht dagegen sei gegenüber solchen verzerrenden Einflüssen gefeit. Dieses Phänomen beschreibt allein die vermuteten, nicht aber die tatsächlichen Unterschiede in der Beeinflussbarkeit individuellen Verhaltens (Brosius & Engel, 1997). Diese Variante verzerrter Alltagswahrnehmung entsteht durch die zumeist irriige Annahme, man selbst sei gegenüber äußeren Einflüssen in seiner Wahrnehmung und Beurteilung besser geschützt als andere Personen. Den eigenen Fähigkeiten wird damit eine ungerechtfertigte Einzigartigkeit („false uniqueness“; Goethals, Messick & Allison, 1991) zugeschrieben, die die Einschätzung der Beeinflussbarkeit anderer Personen färbt. Duck und Mullin (1995, 79) sprechen in diesem Zusammenhang von der „Illusion einzigartiger Unverletzbarkeit“. Innes und Zeitz (1988) begründen vor diesem Hintergrund ein dem Third Person Effect entsprechendes Verhalten balancetheoretisch. Gunther (1991) stellt Bezüge zwischen dem Effekt und dem fundamentalen Attributionsfehler (Ross & Nisbett, 1991) her. Danach mißachtet die Einschätzung des Medieneinflusses auf Dritte die Bedeutung situativer Faktoren, wogegen die Einschätzung der Beeinflussung des eigenen Verhaltens den situativen Kontext ausführlich berücksichtigt.

Mutmaßliche Quelle einer Beeinflussung des Verhaltens anderer Personen, nicht aber des eigenen Verhaltens, ist im Rahmen dieser Untersuchung die Präsenz einer Fernsehkamera im Gerichtssaal, also eine Veränderung der Rahmenbedingungen, unter denen eine Gerichtsverhandlung abläuft. Hier wird ein Third Person Effect erwartet, der nicht auf die Wirkungen des Ergebnisses der Fernsehberichterstattung, also der entstandenen Fernsehsendung abstellt, sondern auf deren Produktion im Gerichtssaal. Die Übertragbarkeit des Effekts auf den hier beleuchteten Kontext erscheint gerechtfertigt: Hier wie dort geht es um die Einschätzung der Beeinflussbarkeit des Verhaltens Dritter im Unterschied zum eigenen Verhalten (Gunther, 1991; Innes & Zeitz, 1988).

Der Third Person Effect kann dazu führen, daß eine Person Dritte vor den vermuteten Gefahren einer Beeinflussung von Einstellungen und Verhalten bewahren möchte (Innes & Zeitz, 1988). Tatsächlich könnte für die Gegner des Gerichtsfernsehens gelten, daß sie Prozeßbeteiligte vor den befürchteten Wirkungen der Filmaufnahmen schützen möchten (Gehring, 1998). Dies mag insbesondere für die häufig oder regelmäßig an Gerichtsverfahren beteiligten Gruppen der Richter und Rechtsanwälte gelten, denen ein besonders hohes Involvement in der Diskussion um die Einführung von Gerichtsfernsehen unterstellt werden kann (Gunther, 1991; vgl. auch Perloff, 1989).



### 3.2 Methodisches Vorgehen

*Stichproben.* Die Befragung der Juristen wurde im ersten Halbjahr 1998 durchgeführt. Im einzelnen wurden vier Gruppen befragt: (1) an einem Seminar der Richterakademie in Trier teilnehmende Verwaltungsrichter, (2) die Richter des Amtsgerichtsbezirks Osnabrück, (3) per Zufallsverfahren ausgewählte, im Amtsgerichtsbezirk Osnabrück zugelassene Rechtsanwälte und (4) alle Professoren und wissenschaftlichen Mitarbeiter des Fachbereichs Rechtswissenschaften der Universität Osnabrück. Insgesamt gingen 115 vollständig ausgefüllte Fragebögen (Rücklaufquote 31,5%) in die Auswertung ein, darunter 20 von Verwaltungsrichtern (davon 5 weiblich; Durchschnittsalter 39,6 Jahre), 15 von Richtern des Amtsgerichtsbezirks Osnabrück (davon 5 weiblich; Durchschnittsalter 46,0 Jahre), 60 von Rechtsanwälten aus dem Amtsgerichtsbezirk Osnabrück (davon 11 weiblich; Durchschnittsalter 46,8 Jahre) und 20 von Professoren und wissenschaftlichen Mitarbeitern des rechtswissenschaftlichen Fachbereichs der Universität Osnabrück (davon 3 weiblich; Durchschnittsalter 48,0 Jahre). Das Durchschnittsalter der Gesamtstichprobe beträgt 45,3 Jahre.

Die Befragung der Rechtslaien wurde im zweiten Halbjahr 1999 durchgeführt. Die Auskunftspersonen wurden aus den Kunden eines Osnabrücker Einkaufszentrums anhand grober Quotenvorgaben hinsichtlich Geschlecht und Alter über einen Zeitraum von zwei Wochen ausgewählt. Die fragebogengestützten Befragungen dauerten durchschnittlich ca. 10 Minuten. Insgesamt nahmen an der Befragung 226 Personen teil (92 Frauen, 134 Männer). Das Durchschnittsalter dieser Stichprobe beträgt 40,7 Jahre.

*Fragebögen.* Der den Juristen vorgelegte Fragebogen enthält zehn Aussagen zu möglichen Auswirkungen von Gerichtsfernsehen (siehe Anhang A). Die Reaktionen der Befragten auf diese Aussagen wurde anhand vierstufiger Antwortskalen erfasst. Die Aussagen fragten nach der grundsätzlichen Zustimmung bzw. Ablehnung von Gerichtsfernsehen, nach vermuteten Auswirkungen von Gerichtsfernsehen auf Verlauf und Ergebnis einer Verhandlung und nach vermuteten Auswirkungen der Fernsehaufnahmen auf das Verhalten der befragten Person in ihrer Tätigkeit als Richter bzw. Rechtsanwalt. Zudem wurden Richtern Aussagen über mögliche Einflüsse der Fernsehaufnahmen auf das Verhalten der Rechtsanwälte und umgekehrt den Rechtsanwälten entsprechende Aussagen zu möglichen Verhaltensbeeinflussungen bei Richtern vorgelegt. Die den Universitätsangehörigen vorgelegten Fragebögen enthalten Aussagen über Verhaltensreaktionen sowohl der Richter als auch der Rechtsanwälte. Schließlich wurden Alter und Geschlecht des Befragten, der durchschnittliche tägliche Fernsehkonsum und die gegenwärtige juristische Tätigkeit erhoben. Der den Rechtslaien vorgelegte Fragebogen enthält ebenfalls zehn Aussagen zu möglichen Wirkungen von Gerichtsfernsehen (siehe Anhang B). Die Aussagen beziehen sich auf mögliche Auswirkungen auf Verlauf und Ergebnis einer Verhandlung und mögliche Verhaltensänderungen bei Richtern bzw. Rechtsanwälten. Der Wortlaut von fünf dieser Aussagen stimmt mit den entsprechenden Aussagen des zur Befragung der Juristen verwendeten Fragebogens überein. Zwei Aussagen beziehen sich auf die vermuteten Konsequenzen des Gerichtsfernsehens für den Angeklagten. Zudem wurden die Befragten nach ihrer Bereitschaft gefragt, als Zeugin/Zeuge oder Angeklagte(r) an einem zum Zweck der Fernsehberichterstattung gefilmten Gerichtsverfahren teilzunehmen. Die Reaktionen der Befragten wurden anhand vierstufiger Antwortskalen erfasst. Zudem wurden soziodemographische Merkmale (Alter, Geschlecht, Schulabschluß und Berufsbildung, derzeitige berufliche Berufstätigkeit), der durchschnittliche tägliche Fernsehkonsum und juristische Vorkenntnisse erhoben. Abschließend wurde gefragt, ob die Auskunftsperson schon einmal als Beteiligter oder als Zeuge an einer Gerichtsverhandlung teilgenommen hat.

## 4 Ergebnisse

Auskunftspersonen mit unterschiedlichen soziodemographischen Merkmalen oder unterschiedlichem durchschnittlichen täglichen Fernsehkonsum unterscheiden sich hinsichtlich der vermuteten Wirkungen von Gerichtsfernsehen nicht signifikant voneinander. Dies gilt für die Gruppe der Juristen ebenso wie für die Gruppe der Rechtslaien. Auf Differenzierungen der beiden Stichproben entlang dieser Merkmale wird bei der Darstellung der Ergebnisse daher verzichtet.

*Einstellungen zum Gerichtsfernsehen.* Von den befragten Juristen wird das Gerichtsfernsehen deutlich abgelehnt. Im einzelnen waren alle Verwaltungsrichter, 95% der Universitätsangehörigen, 86,7% der Amtsrichter und 81,7% der Rechtsanwälte der Ansicht, daß Gerichtsverhandlungen grundsätzlich nicht im Fernsehen zu sehen sein sollten. Zudem lehnten es 70% der Verwaltungsrichter, 66,7% der Amtsrichter und 51,7% der Rechtsanwälte ab, in Ausübung ihrer beruflichen Tätigkeit im Fernsehen gesehen werden zu können. Dagegen deckt die Befragung ein beträchtliches Interesse der Rechtslaien am Gerichtsfernsehen auf: 51,8% von ihnen würden die Zulassung von Gerichtsfernsehen begrüßen und sich im Fernsehen Berichterstattungen aus Gerichtsverhandlungen ansehen, würden sie gesendet werden. Trotz dieses Interesses am Gerichtsfernsehen zeigten sich nur wenige Befragte bereit, in einer gefilmten Gerichtsverhandlung als Zeuge (29,2%) oder als Angeklagter (16,0%) gesehen zu werden.

*Vermutete Auswirkungen des Gerichtsfernsehens.* Tabelle 1 zeigt die Mittelwerte und Standardabweichungen der Reaktionen der befragten Juristen auf die sieben Aussagen zu vermuteten Wirkungen des Gerichtsfernsehens auf Ablauf und Ergebnisse einer Gerichtsverhandlung.

Tabelle 1: Von Juristen vermutete Wirkungen von Gerichtsfernsehen auf Ablauf und Ergebnis von Gerichtsverhandlungen

	Verwaltungsrichter n=20	Amtsrichter n=15	Rechtsanwälte n=60	Universitätsangehörige n=20
<u>Ablauf der Verhandlung</u>				
Ablauf wird behindert	3.40 (.82)	3.40 (.74)	3.25 <sup>a</sup> (.99)	3.70 <sup>a</sup> (.47)
Urteilsfindung wird behindert	2.45 <sup>b</sup> (.89)	2.67 (.98)	3.02 (1.00)	3.30 <sup>b</sup> (.86)
Mediengerechtes Verhalten von Rechtsanwälten	3.65 (.49)	3.80 (.41)	-----	3.75 (.44)
Mediengerechtes Verhalten von Richtern	-----	-----	3.17 (.69)	3.20 (.62)
<u>Ergebnis der Verhandlung</u>				
Höhere Strafen	2.30 (1.03)	2.40 (.74)	2.63 (.94)	2.60 (.82)
Urteile gemäß öffentlicher Meinung	2.45 (.76)	2.67 (1.11)	2.78 (.94)	2.95 (.85)

Angeklagte werden an Pranger gestellt	3.45 (.60)	3.67 (.49)	3.32 (.81)	3.40 (.82)
--	------------	------------	------------	------------

*Hinweis:* Die Werte weisen die Mittelwerte der Zustimmung zu Aussagen über mögliche Wirkungen von Gerichtsfernsehen aus. Das Ausmaß der Zustimmung wurde anhand 4-stufiger Rating-Skalen gemessen (1=keine Zustimmung; 4=hohe Zustimmung). Die Werte in Klammern geben die zugehörigen Standardabweichungen an. Mittelwerte mit zeilenweise gleichem Index sind statistisch signifikant voneinander verschieden ( $p < .01$ ).

Die vier befragten Gruppen von Juristen äußerten sich in sehr ähnlicher Weise zu den vorgelegten Aussagen. Der Vergleich der Antworten zwischen den vier Gruppen zeigt für nur zwei Aussagen statistisch signifikante Differenzen: Die Angehörigen des rechtswissenschaftlichen Fachbereichs der Universität Osnabrück erwarten deutlich stärkere Behinderungen des Ablaufs einer Gerichtsverhandlung infolge der Filmaufnahmen als die befragten Rechtsanwälte ( $t(68,235)=2,73$ ;  $p < 0,01$ ). Unter den Universitätsangehörigen ist zudem die Einschätzung, Gerichtsfernsehen behindere die Urteilsfindung, deutlich stärker als unter den befragten Verwaltungsrichtern ( $t(38)=3,07$ ;  $p < 0,01$ ).

Die größte Zustimmung durch die befragten Juristen erfahren die Aussagen zu Behinderungen des Ablaufs eines Gerichtsprozesses, zu mediengerechtem Verhalten von Rechtsanwälten sowie zur Prangerwirkung. Nur wenige der befragten Juristen konnten sich dagegen vorstellen, daß infolge von Filmaufnahmen im Gerichtssaal höhere Strafen verhängt oder Urteile gemäß der öffentlichen Meinung abgefaßt werden. Insgesamt vermuten die Juristen stärkere Einwirkungen des Gerichtsfernsehens auf den Ablauf von Gerichtsprozessen als auf deren Ergebnisse. Eine Ausnahme bildet hier die vermutete Prangerwirkung gegenüber Angeklagten, die von allen befragten Juristen als hoch eingeschätzt wird.

Als einzige der befragten Gruppen wurden die an der Universität Osnabrück lehrenden Juristen zu den von ihnen vermuteten Wirkungen sowohl auf das Verhalten von Richtern als auch auf das Verhalten von Rechtsanwälten befragt. Diese Gruppe vermutet mediengerechtes Verhalten bei Rechtsanwälten in weit höherem Maße ( $M=3,75$ ) als bei Richtern ( $M=3,20$ ;  $t(19) = 3,58$ ,  $p < 0,01$ ).

Gelangen Rechtslaien zu denselben Einschätzungen der Wirkungen von Gerichtsfernsehen auf Ablauf und Ergebnis von Gerichtsverhandlungen? Tabelle 2 zeigt die Mittelwerte der Antworten für Rechtslaien und für die Juristen, deren Antworten wegen der festgestellten Übereinstimmungen in den Einschätzungen aggregiert wurden. Dabei wurden nur jene Fragen berücksichtigt, die in beiden Fragebögen enthalten waren.

Tabelle 2: Vermutete Wirkungen von Gerichtsfernsehen auf Ablauf und Ergebnis von Gerichtsverhandlungen: Vergleich der Antworten von Juristen und Rechtslaien

	Juristen n=115	Rechtslaien n=226	t (d.f.)
<u>Ablauf der Verhandlung</u>			
Ablauf wird behindert	3.37 (.86)	2.70 (1.05)	6.31**** (271,513)
Mediengerechtes Verhalten von Rechtsanwälten	3.73 <sup>a</sup> (.45)	3.13 (.99)	6.69**** (192,798)
Mediengerechtes Verhalten von Richtern	3.18 <sup>b</sup> (.67)	2.89 (.99)	2.82** (205,258)
<u>Ergebnis der Verhandlung</u>			
Höhere Strafen	2.54 (.91)	2.01 (.96)	4.85**** (336)

Urteile gemäß öffentlicher Meinung	2.74 (.92)	2.14 (1.07)	5.29**** (260,137)
Angeklagte werden an Pranger gestellt	3.40 (.75)	3.26 (.98)	1.49 (288,959)

*Hinweis:* Die Werte der ersten beiden Spalten weisen die Mittelwerte der Zustimmung zu Aussagen über mögliche Wirkungen von Gerichtsfernsehen und die zugehörigen Standardabweichungen aus. Das Ausmaß der Zustimmung wurde anhand 4-stufiger Rating-Skalen gemessen (1=keine Zustimmung; 4=hohe Zustimmung). Die rechte Spalte weist die t-Werte der Mittelwertdifferenzentests und die zugehörigen Freiheitsgrade aus (\*\*\*\* p<.0001; \*\* p<.01).

<sup>a</sup> n=55 (Verwaltungs-, Amtsrichter, Universitätsangehörige)

<sup>b</sup> n=80 (Rechtsanwälte, Universitätsangehörige)

Die Vermutungen von Juristen und Rechtslaien zu Wirkungen von Gerichtsfernsehen weichen bei fünf von sechs Aussagen markant voneinander ab. Die Rechtslaien vermuten signifikant geringere Wirkungen des Gerichtsfernsehens auf den Ablauf der Verhandlung und auf das Verhalten von Rechtsanwälten und Richtern als die befragten Juristen. Insbesondere erwarten die Rechtslaien geringere Wirkungen auf das Verhalten von Richtern und auf den Verhandlungsablauf; die entsprechenden Antwortmittelwerte liegen nahe am Skalenmittelwert von 2,50. Wirkungen des Gerichtsfernsehens auf die Strafzumessung können sich die befragten Rechtslaien eher nicht vorstellen; die Antwortmittelwerte liegen für beide Aussagen unterhalb des Skalenmittelwertes. Zwar stimmen auch die befragten Juristen diesen Aussagen in geringerem Maße zu als den Aussagen zu Wirkungen auf den Ablauf einer Verhandlung, gleichwohl unterscheiden sich die Antworten statistisch hoch signifikant voneinander. Allein mit Blick auf die Prangerwirkung des Gerichtsfernsehens für die Angeklagten stimmen die Vermutungen von Juristen und Rechtslaien überein: Eine solche Wirkung wird von beiden Gruppen erwartet.

Die Bereitschaft, selbst an einer für das Fernsehen gefilmten Gerichtsverhandlung teilzunehmen, ist bei den befragten Richtern und Rechtsanwälten relativ groß (M=2,31), unter den Rechtslaien dagegen gering: Die Bereitschaft, als Angeklagter an einer solchen Verhandlung teilzunehmen (M=3,52), ist signifikant geringer als die Bereitschaft, als Zeuge an einer solchen Verhandlung teilzunehmen (M=3,04; t(224) = 6,65, p<0,0001).

*Vermutete Wirkungen von Gerichtsfernsehen auf eigenes Verhalten und auf das Verhalten anderer Prozeßbeteiligter.* Amtsrichter, Verwaltungsrichter sowie Rechtsanwälte wurden sowohl nach den vermuteten Wirkungen von Gerichtsfernsehen auf das eigene Verhalten befragt als auch um die Einschätzung von Wirkungen auf das Verhalten anderer Prozeßbeteiligter gebeten. Tabelle 3 zeigt die Mittelwerte der Antworten zu diesen Fragen.

Tabelle 3: Von Juristen vermutete Wirkungen von Gerichtsfernsehen auf eigenes Verhalten und auf das Verhalten anderer Prozeßbeteiligter

	Verwaltungsrichter n=20	Amtsrichter n=15	Rechtsanwälte n=60
<u>Eigenes Verhalten</u>			
Urteilsbegründung	2.40 <sub>a</sub> (.88)	2.13 <sub>a</sub> (.99)	-----
Verhandlungsführung	2.85 <sub>b</sub> (.67)	2.87 <sub>b</sub> (.64)	-----
Formulierung des Plädoyers	-----	-----	2.43 <sub>a</sub> (1.00)

Verhalten während der Verhandlung	-----	-----	3.02 <sub>b</sub> (.95)
<hr/>			
<u>Verhalten anderer Prozeßbeteiligter</u>			
Rechtsanwälte	3.65 <sub>c</sub> (.49)	3.80 <sub>c</sub> (.41)	-----
Richter	-----	-----	3.17 <sub>b</sub> (.69)

*Hinweis:* Die Werte weisen die Mittelwerte der Zustimmung zu Aussagen über mögliche Wirkungen von Gerichtsfernsehen aus. Das Ausmaß der Zustimmung wurde anhand 4-stufiger Rating-Skalen gemessen (1=keine Zustimmung; 4=hohe Zustimmung). Die Werte in Klammern geben die zugehörigen Standardabweichungen an. Mittelwerte mit unterschiedlichen Indizes sind statistisch signifikant voneinander verschieden ( $p < .01$ ).

Sowohl bei den Verwaltungsrichtern als auch bei den Amtsrichtern ist der Grad der Zustimmung zu den Aussagen, die mögliche Wirkungen des Gerichtsfernsehens auf das eigene Verhalten beschreiben, signifikant geringer als zu Aussagen über mögliche Verhaltensänderungen bei Rechtsanwälten. Besonders ausgeprägt sind die Unterschiede, wenn es beim eigenen Verhalten um die Formulierung der Urteilsbegründung geht. Hier vermuten die Verwaltungsrichter und erst recht die Amtsrichter eher keine Auswirkungen des Gerichtsfernsehens auf das eigene Verhalten. Alle befragten Richter halten dagegen Wirkungen auf die eigene Verhandlungsführung eher für möglich.

Die befragten Rechtsanwälte können sich eher nicht vorstellen, daß Gerichtsfernsehen sie bei der Formulierung ihrer Plädoyers beeinflusst. Dagegen halten sie Wirkungen laufender Fernsehaufnahmen aus dem Gerichtssaal auf das eigene Verhalten während einer Verhandlung weit eher für möglich. Die Rechtsanwälte sehen ihr eigenes Verhalten dabei kaum weniger stark durch Gerichtsfernsehen beeinflusst als das Verhalten von Richtern; die entsprechenden Antwortmittelwerte ( $M=3,02$  für vermutete Wirkungen auf eigenes Verhalten,  $M=3,17$  für vermutete Wirkungen auf das Verhalten von Richtern) unterscheiden sich nicht signifikant voneinander.

Die Ergebnisse deuten darauf hin, daß Richter Verhaltensauswirkungen durch Fernsehaufnahmen in Gerichtsverhandlungen in weit stärkerem Maße bei Rechtsanwälten als bei sich selber sehen. Rechtsanwälte dagegen vermuten geringe Wirkungen von Gerichtsfernsehen auf ihre Plädoyers, erwarten ansonsten aber keine Unterschiede in den Wirkungen der Filmaufnahmen auf das eigene Verhalten und das Verhalten von Richtern.

## 5 Diskussion

Die Ergebnisse dieser Untersuchung zeigen erstens, daß die im rechtswissenschaftlichen Schrifttum vertretene und ausdrücklich als „herrschende Meinung“ bezeichnete Ablehnung des Gerichtsfernsehens (Gündisch & Dany, 1999) unter den befragten Juristen tatsächlich vorherrscht: Die große Mehrheit der befragten Juristen lehnt es ab, Fernsehaufnahmen aus Gerichtsverhandlungen zuzulassen. Zweitens unterscheiden sich die von Richtern, Rechtsanwälten und in Forschung und Lehre tätigen Juristen vermuteten Wirkungen einer Einführung von Gerichtsfernsehen nur wenig voneinander. Auffällig ist indes, daß die Universitätsangehörigen in hohem Maße Auswirkungen insbesondere auf den Ablauf gefilmter Gerichtsverhandlungen erwarten. Drittens erwarten die befragten Juristen eher Auswirkungen auf den Verlauf von Gerichtsverhandlungen, weniger jedoch auf die Ergebnisse von Verhandlungen. Eine Prangerwirkung für die Angeklagten durch das Gerichtsfernsehen wird von allen befragten Juristengruppen befürchtet.

Viertens vermuten die befragten Amts- und Verwaltungsrichter geringere Auswirkungen auf ihr eigenes Verhalten als auf das Verhalten anderer Prozeßbeteiligter, speziell der Rechtsanwälte. Dies gilt umgekehrt für das Verhalten der Richter aus Sicht der Rechtsanwälte nur eingeschränkt. Mit Blick auf die mutmaßliche Beeinflußbarkeit eigenen und fremden Verhaltens wird damit eine berufsgruppenbezogene Wahrnehmungsverzerrung deutlich: Die seitens der Richter geäußerten Erwartungen mediengerechten Verhaltens von Rechtsanwälten sind erheblich größer als die Erwartungen der Rechtsanwälte im Hinblick auf mediengerechtes Verhalten von Richtern. Auch die in Forschung und Lehre tätigen Juristen erwarten mediengerechtes Verhalten durch Gerichtsfernsehen weniger bei Richtern, sehr wohl indes bei Rechtsanwälten. Richtern wird von Fachkollegen offenbar in weit höherem Maße zugetraut, die Filmaufnahmen zu ignorieren als Rechtsanwälten. Dieses Ergebnis steht im Einklang mit den im deutschsprachigen juristischen Schrifttum immer wieder geäußerten Befürchtungen, gerade Rechtsanwälte seien bestrebt, das Gericht bei gefilmten Verhandlungen als Bühne zur Selbstdarstellung zu nutzen (Wyss, 1996; Zacharias, 1996; Zuck, 1995).

Ein Vergleich der hier vorgestellten Ergebnisse von Befragungen deutscher Juristen und Rechtslaien mit denen der Befragungen des *New York State Committee* (1997) offenbart auf den ersten Blick deutliche Unterschiede. In dieser amerikanischen Untersuchung vermuteten die Befragten stärkere Einflüsse der Filmaufnahmen auf das Verhalten der Richter als auf das Verhalten der Rechtsanwälte. Die Ergebnisse sind indes nur begrenzt vergleichbar, da Richter und Rechtsanwälte in der amerikanischen adversarischen Prozeßvariante und in der kontinentaleuropäischen nichtadversarischen Prozeßvariante unterschiedliche Rollen ausfüllen. Im adversarischen Prozeßmodell nehmen insbesondere die Anwälte Einfluß auf den Verhandlungsverlauf, im nichtadversarischen Prozeßmodell gestalten dagegen die Richter maßgeblich den Verhandlungsverlauf (Bierbrauer & Gottwald, 1987; Kerr & Bray, 1982). Es zeigt sich hier, daß im Urteil der jeweils übrigen Prozeßbeteiligten Auswirkungen auf das Verhalten der das Geschehen in Gerichtsverhandlungen maßgeblich beeinflussenden Akteure eher nicht oder in nur geringem Maße zu erwarten sind. Dieses Ergebnis deutet darauf hin, daß, wenn überhaupt, nur geringe Auswirkungen des Gerichtsfernsehens auf die prozedurale Gerechtigkeit von Gerichtsverhandlungen erwartet werden.

Fünftens zeigen die Erhebungen deutliche Unterschiede im Urteil von Juristen und Rechtslaien zu den mutmaßlichen Wirkungen von Gerichtsfernsehen auf Verlauf und Ergebnisse von Gerichtsverhandlungen. Rechtslaien können sich zwar Beeinflussungen des Verhandlungsverlaufs in geringem Maße vorstellen, halten indes Auswirkungen auf die Strafzumessung für wenig wahrscheinlich. Dieses Ergebnis mag ein Indiz für die hohen Erwartungen, vielleicht sogar für das große Vertrauen der Bevölkerung in die Qualität der Rechtsprechung sein. Häufige Kommentare der Befragten während der Interviews ließen erkennen, daß die Unabhängigkeit der Richter bei ihrer Urteilsfindung für weite Teile der Bevölkerung außer Frage steht und durch Filmaufnahmen in keiner Weise als gefährdet betrachtet wird. Allein mit Blick auf die Prangerwirkung gegenüber Angeklagten erwarten Rechtslaien beträchtliche Auswirkungen durch das Gerichtsfernsehen, die sich von denen der Juristen nicht unterscheiden. Diese Befürchtungen mögen erklären, warum sich trotz der eher geringen erwarteten Einflüsse des Gerichtsfernsehens auf Ablauf und Ergebnisse von Verhandlungen nur ein geringer Teil der befragten Rechtslaien dazu bereit erklärte, als Zeuge oder gar als Angeklagter an einer gefilmten Verhandlung teilzunehmen.

## **6 Fazit und Forschungsperspektiven**

Zwei Ergebnisse der Befragungen verdienen besondere Beachtung:

- Der Vergleich der Antworten amerikanischer und deutscher Juristen zeigt, daß sich vor dem Hintergrund unterschiedlicher eigener Erfahrungen mit Gerichtsfernsehen deutliche

Unterschiede hinsichtlich der vermuteten Auswirkungen von Fernsehberichterstattung aus Gerichtsverhandlungen einstellen. Personen, die direkte Erfahrungen gemacht haben, erwarten geringere Auswirkungen als Personen ohne solche Erfahrungen. Zwar lassen sich aus den erfahrungsgestützten Vermutungen der amerikanischen Juristen keine zuverlässigen Schlußfolgerungen auf die tatsächlich zu erwartenden Verhaltensfolgen des Gerichtsfernsehens ableiten, da sie letztlich auf ihren subjektiven Wahrnehmungen erlebter Verhandlungen gründen. Der Vergleich mit den Reaktionen deutscher Juristen deckt indes ein typisches Phänomen alltagspsychologischer Erkenntnisgewinnung auf: Fehlen eigene Erfahrungen über Auswirkungen situativer Faktoren, wird dennoch nicht darauf verzichtet, Aussagen über deren „wahrscheinliches“ Ausmaß zu treffen. Dabei werden einfache Ursache-Wirkung-Muster bevorzugt (Ross & Nisbett, 1991). Daher ist es kaum verwunderlich, wenn in der deutschen juristischen Diskussion um das Gerichtsfernsehen fehlende eigene Erfahrungen und wissenschaftliche Erkenntnisse durch alltagspsychologische Erkenntnisbildung ersetzt werden.

- Die Antworten der Rechtslaien erweisen überdies ein beträchtliches Vertrauen der Bevölkerung in die Unabhängigkeit der deutschen Rechtsprechung. In ihrem Urteil schmälert Gerichtsfernsehen die prozedurale und erst recht die distributive Gerechtigkeit von Prozessen kaum oder gar nicht. Das beträchtliche Interesse an einer Filmberichterstattung aus Gerichtsverhandlungen läßt zudem Wissensdefizite sowie den Wunsch erkennen, diese zu verringern.

Sicherlich lassen die Ergebnisse der Befragungen keinerlei Folgerungen darüber zu, welche Auswirkungen tatsächlich von Gerichtsfernsehen ausgehen. Gesicherte Erkenntnisse hierüber stehen bis heute aus. Ihre Gewinnung bedarf sowohl weiterer Erfahrungen mit dem Gerichtsfernsehen als insbesondere auch experimenteller Untersuchungen auf Grundlage einschlägiger sozialpsychologischer Theorien, etwa der Theorie der Selbstdarstellung (Laux & Schütz, 1996; Leary, 1995), der Theorie der Sozialen Aktivierung und Hemmung (Zajonc, 1965) oder der Theorie der Selbstaufmerksamkeit (Wicklund, 1975). Diese Theorien bieten Ansatzpunkte für empirische Untersuchungen über Verhaltensbeeinflussungen durch Gerichtsfernsehen. Die weitere Diskussion um die Einführung von Gerichtsfernsehen in Deutschland benötigt Erkenntnisse aus entsprechenden Untersuchungen über tatsächliche Auswirkungen von Fernsehkameras im Gerichtssaal. Solange sie nicht zur Verfügung stehen, bleiben die in der juristischen Kontroverse um Gerichtsfernsehen ausgetauschten Argumenten der Gegner und Befürworter von Fernsehaufnahmen aus Gerichtsverhandlungen weiterhin Spekulationen.

## **Literatur**

- Barber, S. (1987). *News cameras in the courtroom. A free press – fair trial debate*. Norwood, NJ: Ablex Publishing Corporation.
- Bierbrauer, G. & Gottwald, W. (1987). Psychologie und Recht – Brückenschlag zwischen Fakten und Fiktion. In J. Schultz-Gambard (Ed.), *Angewandte Sozialpsychologie*. München: Psychologie Verlags-Union.
- Bierbrauer, G. & Klinger, E. W. (im Druck). Akzeptanz von Entscheidungen durch faire Verfahren. Einige Überlegungen und Befunde aus der Forschung über Verfahrensfairneß. In S. Wesche, H. Hof & F. Haft (Eds.), *Verhaltensgrundlage des Rechts*.
- Braun, Y. (1998). *Medienberichterstattung über Strafverfahren im deutschen und englischen Recht*. Freiburg/Breisgau: Max-Planck-Institut für ausländisches und internationales Strafrecht („edition iuscrim“).

- Brosius, H.-B. & Engel, D. (1997). Die Medien beeinflussen vielleicht die anderen, aber mich doch nicht: Zu den Ursachen des Third-Person-Effekts. *Publizistik*, 38, 325–345.
- Cohn, M. & Dow, D. (1998). *Cameras in the courtroom: Television and the pursuit of justice*. New York: McFarland.
- Deutscher Richterbund (1996). DRB strikt gegen >court tv<. *Deutsche Richterzeitung*, 6, 246–250.
- Duck, J. M. & Mullin, B.-A. (1995). The perceived impact of the mass media: Reconsidering the third person effect. *European Journal of Social Psychology*, 25, 77–93.
- Eberle, R. G. (1996). Verwaltungsgerichte in der Medienberichterstattung am Beispiel von Tageszeitungen in Hessen. *Zeitschrift für Rechtssoziologie*, 17, 300-309.
- Gehring, H. (1998). Sozialpsychologische Aspekte von Fernsehberichterstattung aus Gerichtsverhandlungen. *Zeitschrift für Rechtspolitik*, 31, 8–10.
- Gibbon, P. & Durkin, K. (1995). The third person effect: Social distance and perceived media bias. *European Journal of Social Psychology*, 25, 597–602.
- Goethals, G. R., Messick, D. M. & Allison, S. T. (1991). The uniqueness bias: Studies of constructive social comparison. In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research*. Hillsdale, NJ: Erlbaum.
- Gündisch, J. & Dany, P. (1999). Rundfunkberichterstattung aus Gerichtsverhandlungen. *Neue Juristische Wochenschrift*, 4, 256–260.
- Gunther, A. (1991). What we think others think. *Communication Research*, 18, 355–372.
- Innes, J. M. & Zeitz, H. (1988). The public's view of the impact of the mass media: A test of the 'third person' effect. *European Journal of Social Psychology*, 18, 457–463.
- Kerr, N. L. & Bray, R. M. (1982). *The psychology of the courtroom*. New York: Academic Press.
- Kortz, H. (1997). Ausschluß der Fernsehöffentlichkeit im Gerichtsverfahren. *Archiv für Presserecht*, 1, 443–450.
- Laux, L. & Schütz, A. (1996). *Wir, die wir gut sind: Die Selbstdarstellung von Politikern zwischen Glorifizierung und Glaubwürdigkeit*. München: dtv.
- Leary, M. R. (1995). *Self-Presentation*. Boulder: Westview.
- Lind, E. A. & Tyler, T. R. (1988). *The social psychology of procedural justice*. New York: Plenum Press.
- Machura, S. (1998). Rechtsfilm und Rechtsrealität. *Richter ohne Robe*, 10, 39–42.
- Merten, K. (1997). Die Rolle der Medien bei der Vermittlung zwischen Recht und Gesellschaft. *Zeitschrift für Rechtssoziologie*, 18, 16-30.
- New York State Committee (1997). An open courtroom. Cameras in New York courts 1995 – 1997. *New York Law Journal*, special edition April.
- Perloff, R. M. (1989). Ego-involvement and the third person effect of televised news coverage. *Communication Research*, 16, 2, 236–262.
- Ross, L. & Nisbett, R. E. (1991). *The person and the situation*. New York: McGraw-Hill.
- Thaler, P. (1994). *The watchful eye. American justice in the age of television trials*. Westport: Praeger.
- Walther, S. (1998). Mehr Publizität oder mehr Diskretion? *Juristenzeitung*, 23, 1145-1153.
- Westfeldt, W. & Wicker, T. (1998). *Indictment: The news media and the criminal justice system*. Nashville: First Amendment Center.
- Wicklund, R. A. (1975). Objective self awareness. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, 7, 319–342.
- Wyss, M. P. (1996). Öffentlichkeit von Gerichtsverfahren und Fernsehberichterstattung. *Europäische Grundrechte Zeitschrift*, 1-2, 1–17.
- Zacharias, T. (1996). Das Öffentlichkeitsprinzip im Strafprozeß. *Österreichische Juristen-Zeitung*, 18, 681–694.
- Zajonc, R. B. (1965). Social facilitation. *Science*, 149, 6, 269–274.



Zuck, R. (1995). Court TV: Das will ich sehen! *Neue Juristische Wochenschrift*, 32, 2082–2083.

## *Anhang A*

### **Items der Befragung der Juristen**

#### Allgemeine Einstellung zum Gerichtsfernsehen

Grundsätzlich sollten Gerichtsverhandlungen im Fernsehen gezeigt werden.

#### Eigene Bereitschaft zur Teilnahme

Ich hätte nichts dagegen, wenn mich jeder in Ausübung meiner Tätigkeit im Fernsehen sehen könnte.

#### Ablauf einer Verhandlung

Wenn eine Gerichtsverhandlung fürs Fernsehen gefilmt wird, dann behindert das den Ablauf der Verhandlung.

#### Ergebnis und Auswirkungen einer Verhandlung

Wenn eine Gerichtsverhandlung fürs Fernsehen gefilmt wird, dann behindert das die Urteilsfindung.

Wenn eine Gerichtsverhandlung fürs Fernsehen gefilmt wird, dann besteht die Möglichkeit, daß das Urteil entsprechend der öffentlichen Meinung abgefaßt wird.

Wird eine Gerichtsverhandlung fürs Fernsehen gefilmt, dann besteht die Möglichkeit, daß höhere Strafen ausgesprochen werden.

Wenn eine Gerichtsverhandlung fürs Fernsehen gefilmt wird, dann werden die Angeklagten dadurch an den Pranger gestellt.

#### Verhalten der Prozeßbeteiligten

Wird eine Gerichtsverhandlung fürs Fernsehen gefilmt, dann werden die Richter / Anwälte sich mediengerecht verhalten.

Wenn eine Verhandlung fürs Fernsehen gefilmt wird, dann nimmt das Einfluß auf meine Verhandlungsführung. (nur bei Befragung der Richter)

Wenn ich bei der Urteilsbegründung im Fernsehen zu sehen wäre, würde ich diese anders formulieren. (nur bei Befragung der Richter)

Wenn eine Verhandlung fürs Fernsehen gefilmt wird, dann nimmt das Einfluß auf mein Verhalten in dieser Verhandlung. (nur bei Befragung der Anwälte)

Wenn ich bei der Verhandlung im Fernsehen zu sehen wäre, würde ich mein Plädoyer anders formulieren. (nur bei Befragung der Anwälte)

## *Anhang B*

### **Items der Befragung der Rechtslaien**

#### Eigene Bereitschaft zur Teilnahme

Wenn ich selber als Zeuge an einem Gerichtsverfahren teilnehmen müßte, dann hätte ich etwas dagegen, wenn mich dabei jeder im Fernsehen sehen könnte.

Wenn ich selber als Angeklagter an einem Gerichtsverfahren teilnehmen müßte, dann hätte ich etwas dagegen, wenn mich dabei jeder im Fernsehen sehen könnte.

#### Ablauf einer Verhandlung

Wenn eine echte Gerichtsverhandlung fürs Fernsehen gefilmt wird, dann behindert das den Ablauf der Verhandlung.

#### Ergebnis und Auswirkungen einer Verhandlung

Wenn eine echte Gerichtsverhandlung fürs Fernsehen gefilmt wird, dann behindert das die Urteilsfindung.

Wenn eine Gerichtsverhandlung fürs Fernsehen gefilmt wird, dann fassen die Richter das Urteil entsprechend der öffentlichen Meinung ab.

Wird eine Gerichtsverhandlung fürs Fernsehen gefilmt, dann werden höhere Strafen ausgesprochen.

Wenn eine Gerichtsverhandlung fürs Fernsehen gefilmt wird, dann werden die Angeklagten dadurch an den Pranger gestellt.

#### Verhalten der Prozeßbeteiligten

Wird eine echte Gerichtsverhandlung fürs Fernsehen gefilmt, dann verhalten die Richter/Anwälte sich mediengerecht.

Wenn eine echte Gerichtsverhandlung fürs Fernsehen gefilmt wird, dann sind die Angeklagten verunsichert.

Wenn eine echte Gerichtsverhandlung fürs Fernsehen gefilmt wird, dann bereiten sich die Rechtsanwälte besser auf die Verhandlung vor.

<sup>1</sup> Wir danken Günter Bierbrauer und Gudrun Schwarzer für wertvolle Anregungen und die kritische Durchsicht des Manuskripts sowie Raphael Hartmann für die Mitarbeit bei der Datenauswertung.



# Psychologische Faktoren der Arbeitssicherheit bei der Polizei am Beispiel Schutzwesten

*Clemens Lorei*

*Redaktion Polizei und Wissenschaft, Frankfurt*

Polizeiliche Einsätze sind seit jeher und global mit dem Risiko verbunden, von Rechtsbrechern verletzt oder getötet zu werden. Als im Jahr 2000 acht Polizisten durch Schuss- oder Stichwaffen im Dienst ermordet wurden, führte dies zu einer entsprechenden Diskussion über Möglichkeiten zur Steigerung der Sicherheit von Polizeibeamten. Im Sinne der Schutzzielhierarchie stellt die Nutzung von Schutzwesten zur Verringerung der Gefährdung eine für die Polizei naheliegende Option dar. Dabei lassen sich verschiedene psychologische Einflussfaktoren auf das Trageverhalten annehmen, die aus industriellen Bereichen bereits bekannt sind. In einer bundesländerübergreifenden, nicht repräsentativen Befragung von 1679 Polizistinnen und Polizisten mittels Fragebogen wurde versucht, Faktoren aus den Bereichen persönliche Gewalterfahrung, Gefahrenkognition, Funktionalitätsüberzeugung, Annahmen über Gefährdung auf Grund einer Schutzweste, Tragekomfort, Verfügbarkeit und Sicherheitskultur zu erfassen. Die Ergebnisse zeigen, dass sowohl die angegebene Nutzung der Schutzausrüstung als auch die erfragten Einflussfaktoren interindividuell deutlich variieren. Es konnten geringe bis mittlere korrelative Zusammenhänge zwischen verschiedenen Faktoren und dem angegebenen Trageverhalten gefunden werden. Besonders variiert die empfundene Beweglichkeitseinschränkung, der Aufwand, der notwendig ist, um diese Ausrüstung zu tragen, und das Trageverhalten der Kollegen mit der Nutzung. Aus den Ergebnissen wird geschlussfolgert, dass das Tragen von Schutzwesten bei der Polizei multifaktoriell beeinflusst wird und auf komplexen und individuellen Entscheidungen basiert. Entsprechend wird ein mehrdimensionales psychologisches Vorgehen zur Steigerung der Tragerate empfohlen.

## **Einleitung**

Acht Polizistenmorde im Jahr 2000 ließen zum wiederholten Male Rufe nach Schutzwesten laut werden. Ausgangspunkt ist die Annahme, dass Schutzwesten die fatalen Konsequenzen der Schüsse (und auch der Messerstiche) auf die Polizisten hätten mildern können. Eine ballistische Beurteilung der Sachverhalte kann und soll hier nicht vorgenommen werden. Klar ist, dass ganz allgemein Schutzwesten bei Beschuss (und Messerangriffen) eine erhöhte Sicherheit - beschränkt auf gewisse Körperregionen und Angriffsintensitäten - bieten können. Dies natürlich nur dann, wenn sie getragen werden. Es ist vor allem aus den Bereichen der Psychologie der Arbeitssicherheit und der Unfallpsychologie bekannt, dass diese Voraussetzung nicht selbstverständlich erfüllt wird. Es scheint auch für den Bereich Schutzwesten bei der Polizei durchaus gerechtfertigt, davon auszugehen, dass die Verfügbarkeit einer Schutzweste nicht automatisch zum Tragen derselben führt (vgl. Heide, 1998; Bechstein, 1998; für USA: Pokojewski, 1998; Brand, 1993). In der Analyse zur Tötung US-amerikanischer Polizisten stellt das Federal Bureau of Investigation (1994, S. 41) entsprechend fest, dass zu selten Schutzwesten von Polizisten getragen werden, und fordert entsprechende Trainingsprogramme, die ein situativ adäquates Tragen schulen.

## **Theorie**

### Schutzzielhierarchie

Sicherheit im Sinne von Unfallverhütung als Beseitigung von Gefahren und Verringerung ihrer Auswirkungen lässt sich hinsichtlich ihrer Wirksamkeit in eine Rangordnung bringen, auch Schutzzielhierarchie genannt (Skiba 1997, S. 39ff.; vgl. Zapf & Dormann, 2001, S. 561):

- Beseitigung der Gefahr
- Beseitigung der Gefährdung

- Verringerung der Gefährdung
- Anpassung an die Gefährdung

Die Beseitigung der Gefahr, also dem Objekt oder der Situation von der eine Gefahr ausgeht, macht eine Schädigung unmöglich. Eine Beseitigung der Gefährdung besteht in der Trennung von Gefahr und Mensch, wodurch die Gefahr nicht negativ auf den Menschen einwirken kann. Bei der Verringerung der Gefährdung durch Abschirmung ist diese weiterhin vorhanden, jedoch werden die potenziellen negativen Auswirkungen vermindert. Bei der Anpassung an die Gefährdung hängt die Auswirkung der Gefahr hauptsächlich von Verhalten und Eigenschaften des Menschen ab. Sind diese sicherheitsgerecht, können Auswirkungen von Gefahren reduziert werden. Die Maßnahmen sollen in ihrer Wirksamkeit mit zunehmender Abhängigkeit vom menschlichen Verhalten abnehmen (Skiba 1997, S. 39ff.). Die Verringerung der Gefährdung wie z.B. durch Schutzbekleidung ist von einer Nutzung abhängig, so dass sie durch mangelnden Gebrauch unwirksam gemacht wird.

Ist ein Zusammentreffen von Mensch und Gefahr weder durch Beseitigung der Gefahr noch durch Beseitigung der Gefährdung zu begegnen, ist der Einsatz von Schutzausrüstung zur Verringerung der Gefährdung einzusetzen (vgl. Weber, 1994). Für Polizeikräfte bedeutet dies, da sie gerade in Gefahrenbereichen zur Beseitigung von Gefahren und Gefährdungen für die Gesellschaft eingesetzt werden, dass auch sie einen Anspruch auf die Verringerung ihrer Gefährdung haben sollten. Es ist dabei davon auszugehen, dass überwiegend oder sogar ausschließlich Maßnahmen zur Verringerung der Gefährdung und Anpassung an die Gefährdung zum Einsatz kommen. Eine Beseitigung der Gefahr, die in der Nichtexistenz von Gewalttättern bestünde, ist zweifellos angestrebt, kann aber nicht als abgeschlossen angesehen werden. Eine Beseitigung der Gefährdung, die in der Trennung von Gefahr und Mensch besteht, kann je nach Situation nicht möglich sein: Noch muss der Polizist in persönlichen Kontakt und Interaktion mit Gewaltverbrechern treten. Deshalb muss die Gefährdung verringert werden, indem Schutzwesten angelegt werden und eine umfassende Anpassung an die Gefährdung in Aus- und Fortbildung stattfindet.

### Tragen von Schutzausrüstung

Pokojewski (1998) gibt an, dass die Tragerate von Schutzwesten bei der Polizei in den USA ca. 25% beträgt. Ähnliches berichtet Brand (1993). Bechstein (1998) stellt ebenso fest, dass in den USA nur 26% derjenigen, die über eine Schutzweste verfügen, diese auch ständig im Dienst tragen. Dabei sind die den Gebrauch einer Schutzweste mitbestimmenden Faktoren in Polizeikreisen nicht unbekannt (vgl. Berka & Sauter, 1995; Bechstein, 1998; Panek, 1998; Van Zijl, 1998; Schiller & Weisswange, 2000; Damm, 2000; Brand, 1993). Die Nichtnutzung von Schutzausrüstung ist jedoch kein polizeispezifisches Problem. Vielmehr ist Gebrauch von Schutzmitteln in der Psychologie der Arbeitssicherheit schon lange ein verbreitetes Thema. Dabei hängt die Verringerung der Gefährdung durch Schutzausrüstung von der Akzeptanz und der Mitarbeit der zu schützenden Person ab. Sie entscheidet, ob die Maßnahme eingesetzt wird. Dabei lässt sich beobachten, dass unabhängig von Vorschriften oder eigener Entscheidung Personen eine Teilnahme an Sicherheit spendenden Maßnahmen ablehnen (Strobel, 1994; Weber, 1994; Strobel & Wittmann, 1998). Weber (1994, S.603) stellt in diesem Zusammenhang fest, dass in der Industrie nur selten die Schutzwirkung unzureichend ist, vielmehr eine mangelnde Tragequote das eigentliche Sicherheitsproblem darstellt. Der Gebrauch und die Akzeptanz von Schutzmitteln scheint in hohem Maße von der Gefahrenkognition (vgl. Musahl, 1997; Strobel, 1994), der Einstellung (Schmidt, 1990, S. 180), der Tragebequemlichkeit (Schmidt, 1990, S. 180) bzw. dem Tragekomfort und der Handhabbarkeit, der optischen Gestaltung, der Kenntnis der Gefahren und Gesundheitsrisiken, dem Wissen um die Funktionalität persönlicher Schutzausrüstung (Strobel, 1994; Weber, 1994; Strobel & Wittmann, 1998), der Sicherheitskultur u.a. im Sinne einer kollektiven Trageverbreitung und der Vorschriftensituation (Weber, 1994) abhängig zu

sein. Im Folgenden sollen hierzu Feststellungen aus dem industriellen Bereich dargestellt werden, um sie anschließend im Rahmen dieser Studie für den polizeilichen Einsatz zu prüfen.

### Gefahrenkognition und Einstellung

Gefahrenkognition entspricht nach Hoyos (1980) der subjektiven Repräsentation von Gefährdungen und Gefahren. Die Einschätzung der Gefährlichkeit einer Situation bestimmt dabei das individuelle Vorsorgeverhalten (Musahl, Groß-Thomas & Müller-Gethmann, 1994) bzw. der Umgang mit Gefahren hängt in einem erheblichem Ausmaß vom Wissen um diese ab (Hoyos, Hirsch & Arnhold, 1991, S. 68). Die subjektive Notwendigkeit des Gebrauchs einer Schutzausrüstung korrespondiert dann nicht unbedingt mit der objektiv vorhandenen Gefahr (Strobel, 1994, S.626). Mangelnde Tragebreitschaft lässt sich deshalb mitunter auf Wissens- bzw. Gefahrenkognitionsdefizite zurückführen. Denn wer von keiner Gefährdung ausgeht, wird sich auch nicht vor ihr schützen wollen. Der Zusammenhang zwischen Gefahrenkognition und Vorsorgeverhalten bzw. Umgang mit Gefahren kann dabei als belegt angesehen werden (Hoyos et al., 1991, S. 68; Musahl et al., 1994).

### Tragekomfort und Handhabbarkeit

Strobel (1994, S. 624) und auch Weber (1994, S. 608) sehen im erlebten Tragekomfort als Summe der wahrgenommenen Trageeigenschaften einen erheblichen Faktor bezüglich der Bereitschaft, Schutzausrüstung zu tragen. Hier wird die Bedeutung einer nur mangelhaften ergonomischen Gestaltung von Schutzausrüstung für die persönliche Sicherheit klar. In der Industrie findet sich wie im polizeilichen Bereich häufig das Argument der klimatischen Unzumutbarkeit. Daneben sind Einschränkungen der Beweglichkeit und allgemeine Passform kritische Eigenschaften von Schutzwesten, die über die Benutzung entscheiden können (vgl. auch Weber, 1994). Hier sei auch auf die teilweise nur ungenügende Berücksichtigung anatomischer Unterschiede zwischen Männern und Frauen hingewiesen. Das Gewicht der Schutzausrüstung scheint immer wieder ein kritischer Faktor zu sein (Weber, 1994). Weiterhin ist neben einer Beweglichkeitseinengung auch eine Behinderung durch die Schutzausrüstung zu beachten. In diesem Zusammenhang wurden des öfteren Probleme mit der Dienstwaffe bzw. dem Holster laut. Nicht zu vergessen ist die Verfügbarkeit und der empfundene Aufwand, der mit einer Beschaffung verbunden ist. Nicht zu vergessen ist natürlich auch der Pflegeaufwand, der notwendig ist, um Ausrüstung zu bewahren und trotz Schmutz- und Geruchseinwirkung tragbar zu halten.

### Optische Gestaltung

Nach Strobel (1994, S. 625) und Strobel & Wittmann (1998) führt auch ein als angenehm erlebtes Erscheinungsbild zu Nutzungsbereitschaft. Verständlich ist dies, da Kleidung ein Statussymbol ist und Aussehen einen nicht unwichtigen Faktor der sozialen Interaktion darstellt.

### Wissen um Funktionalität

Wird die Funktionalität der Schutzmaßnahme in Frage gestellt, ist entsprechend die Nutzung unwahrscheinlicher (vgl. Strobel, 1994; Weber, 1994). Strobel (1994, S. 627) empfiehlt deshalb, Betroffene unbedingt von der Schutzwirkung zu überzeugen.

### Verfügbarkeit

Selbstverständliche Voraussetzung für den Einsatz von Schutzausrüstung ist deren Vorhandensein und Verfügbarkeit. Jeglicher Aufwand, der erforderlich ist, diese anzulegen, macht die regelmäßige Benutzung dieser Sicherheitsausrüstung unwahrscheinlicher.

## Sicherheitskultur

Der Begriff Sicherheitskultur als Teil der Organisationskultur weist auf eine notwendige ganzheitliche Sichtweise hin, die Organisation, Mensch und Technik umfasst (Künzler & Grote, 1996). Sie wird von Künzler & Grote (1996, S. 83) definiert als „die Gesamtheit der von allen Mitgliedern einer Organisation geteilten direkt und indirekt sicherheitsbezogenen Grundannahmen, Werte und Normen, die Ihren Ausdruck im konkreten Umgang mit Sicherheit, aber auch in allen anderen Funktionsbereichen der Organisation finden“. Es wird deutlich, dass nicht der Einzelne allein für Sicherheit verantwortlich ist, sondern innerhalb der Organisation viele Faktoren beteiligt sind. Sicherheit findet also auf allen Ebenen und in allen Bereichen einer Organisation ihren Ursprung (vgl. Künzler & Grote, 1996, S. 84f.).

## Soziale Faktoren, Lernprozesse und Gewohnheiten

Den Einfluss sozialer Faktoren wie Kollegen und Vorgesetzte, Lernprozesse und Gewohnheiten auf sicheres Verhalten sowie das Tragen von Schutzausrüstung macht Burkardts (1981, 1992) Theorie zur Entstehung sicherer und sicherheitswidriger Gewohnheiten deutlich. Basierend auf Lernvorgängen ergeben sich Arbeitsgewohnheiten als Folge von Vor- und Nachteilen von sicherheitsrelevantem Verhalten. Neben diesen Lernprozessen sind natürlich für den Einfluss von Kollegen und Vorgesetzten auch Aspekte der sozial-kognitiven Lerntheorie, des Gruppenzwangs, des Risiko-Shifts, des Gehorsams, der Konformität und Ähnliches potenziell zu berücksichtigen.

## Vorschriften

Die Psychologie zeigt u.a. im Bereich von Verstößen, dass Vorschriften nicht unbedingt immer ausreichen (vgl. Reason, 1994; Zeitlin, 1994). Es sind häufig begleitende Maßnahmen nötig, die ein Befolgen der Sicherheitsvorschriften wahrscheinlicher machen (Zeitlin, 1994).

## Methode

Ziele dieser Studie als Pilotprojekt ist die explorative Deskription potenziell tragerelevanter Faktoren und eine explorative inferenzstatistische Prüfung verschiedener Hypothesen zum Einfluss tragerelevanter Faktoren. Dabei war die Stichprobe nicht auf Repräsentativität ausgerichtet. Hierfür fehlten sowohl finanzielle Mittel als auch die Infrastruktur sowie entsprechende Erkenntnisse über die Population.<sup>1</sup> Eine Generalisierung der Ergebnisse auf die Polizeipopulation ist deshalb nicht möglich.

## Fragebogen

Um mit den zur Verfügung stehenden Mitteln die tragerelevanten Faktoren bei möglichst vielen Polizisten erfassen zu können, fiel die Entscheidung auf die Konstruktion eines standardisierten Fragebogens mit geschlossenen Antworten. Dabei wurde darauf geachtet, dass zwar alle potenziell tragerelevanten Bereiche angesprochen wurden, jedoch der Fragebogen sich insgesamt kurz und knapp darstellt, um eine Beteiligung motivational zu unterstützen.

Der Fragebogen enthielt zu Beginn eine kurze Erläuterung des Zweckes der Studie und Informationen zum weiteren Verlauf. Es wurde darauf hingewiesen, dass die Teilnahme freiwillig und anonym erfolge. Es wurde die Notwendigkeit, alle Fragen zu beantworten,



dargelegt und sich vorab für die Teilnahme bedankt. Anschließend folgten Fragen zu den Bereichen:

- Person (Geschlecht, Alter, Familienstand, Diensterfahrung, Dienstbereich)
- Gewalterfahrung und Vorbereitung (Erfahrungen mit Gewalt gegen die eigene Person, Erfahrung mit Schusswaffengebrauch, Fortbildungs- und Trainingsgewohnheiten)
- Gefahrenkognition (Angst vor eskalierenden Situationen, Schätzung der Wahrscheinlichkeit eskalierende Situationen zu erleben, Ansicht über Häufigkeit von Gewalt gegenüber Polizeibeamten)
- Gefährdungen auf Grund der Schutzweste (Kopfschuss)
- Funktionalität der Schutzweste (Schutzwirkung, Fläche)
- Tragekomfort (Beweglichkeit, Passform, klimatische und modische Eigenschaften und deren Einfluss auf das Trageverhalten)
- Trageverhalten (Besitz einer Schutzweste, Aufwand, Tragehäufigkeit)
- Sicherheitskultur und andere Einflüsse (Vorschrift, Kollegen, Vorgesetzte)

Die zur Verfügung stehenden Antwortmöglichkeiten waren stets verbal formuliert. Eine Frage enthielt die Aufgabenstellung, Eigenschaften der Schutzweste bezüglich ihrer subjektiven Wichtigkeit in eine Rangfolge zu bringen. Es kann nicht davon ausgegangen werden, dass die möglichen Antworten Ordinalskalenniveau überschreiten. Teilweise ist eine nominale Skalierung offensichtlich (Geschlecht, Bundesland). Kennwerte der testtheoretischen Qualität sind unbekannt. Es wird davon ausgegangen, dass diese Testgütekriterien ein ausreichendes Niveau erreichen.

## Durchführung

Befragt wurden Polizeiangehörige im Zeitraum vom 27.11.2000 bis zum 2.2.2001.<sup>2</sup> Der Kontakt zu den Teilnehmern wurde dabei hergestellt durch persönliche Direktansprache (nur in Hessen), ein E-Mailing<sup>3</sup> an 149 E-Mail Adressen von Polizeidienststellen der Länder und des Bundes (zusammengestellt mittels Internet-Recherche) sowie an den Webmaster einer Newsgroup und die Erstellung einer Seite im Internet ([www.polizeipsychologie.de](http://www.polizeipsychologie.de)). Der Rücklauf der ausgefüllten Fragebogen konnte mittels Post, Fax, E-Mail oder persönliche Abholung (nur in Hessen) geschehen. Hierdurch konnten länderübergreifend 1679 Teilnehmer gewonnen werden. Davon waren 200 (11.9%) Frauen (22 ohne Angaben). 434 Personen (25.8%) gehörten der Altersgruppe bis 30 Jahren an. 645 Teilnehmer (38.4%) gaben an, 31-40 Jahre alt zu sein. 40 Jahre und älter waren 582 (34.7%) der Befragten (18 ohne Angaben). 383 (22.8%) verfügten über bis zu 10 Jahren Diensterfahrung. 501 (29.8%) waren bereits 11-20 Jahre im Polizeidienst, während 732 (43.6%) sogar schon über 20 Dienstjahre aufweisen konnten (63 ohne Angaben). 1479 (88.1%) waren Angehörige der Schutzpolizei, 141 (8.4%) der Kriminalpolizei und 50 (3.0%) der Bereitschaftspolizei (9 ohne Angaben). Als die überwiegende dienstliche Tätigkeit gaben 229 (13.6%) den Innendienst, 1211 (72.1%) den Streifendienst, 164 (9.8%) Ermittlungen und 36 (2.1%) Sonstiges an. Die Befragten entstammen mit deutlich unterschiedlicher Beteiligung (von einer Person bis zu 845 Befragten) 10 verschiedenen Bundesländern (7 mal keine Angaben). Als Dienstbereich wurde von 413 (24.6%) eine ländliche, von 598 (35.6%) eine kleinstädtische und von 388 (23.1%) eine großstädtische Gegend angegeben.

## Ergebnisse

### Schutzwestenbesitz und Trageverhalten

1208 (71.9%) der Befragten gaben an, eine dienstliche Weste oder einen Zugang zur ihr zu haben. 204 (12.2%) erklärten, sie besäßen eine private Schutzweste (12 Fälle ohne Angaben). Dabei sind die befragten Personen Angehörige unterschiedlicher Bereiche und gehen in ihrer überwiegenden Zeit unterschiedlichen dienstlichen Tätigkeiten nach. Tatsächlich unterscheiden sich die Angaben zum Schutzwestenbesitz der verschiedenen Polizeibereiche (Schutz-, Kriminal- und Bereitschaftspolizei,  $\chi^2=90.584^{**}$ ,  $df=4$ ,  $N=1658$ ) und dienstlichen Tätigkeiten (Innendienst, Streifendienst, Ermittlungen und Sonstiges;  $\chi^2=111.781^{**}$ ,  $df=6$ ,  $N=1629$ ).

Nie eine Schutzweste zu tragen, geben 119 (7.2%) der Befragten an. 388 (23.5%) schildern, dass sie selten die Schutzweste tragen. Gelegentlich soll die Schutzweste von 397 (22.9%) getragen werden. 346 (20.9%) stellen ihr Trageverhalten als häufig dar. Nach eigenen Angaben tragen 422 (25.5%) immer eine Schutzweste (25 Fälle ohne Angaben). Wie zu erwarten ist, unterscheiden sich die angegebenen Tragehäufigkeiten der verschiedenen Polizeibereiche (Schutz-, Kriminal- und Bereitschaftspolizei,  $\chi^2=130.653^{**}$ ,  $df=8$ ,  $N=1652$ ) und dienstlichen Tätigkeiten (Innendienst, Streifendienst, Ermittlungen und Sonstiges;  $\chi^2=281.382^{**}$ ,  $df=12$ ,  $N=1623$ ).

Besonders auffallend ist die unterschiedliche angegebene Tragehäufigkeit in Abhängigkeit vom Besitz einer Schutzweste (dienstliche, private, weder dienstliche noch private;  $\chi^2=593.219^{**}$ ,  $df=8$ ,  $N=1654$ ). Neben der Trivialität, dass kein Besitz zu entsprechend seltenem Tragen führt - ein Ausleihen einer Weste o.ä. ermöglicht ein gelegentliches Tragen - fällt vor allem auf, dass private Schutzwesten überwiegend immer (60.4%) und insgesamt öfter als dienstliche Schutzwesten getragen werden.

Die Angabe zur Tragehäufigkeit stellt einen momentanen Wert dar. Um zu beurteilen, ob diese Häufigkeit zeitlich stabil oder nur die Reaktion auf bestimmte Ereignisse ist, muss erfasst werden, seit wann diese Tragehäufigkeit vorliegt. Allgemein geben 495 (29.5%) an, schon immer die Schutzweste so häufig zu tragen, wie sie zuvor geantwortet hatten. 559 (33.3%) tragen sie seit langem und 323 (19.2%) seit kurzem mit dieser Häufigkeit. 177 (10.5%) können ein Ereignis nennen, seit dem sie die Schutzweste in der angegebenen Art tragen. Als Ereignisse finden sich insbesondere Angriffe auf den Befragten selbst oder Kollegen, Ermordungen von Polizisten - vor allem diese Ereignisse im Jahr 2000 -, die Geburt eines Kindes und die Anschaffung oder der Erhalt der Schutzweste. Bezogen auf die Häufigkeit des Tragens der Schutzweste ergibt dies, dass seit kurzem bzw. seit bestimmten Ereignissen - hier seien besonders die Angriffe auf Polizisten und Ermordungen im Jahr 2000 erwähnt - die angegebene Tragehäufigkeit zugenommen hat.

#### Wichtigkeit verschiedener Eigenschaften

Die Befragten wurden aufgefordert, 11 Eigenschaften einer Schutzweste in eine Rangfolge zu bringen, die ausdrückt, wie wichtig ihnen eine gute Ausprägung dieser Eigenschaften ist.<sup>4</sup> Die Rangfolge wird angeführt von der Schutzwirkung gegen Geschosse und Stiche (siehe Tabelle 1). Offensichtlich steht der eigentliche Zweck einer Schutzweste tatsächlich an vorderster Stelle. Die Meinungen hierüber gehen vor allem bei der Wirkung gegen Geschosse kaum auseinander, was durch die Kurtosis der Verteilung bestätigt wird. Das Mittelfeld bezüglich guter Ausprägung von Eigenschaften wird durch die Größe der geschützten Fläche, Gewicht der Weste, die Passform und Tragebequemlichkeit, die trotz Schutzweste vorhandene Beweglichkeit und den Splitterschutz gebildet. Hier ähneln die Meinungen schon eher einer Normalverteilung und gehen auseinander: Während einige ein gutes Gewicht für wichtiger halten, sind anderen angenehme klimatische Verhältnisse bedeutsamer. Am Ende der Wichtigkeitsskala stehen Sichtbarkeit/Auffälligkeit, Pflege- und Reinigungsaufwand und modische Eigenschaften.<sup>5</sup>

Zu bemerken ist, dass fast für alle Eigenschaften die Bandbreite von Antwortmöglichkeiten, d.h. von Rängen ausgeschöpft wurde. Dies bedeutet, dass die Geschosswirkung trotz des Gesamtranges von Platz 1 mitunter auf deutlich niedrigeren Rängen angesiedelt oder mit anderen Eigenschaften gleich gesetzt wurde sowie den modischen Eigenschaften einer Schutzweste vereinzelt die höchste Wichtigkeit zugewiesen wurde. Eine gewisse Individualität der Wichtigkeit von Eigenschaften einer Schutzweste ist damit zu schlussfolgern.

Tabelle 1: Resultierende Rangfolge der genannten Wichtigkeit verschiedener Eigenschaften einer Schutzweste

Rang	Eigenschaft	M	Min.	Max.	SD	N	Kurtosis	Schiefe
1	Schutzwirkung gegen Geschosse	1.37	1	11	.986	1597	22.300	4.079
2	Stichschutz	2.57	1	11	1.537	1593	8.339	2.691
3	Größe der geschützten Fläche	4.81	1	10	1.881	1585	-.663	.422
	Gewicht	4.82	1	11	1.625	1582	.024	.431
	Passform und Tragebequemlichkeit	4.90	1	10	1.774	1578	.052	.014
6	Beweglichkeit	5.25	1	11	2.019	1581	-.185	.231
7	klimatische Eigenschaften	6.51	1	11	1.809	1577	-.267	-.381
8	Splitterschutz	7.51	2	11	2.699	1558	-.983	-.418
9	Sichtbarkeit /Auffälligkeit	8.58	1	11	1.689	1582	2.235	-1.359
10	Pflege- und Reinigungsaufwand	9.32	1	11	1.354	1577	5.634	-.1667
11	modische Eigenschaften	10.17	1	11	1.287	1597	8.973	-2.431

### Gewalterfahrung

Die Befragten waren aufgefordert anzugeben, wie häufig sie im Dienst angegriffen oder bedroht wurden. Es wurde berichtet, dass sie im Mittel „1-2 mal in der gesamten Dienstzeit“ geschlagen oder getreten wurden. Im Mittel zwischen „1-2 mal in der gesamten Dienstzeit“ und „nie“ wurden die Befragten nach eigener Aussage mit Messer oder Schlagwaffen angegriffen. Beschossen wurden sie im Mittel „nie“. Die Androhung dieser Angriffe erfuhren die Befragten statistisch sehr signifikant jeweils häufiger (Wilcoxon-Tests:  $Z_1=-25.313^{**}$ ,  $Z_2=-21.950^{**}$ ,  $Z_3=-16.412^{**}$ ). Die berichteten Mittelwerte dürfen jedoch nicht darüber hinwegtäuschen, dass einzelne Polizeibeamte nach eigenen Angaben wesentlich häufiger entsprechend bedroht oder angegriffen wurden.

### Gefahrenkognition

Die Angst, in eine Situation zu geraten, die eskaliert, ist im Mittel nur minimal bis wenig ausgeprägt. Wobei jedoch zu beachten ist, dass 6.5% bzw. 1.4% angeben, ausgeprägte bzw. sehr große Angst davor zu haben. In eine eskalierende Situation zu geraten, wird im Mittel für „wahrscheinlich“ bis „ganz sicher“ gehalten, wobei nur wenigen dies „äußerst unwahrscheinlich“ (1.7%) oder „unwahrscheinlich“ (3.8%) erscheint. Die überwiegende Mehrheit (90.2%) der Befragten gibt an, dass Gewalt gegen Polizeibeamte in der letzten Zeit leicht bis stark zugenommen hat. Hingegen meinen 6.8%, dass sie relativ gleich geblieben ist.

Der Aussage „Wenn man eine Schutzweste trägt, so schießen Täter eher in den Kopf!“ wurde im Mittel „weder zu noch nicht“ zugestimmt bis „eher nicht“ zugestimmt. 90.5% gaben an, dass diese Aussage sie in ihrem Trageverhalten nicht beeinflusse.

### Funktionalität der Schutzweste

Die Teilnehmer wurden befragt nach ihrer Beurteilung des Schutzes, den die geschützte Fläche im Allgemeinen, gegen die häufigsten Geschosse, gegen Spezialgeschosse und gegen Messer bietet. Der Allgemeine Schutz wurde im Mittel für ausreichend bis sehr gut gehalten. Der Schutz gegen die häufigsten Geschosse wurde leicht, aber statistisch signifikant besser eingeschätzt (Wilcoxon-Test:  $Z=-7.137^{**}$ ). Für jeweils etwas schlechter (Wilcoxon-Test  $Z_1=-$

24.783<sup>\*\*</sup>,  $Z_2 = -25.373^{**}$ ), nämlich mittelmäßig bis ausreichend wurde der Schutz gegen Messer beurteilt. Am signifikant schlechtesten schnitt der Schutz gegen Spezialgeschosse in der Beurteilung ab: Er wurde für ungenügend bis mittelmäßig gehalten (Wilcoxon-Test  $Z_1 = -31.455^{**}$ ,  $Z_2 = -31.821^{**}$ ,  $Z_3 = -19.972^{**}$ ). Weiterhin sollte angegeben werden, inwieweit die durch die Schutzweste geschützte Fläche als ausreichend empfunden wird. Im Mittel wurde sie für „ausreichend“ bis „ein wenig zu klein“ eingeschätzt. 1.5% gaben an, dass die Schutzweste ein bisschen oder sogar erheblich kleiner sein sollte, während 4.8% sie erheblich vergrößert sehen wollten.

### Tragekomfort

Die Teilnehmer wurden gebeten, die Einschränkung der Beweglichkeit durch das Tragen einer Schutzweste, die Passform, klimatische Merkmale und modische Eigenschaften der Schutzweste zu beurteilen. Im Mittel wurde angegeben, dass ein Tragen der Schutzweste die Beweglichkeit geringfügig bis deutlich hemmt. 4.6% meinten nicht eingeschränkt zu werden, während 7.3% die Einschränkung als sehr stark mitteilten. 35.8% gaben an, wegen dieses Ausmaßes an Beweglichkeitseinschränkung die Schutzweste seltener zu tragen. 56.7% äußerten, davon unbeeinflusst zu sein. 1.6% fühlten sich durch die Intensität der Beweglichkeitseinschränkung zu einem häufigeren Tragen veranlasst.

Die Passform der Schutzweste wurde im Mittel als „neutral“ bis „eher schlecht“ beurteilt. 7.5% meinen, dass sie sehr schlecht sei, während 12.7% sie für eher gut oder sogar sehr gut (3.2%) hielten. Im Mittel fühlen sich die Befragten von der Passform geringfügig bis deutlich beeinflusst.

Die klimatischen Eigenschaften wurden im Mittel als „eher schlecht“ angesehen, wobei 21.4% diese sogar als sehr schlecht bewerteten. Nur 2.3% hielten die klimatischen Merkmale der Schutzweste für „eher gut“ oder 0.5% sogar für „sehr gut“.

### Modische Eigenschaften

Die modischen Eigenschaften wurden im Mittel als „neutral“ bis „eher schlecht“ bewertet. „Eher gut“- und „sehr gut“- Bewertungen traten nur in 1.8% bzw. 0.5% der Fälle auf. „Sehr schlecht“ fiel in 14.5% der Fälle das Urteil aus. Im Mittel meinten die Befragten, „überhaupt nicht“ durch dieses Attribut beeinflusst zu werden.

### Aufwand

Den Aufwand, der notwendig ist, um eine Schutzweste im Dienst tragen zu können, geben die Befragten im Mittel mit „geringfügig“ an. Dabei beschreiben diesen Aufwand 25.3% als nicht vorhanden, hingegen 16.6% als deutlich bzw. 3.5% als erheblich.

### Sicherheitskultur

„Ob eine Vorschrift, die das Tragen von Schutzwesten im Dienst vorschreibt, ihr Trageverhalten ändern würde“, wurde im Mittel mit „weiß nicht“ bis „eher nicht“ beantwortet. Als Hintergrund zur Einschätzung der Angaben der Frage sollten die Antworten unbedingt im Zusammenhang mit dem Trageverhalten berücksichtigt werden (siehe Abbildung 1).

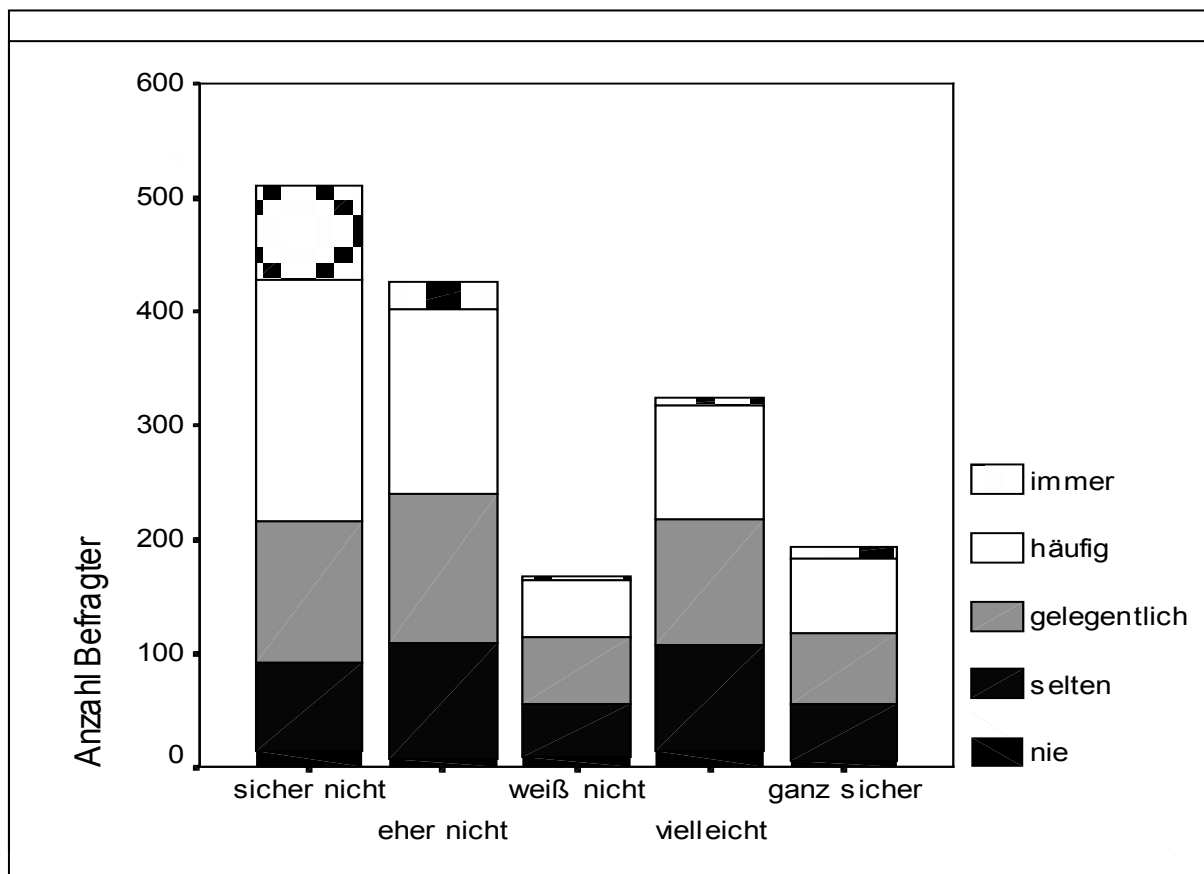


Abbildung 1: Angegebene Bereitschaft, das Trageverhalten auf Grund einer Vorschrift zu ändern

Im Mittel wurde angegeben, dass die meisten Kollegen gelegentlich bis häufig eine Schutzweste tragen. Es finden sich dabei 3% Nennungen von der Häufigkeit „nie“, 22.5% „selten“, 29.4% „gelegentlich“, 35.3% „häufig“ und 7.7% „immer“. Dabei geben 87.0% der Teilnehmer an, sich nicht von den Kollegen im Trageverhalten beeinflussen zu fühlen. 2.0% meinen, auf Grund der Kollegen die Schutzweste seltener zu tragen, während 7.6% bemerken, dass Sie durch ihre Kollegen zu einer höheren Tragerate animiert werden.

Als vom Vorgesetzten im Tragen der Schutzweste unbeeinflusst beschreiben sich 94.0%. Nur 2.0% tragen angeblich auf Grund des Einflusses des Vorgesetzten ihre Schutzweste häufiger, während sich 0.8% durch den Einfluss des Vorgesetzten veranlasst sehen, ihre Schutzweste seltener zu tragen.

### Zusammenhänge

Es soll nun versucht werden, einen allgemeinen Zusammenhang zwischen Bewertungen bzw. Aussagen und dem angegebenen Trageverhalten zu bestimmen. Zu beachten ist, dass das dabei verwendete korrelative Verfahren (Kendalls-Tau-b) eine kausale Interpretation verbietet und lediglich statistische Zusammenhänge, d.h. gemeinsame Variation bzw. die Ähnlichkeit der Rangordnung beschreibt.

Auf Grund der Unterschiede der Angaben von Befragten unterschiedlicher Polizeibereiche und Teilnehmer mit verschiedenen dienstlichen Schwerpunkt-Tätigkeiten, sollen Zusammenhänge zwischen dem Trageverhalten und den hier erhobenen potenziell tragerelevanten Faktoren für die Substichprobe der Schutzpolizisten mit überwiegendem Streifendienst (N=1165) und die Gesamtstichprobe bestimmt werden.

Insgesamt fanden sich für die Antworten im Fragebogen geringe bis mittlere Zusammenhänge mit dem berichteten Trageverhalten (siehe Tabelle 2). Augenscheinlich liegen die

Zusammenhangskoeffizienten der Substichprobe und der Gesamtstichprobe meist sowohl in gleicher Richtung als auch in gleicher Größenordnung. Auf gesonderte Betrachtungen wird deshalb verzichtet.

Die Tragehäufigkeit variiert vor allem mittelmäßig mit der Beweglichkeitseinschränkung durch die Schutzweste, mit dem angegebenen Einfluss dieser Beweglichkeitseinschränkung, mit der Passform, mit dem subjektiven Aufwand, der notwendig erscheint, um eine Schutzweste im Dienst tragen zu können, mit der Bereitschaft, auf Grund der Vorschrift das Trageverhalten zu ändern, und der Häufigkeit des Tragens der Schutzweste bei den meisten Kollegen (siehe Tabelle 2). Das angegebene Trageverhalten scheint also vor allem mit Faktoren des Tragekomforts, mit der Verfügbarkeit bzw. dem subjektiven Aufwand und mit Aspekten der Sicherheitskultur zusammenzuhängen (dies ist auf Grund des Designs der Studie nicht kausal interpretierbar). Aber es finden sich auch - zwar nur in geringer Größenordnung - statistische Zusammenhänge mit Gewalterfahrung, Fortbildungsaktivität, Gefahrenkognition und modischer Bewertung. Nicht gefunden wurden selbst geringfügige Zusammenhänge ( $\tau\text{-}b < .1$ ) der Funktionalitätsüberzeugung mit der Tragehäufigkeit.

Tabelle 2: Statistisch sehr signifikante Zusammenhänge zwischen der angegebenen Tragehäufigkeit und verschiedenen Angaben im Fragebogen (Kendalls  $\tau$ /Tau-b), die einen Wert von 0.1 übersteigen

	<b>Schutzpolizisten mit überwiegend Streifendienst</b>		<b>alle Befragten</b>	
	$\tau$	N	$\tau$	N
Angriffe im Dienst: Schläge angedroht	.107	1139	.124	1633
Genutzte Fortbildung: Schießen	.101	1109	.134	1588
Wahrscheinlichkeit für eskalierende Situation	.109	1145	.163	1647
Veränderung der Gewalt gegen Polizeibeamte	.102	1146	.112	1645
Einfluss „wenn man Schutzweste trägt, schießt Täter eher in den Kopf“	.113	1143	.110	1633
Beweglichkeitseinschränkung durch die Schutzweste	-.374	1144	-.351	1631
Einfluss der Beweglichkeitseinschränkung durch die Schutzweste	.592	1138	.511	1622
Bewertung Passform	.260	1100	.220	1538
Beeinflussung durch die Passform	-.363	1090	-.313	1522
Beeinflussung durch die klimatischen Eigenschaften	-.232	1069	-.194	1477
Bewertung modische Eigenschaften	.165	911	.155	1296
Beeinflussung durch die modischen Eigenschaften	-.189	974	-.180	1361
Notwendiger Aufwand um Schutzweste zu tragen	-.393	1115	-.398	1593
Änderung des Trageverhaltens durch Vorschrift	-.408	1148	-.368	1638
Häufigkeit des Tragens der Schutzweste bei den meisten Kollegen	.351	1144	.378	1633
Wichtigkeit der Schutzwirkung gegen Geschosse	-.126	1117		
Wichtigkeit der Größe der geschützten Fläche	-.130	1114		
Wichtigkeit der modischen Eigenschaften	.124	1105	.117	1571

## Diskussion

Im Vergleich zu Angaben aus den USA geben deutsche Polizistinnen und Polizisten an, häufiger eine Schutzweste zu tragen. Dabei muss jedoch berücksichtigt werden, dass die Polizistenmorde des Jahres 2000 mit der in diesem Zusammenhang geführten Diskussion um Schutzwesten zu einer kurzfristigen Erhöhung der Tragehäufigkeit geführt haben kann, was die Angaben vieler Befragter auch bestätigen. Es bleibt abzuwarten, ob dieser Effekt als Strohfeder bald seine Wirkung verliert oder aber eine stabile Änderung bewirkt hat. Es darf weiterhin nicht vergessen werden, dass die Tragehäufigkeit von den Befragten selbst

angegeben wurde. Inwieweit die Antworten Tendenzen unterliegen oder die Realität exakt abbilden, bleibt unbeantwortet.

Es finden sich teilweise starke individuelle Differenzen aber auch große Ähnlichkeiten bezüglich der Erfahrungen, Bewertungen und Präferenzen im Zusammenhang mit dem Tragen einer Schutzweste bei der Polizei. Ein standardmäßiges Vorgehen bzgl. des Tragens einer Schutzweste erscheint deshalb ungeeignet. Vielmehr kann es sinnvoll sein, individuelle Präferenzen und Eigenschaften zu berücksichtigen. In diesem Rahmen ist es vorstellbar, dass die Schutzweste als individueller Ausrüstungsgegenstand eingesetzt, d.h. persönlich zugewiesen und zugeschnitten, mit einer erhöhten Tragerate einhergeht. Dafür sprechen vor allem die Korrelationen der Tragehäufigkeit mit dem Tragekomfort und dem subjektiv notwendigen Aufwand.

Erstaunlich ist, dass die Tragehäufigkeit mit modischen Bewertungen - wenn auch nur gering - korreliert, da gerade die Aufnahme dieses Faktor in den Fragenkatalog zu vielen missbilligenden Bemerkungen und Kommentaren geführt hat. Auch hier sollte sich keinesfalls zu einer kausalen Interpretation entschlossen werden. Genauere Erkenntnisse zum Einfluss dieses wie auch der anderen Faktoren benötigen ein experimentelles Vorgehen. Umgekehrt scheint es aber auch falsch zu sein, modische Aspekte bzgl. des Tragens einer Schutzweste zu ignorieren.

Die Ergebnisse lassen vermuten, dass Sicherheit insbesondere durch Schutzwesten in der Organisationskultur noch keinen festen Platz hat: Während das Trageverhalten von Kollegen - eventuell durch gruppenspezifische Phänomene vermittelt - mit dem Trageverhalten der Befragten zusammenhängt, scheint dies noch nicht systematisch durch Kollegen oder Vorgesetzte genutzt zu werden (siehe Angaben zum Einfluss von Kollegen und Vorgesetzten auf das Trageverhalten).

Ein Tragen der Schutzweste durch Vorschriften quasi erzwingen zu können, erscheint auf Grund der Ergebnisse fraglich: Nur 31.5% geben an, vielleicht oder ganz sicher ihr Trageverhalten in Folge einer Vorschrift zu ändern.

Dass sich etliche mittlere und geringe statistische Zusammenhänge zwischen verschiedenen Faktoren und dem Trageverhalten gefunden haben, zeigt die Komplexität des Tragens dieser Schutzausrüstung. Ob die Schutzweste getragen wird, scheint im Allgemeinen nicht von einer Eigenschaft alleine abzuhängen. Vielmehr ist vermutlich ein gewisses Zusammenspiel notwendig. Ebenso scheinen z.B. schlechte Bewertungen unterschiedlicher Eigenschaften nicht hinreichend für eine geringe Tragehäufigkeit zu sein. Es ist eher anzunehmen, dass verschiedene Personen eine Schutzweste z.B. trotz klimatisch ungünstiger Eigenschaften häufig oder immer tragen. Ebenso kann z.B. eine Person sich gegen das Tragen einer Schutzweste sträuben, obwohl sie überzeugt ist, Gewalt gegen Polizeibeamte habe stark zugenommen, und sie sich ganz sicher ist, in eine Situation zu geraten, die eskaliert.

Dass einige theoretische Zusammenhänge nicht statistisch gefunden werden konnten oder nur eine geringe Größe besaßen, muss nicht bedeuten, dass sie im Allgemeinen keinen Einfluss auf das Trageverhalten haben. Insbesondere sei hier die Funktionalitätsüberzeugung genannt. Es ist theoretisch wie logisch einleuchtend, dass die Überzeugung, eine Weste schütze gut oder sehr gut, eher zu einem Tragen führt als die Annahme, die Weste könnte nicht oder nur schlecht schützen. Ein entsprechender statistischer Zusammenhang konnte hier jedoch nicht gefunden werden. Deshalb ist aber nicht auf eine fehlerhafte Theorie zu schließen, vielmehr liegt hier eher ein methodisches Artefakt vor: Zusammenhangsanalysen benötigen ein Mindestmaß an Variation in den Ausprägungen der Merkmale, d.h. es müssen alle Antwortkategorien in gewissem Umfang gewählt werden. Sammeln sich aber die Antworten wie bei der Funktionalitätsüberzeugung in einem sehr engen Bereich mit nur geringer

Abweichung davon, kann kein statistisch bedeutsamer Zusammenhang entdeckt werden. Dies kann für einige Fragen hier angenommen werden.

## Schlussfolgerungen

Die hier vorgestellten Ergebnisse stellen einen ersten Ausschnitt einer Pilotstudie dar, die zur systematischen Erforschung dieses Themenbereiches beitragen möchte, um die Sicherheit beim polizeilichen Einsatz zu erhöhen. Deshalb kann und muss ein Schlussfolgern nur sehr zurückhaltend und allgemein erfolgen. Nicht vergessen werden sollte auch, dass die Repräsentativität der Stichprobe ungeklärt ist und damit Generalisierungen auf größere Bereiche wenn überhaupt, dann nur mit entsprechender Vorsicht, möglich sind. In diesem Zusammenhang sei auch erwähnt, dass die Stichprobe auf Grund der Freiwilligkeit der Teilnahme selektiv bzgl. relevanter Eigenschaften zusammengesetzt sein kann.

Zweifellos ist eine Befragung immer anfällig gegen motivational gesteuerte Antworten, Antworten im Sinne der (sozialen oder hier polizeilichen) Erwünschtheit oder mangelhafte Selbstkenntnis und -auskunft. Ob letztendlich die angegebenen Tragehäufigkeiten die Realität widerspiegeln und die Beurteilungen und Angaben mit der Wirklichkeit harmonieren und ursächlich für die Tragerate sind, kann nur durch Beobachtungen und experimentelle Studien geklärt werden. Nichts desto trotz liefert diese Erhebung umfassendes Material und Erkenntnisse, die auf dieses Ziel und damit die systematische Verbesserung der Sicherheit des Polizisten im Dienst hinarbeiten. Es ist unbedingt zu fordern, dass das Trageverhalten entsprechend weiter erforscht wird.

Fakt ist, dass der ballistische Schutz einer Schutzweste die Wirkung von Geschossen mindern kann und damit einen potenziellen Beitrag zur Sicherheit leistet. Um diesen nicht zu reduzieren, aufzuheben oder etwa ins Gegenteil zu kehren, ist es notwendig, dass die Schutzweste in geeigneter Weise getragen und in ihrer Wirkung nicht durch riskanteres Verhalten kompensiert wird (vgl. Theorie der Risikokompensation bzw. Risiko-Homöostase, Wilde, 1981). Da das Trageverhalten mit einigen Faktoren zusammenzuhängen scheint, sind diese weiter zu erforschen - Strobel (1994) sieht als Voraussetzung für Motivationsprogramme zur Steigerung des Trageverhaltens die Erkundung der Gründe mangelnden Einsatzes von Körperschutzmitteln - und kann es sinnvoll sein, sie in entsprechende Entwicklungen, Beschaffungs- und Einsatzmaßnahmen einfließen zu lassen. Weiterhin ist vorstellbar, dass neben der Ausgabe von Schutzwesten begleitend eine fundierte psychologische Schulung und (Um-)Organisation auf Individuums-, Gruppen- sowie Organisationsebene ein Trageverhalten positiv beeinflussen kann (vgl. Zimolong, 1991, S. 251ff.). Ein „wortloses“ Übergeben oder Bereitstellen dieses Ausrüstungsgegenstandes ohne Beachtung vieler psychologischer und individueller Faktoren kann nicht systematisch zu höherer Sicherheit eines Polizisten beitragen. Man überließe es dem Zufall, welche Wirkung letztendlich die Schutzweste hätte.

## Literatur

- Bechstein, D. (1998). Die Überlebenden mahnen zur Wachsamkeit. *Polizei heute*, 27 (4), 132-134.
- Berka, W. & Sauter, S. (1995). Kugelschutzwesten - Je nach Lage tragen? *Kriminalistik*, 8-9, 585-587.
- Bortz, J. & Döring, N. (1995). *Forschungsmethoden und Evaluation*. Berlin: Springer.
- Bortz, J., Lienert, A. & Boehnke, K. (1990). *Verteilungsfreie Methoden in der Biostatistik*. Berlin: Springer.



- Brand, M. I. (1993). *Why police officers don't wear body armor*. Du Pont Body Armor Symposium.
- Burkardt, F. (1981). *Information und Motivation zur Arbeitssicherheit*. Wiesbaden: Universum-Verl.-Anst..
- Damm, H.-R. (2000). Zum Schutz vor lebensbedrohenden Angriffen. *Deutsche Polizei*, 12, 20-22.
- Federal Bureau of Investigation (1994). *Killed in the Line of Duty*. Boulder, Colorado: Paladin Press.
- Hoyos, C. Graf (1980). *Psychologische Unfall- und Sicherheitsforschung*. Stuttgart: Kohlhammer.
- Hoyos, C. Graf, Bernhardt, U., Hirsch, G. & Arnhold, T. (1991). Vorhandenes und erwünschtes sicherheitsrelevantes Wissen in Industriebetrieben. *Zeitschrift für Arbeits- und Organisationspsychologie*, 35 (2), 68-76.
- Künzler, K. & Grote, G. (1996). SAM – ein Leitfaden zur Bewertung von Sicherheitskulturen in Unternehmen. In B. Ludborz, H. Nold & B. Rüttinger (Hrsg.), *Psychologie der Arbeitssicherheit. 8. Workshop 1995* (S. 78-93). Heidelberg: Asanger.
- Musahl, H.-P. (1997). *Gefahrenkognition*. Heidelberg: Asanger.
- Musahl, H.-P., Groß-Thomas, C. & Müller-Gethmann, H. (1994). Gefahrenkenntnis und Arbeitssicherheit - Entwicklung und Evaluation eines „top-down“-Programms. In F. Burkardt & C. Winklmeier (Hrsg.), *Psychologie der Arbeitssicherheit. 7. Workshop 1993* (S. 559-571). Heidelberg: Asanger.
- Panek, C. (1998). Schutzwesten aus Kevlar haben noch nie versagt. *Polizei heute*, 27 (4), 120-121.
- Pokojewski, B. (1998). Ballistischer Schutz in den USA - Maßstab für Deutschland? *Polizei heute*, 27 (4), 125-132.
- Schmidt, D. (1990). *Arbeitssicherheit. Physiologische, psychologische, organisatorische und rechtliche Grundlagen*. Heidelberg: R. v. Decker's Verlag.
- Skiba, R. (1997). *Taschenbuch der Arbeitssicherheit*. Bielefeld: Erich Schmidt.
- Strobel, G. & Wittmann, M. (1998). Motivation zum Tragen persönlicher Schutzausrüstung – Qualifizieren durch Leitfäden für die betriebliche Praxis. In H. v. Benda & D. Bratge (Hrsg.), *Psychologie der Arbeitssicherheit. 9. Workshop 1997* (S. 141-145). Heidelberg: Asanger.
- Strobel, G. (1994). Persönliche Schutzausrüstung: Motive mangelnder Tragebereitschaft und Maßnahmen zur Förderung der Akzeptanz. In F. Burkardt & C. Winklmeier (Hrsg.), *Psychologie der Arbeitssicherheit. 7. Workshop 1993* (S. 623-630). Heidelberg: Asanger.
- Van Zijl, N. (1998). Über die Klinge gesprungen. *Polizei heute*, 27 (4), 142-143.
- Weber, P. (1994). Die Beurteilung des Tragekomforts und der Trageeigenschaften von persönlichen Schutzausrüstungen aus psychologischer Sicht. In F. Burkardt & C. Winklmeier (Hrsg.), *Psychologie der Arbeitssicherheit. 7. Workshop 1993* (S. 603-614). Heidelberg: Asanger.
- Wilde, G. J. S. (1981). Objektive und subjektive Sicherheit im Wechselwirkungsbereich Fahrer-Straße: Die Theorie der Risiko-Homöostase. In Daimler Benz AG. (Hrsg.), *Verkehrssicherheit: Fahrverhalten und die Bedeutung der Risikotheorien für die Sicherheit im Straßenverkehr*. Stuttgart: Daimler Benz AG.
- Zapf, D. & Dormann, C. (2001). Gesundheit und Arbeitsschutz. In H. Schuler (Hrsg.), *Lehrbuch der Personalpsychologie* (S. 559-587). Göttingen: Hogrefe.
- Zeitlin, L. R. (1994). Failure to Follow Safety Instructions: Faulty Communication or Risky Decisions? *Human Factors*, 36 (1), 172-181.
- Zimolong, B. (1991). Ein psychologisches Qualifizierungskonzept für Führungskräfte im Arbeits- und Umweltschutz. In H. Schmöle (Hrsg.), *Lernziel Sicherheit: psychologische Beiträge zur Aus- und Fortbildung im Arbeitsschutz* (S. 7-23). Bonn: Deutscher Psychologen-Verlag.

<sup>1</sup> Repräsentativität (vgl. Bortz & Döring, 1995, S. 371ff.) einer Stichprobe ist nur in Bezug zu einer Population und nicht allein durch Anlage der Studie oder den Stichprobenumfang beurteilbar. Eine diesbezügliche Bewertung der hier befragten Stichprobe setzt voraus, dass sowohl für die Befragung relevante Eigenschaften als auch die Population bekannt ist. Dies ist aber nicht der Fall und z.Zt. auch nicht vorstellbar: Die Kenntnis, wen und wann die Schutzwestenproblematik angeht, ist genau eines der strittigen und bisher unerforschten Gebiete. Die relevanten Eigenschaften und Merkmale der Population, hinsichtlich derer die Stichprobe repräsentativ, d.h. als der Gesamtheit ähnlich sein soll, sind noch unklar. Gerade deshalb wurde diese Studie durchgeführt. Somit ist ein Design zur systematischen Erlangung einer Repräsentativität unmöglich. Repräsentativität mittels Zufallsziehung der Stichprobe scheint nicht hinreichend, da von einer selektiven Erreichbarkeit der Studienobjekte ausgegangen werden muss. Eine bewusst vorsichtig vollzogenen Interpretationen und Generalisationen kann dem Rechenschaft tragen.

<sup>2</sup> Die E-Mailbefragung fand im November/Dezember 2000 statt, die persönliche Ansprache hessischer Dienststellen erfolgte im Januar 2001.

<sup>3</sup> Es konnte nicht von allen via E-Mail erreichten Dienststellen der Fragebogen entschlüsselt werden. Ein klassischer Versand via Post wurde entsprechend angeboten und auch teilweise genutzt. Es zeigte sich neben einer engagierten Weiterleitung als erste Reaktion eine teilweise ministeriell/institutionell angeordnete Teilnahmeverweigerung.

<sup>4</sup> Gaben die Befragten für mehrere Merkmale einen gleichen Rang an, wurde ein mittlerer gemeinsamer Rang für die Auswertung berechnet (=Rangaufteilung, vgl. Bortz, Lienert & Boehnke, 1990, S.69).

<sup>5</sup> Auf die Frage nach der Wichtigkeit und an anderer Stelle nach der Bewertung der modischen Eigenschaften einer Schutzweste reagierten einige der Befragten mit Bemerkungen und Beschimpfung der Studie.

# Ergebnisse einer Bürgerbefragung zur Akzeptanz von polizeilichen Kontroll- und Eigensicherungsmaßnahmen

*Clemens Lorei*

*Redaktion Polizei und Wissenschaft, Frankfurt*

Polizeiliches Handeln im Kontakt mit Bürgern und Rechtsbrechern erfordert mitunter eine Gratwanderung zwischen Bürgerfreundlichkeit und notwendigen Maßnahmen, um die Sicherheit der eingesetzten Polizisten zu gewährleisten. Infolge der Tötung von acht Polizisten im Jahr 2000 wurden sogenannte "amerikanische Verhältnisse" erwogen, das heißt Kontrollbedingungen und Maßnahmen, die eine verstärkte Sicherheit der eingesetzten Polizeikräfte bewirken sollen. Unabhängig von einer taktischen Bewertung dieser Eigensicherungsmaßnahmen scheint es im Sinne sowohl der Bürgerfreundlichkeit als auch der Kooperationsbereitschaft von bürgerlicher Seite aus wichtig, die Akzeptanz von entsprechenden polizeilichen Kontroll- und Eigensicherungsmaßnahmen zu erfragen. Es wurde in einer nicht repräsentativen Befragung von 431 hessischen Bürgerinnen und Bürgern versucht, mittels Fragebogen eine Bewertung polizeilicher Kontroll- und Eigensicherungsmaßnahmen, die Einstellung gegenüber der Polizei und die psychologische Reaktanz zu erfassen. Die Ergebnisse zeigen, dass Bewertungen verschiedener Maßnahmen interindividuell unterschiedlich variieren. Die mittleren Beurteilungen zeigen trotz überwiegend positiver Einstellung gegenüber der Polizei teilweise deutliche Ablehnung. Fast alle Kontroll- oder Eigensicherungsmaßnahmen bei Fußgängerkontrollen werden geringfügig, aber statistisch signifikant negativer bewertet als bei Fahrzeugkontrollen. Insgesamt zeigten sich höchstens geringe korrelative Zusammenhänge zwischen verschiedenen Maßnahmen und der Einstellung gegenüber der Polizei und der Reaktanzneigung. Es wird deshalb daraus geschlossen, dass die polizeilichen Maßnahmen nicht auf Grund einer allgemeinen Haltung bewertet werden. Der Versuch, die verschiedenen Maßnahmen faktorenanalytisch zu gruppieren, legt nahe, dass neben einer starken allgemeinen Bewertung vier Gruppen von polizeilichen Handlungsweisen existieren, die sich inhaltlich mehr oder minder gut interpretieren lassen und Rückschlüsse über die Beurteilungen erlauben.

## **Einleitung**

Der Polizeiberuf bringt sowohl eine enorme Aufgabenvielfalt mit sich als auch vielfältige Kontakte zu unterschiedlichen Personengruppen. Da die Aufgabe der Polizei u.a. in der Gefahrenabwehr und Strafverfolgung besteht, ist mit einer gewissen Regelmäßigkeit damit zu rechnen, dass der Polizist in potenziell gefährliche, mitunter lebensbedrohliche Auseinandersetzungen gerät. Dabei kann ein Polizeibeamter nicht - im Gegensatz zu den klassischen Annahmen Ernst Kretschmers - an der äußeren Erscheinung seines Gegenübers erkennen, ob dieser ein gewaltbereiter Täter oder ein friedfertiger (Normal-)Bürger ist. Um mit einem Mindestmaß an Sicherheit polizeiliche Aufgaben erledigen zu können, ist deshalb ein präventives (und auch reaktives) Verhalten notwendig, das Eskalationen vorbeugt oder im Falle gewalttätiger Interaktionen es den Polizeibeamten erlaubt, alles notwendige zur Sicherheit des Bürgers und natürlich auch des eigenen Lebens zu tun. Aus diesem Grund hat der Gesetzgeber die Polizei mit einer entsprechenden Gesetzesgrundlage ausgestattet, die polizeiliches Handeln u.a. zur Sicherheit der Polizeibeamten, der sog. Eigensicherung, erlaubt. In Folge der Ereignisse des vergangenen Jahres, in dem mit acht ermordeten Polizisten ein trauriger „Rekord“ nach den durch Terrorismus geprägten 80er Jahren aufgestellt wurde, wurden häufig sog. „amerikanische Verhältnisse“ diskutiert (vgl. Hasselmann, 2000). Unabhängig davon, welche Verhältnisse in der USA im Bereich von Gewalt und bei der polizeilichen Arbeit tatsächlich vorliegen und welche taktischen Vorteile entsprechende Maßnahmen der Eigensicherung dem Beamten auf der Straße verleihen, ist klar, dass diese in Bezug zur Bürgerfreundlichkeit zu setzen sind (vgl. Schmalzl, 1996a). Dabei gilt, dass die empfundene Fairness bei solchen Polizei-Bürger-Interaktionen unabhängig vom inhaltlichen Ausgang des Polizei-Bürger-Kontaktes ist und sogar wichtiger als dieser sein kann, selbst wenn das Polizeigegenüber für eine Ordnungswidrigkeit oder

Gesetzesübertretung zur Verantwortung gezogen wird (Tyler & Folger, 1980). Dies gilt nach Tyler & Folger (1980) insbesondere für Fahrzeugkontrollen. Sie schlussfolgern, dass eine faire Behandlung durch die Polizei den Eindruck negativer Folgen dieses Kontaktes, z.B. in Form einer Bestrafung, reduzieren kann. Entsprechend erfordert polizeiliches Handeln im Kontakt mit Bürgern und Rechtsbrechern eine Gratwanderung zwischen Sicherheit und Bürgerfreundlichkeit, denn viele der über die Polizistenmorde eben noch empörten Bürger befürchten wenig später, wie Schwerverbrecher behandelt oder sogar versehentlich erschossen zu werden. So zeigte eine Online-Befragung des Tagesspiegels (3.3.2001), dass 60.4% von 763 beteiligten Personen dagegen sind, dass Verkehrskontrollen künftig zum Schutz von Beamten mit gezogener Waffe durchgeführt werden sollten (Frage: „Soll die Polizei Autofahrer mit gezückter Waffe kontrollieren?“). Kommentare zu dieser Maßnahme fallen höchst unterschiedlich aus (siehe Tabelle 1).

Tabelle 1: Pro & Contra Meinungen der Online Befragung des Tagesspiegels (3.3.2001)

- Sollen noch mehr Polizisten von besoffenen, zugehörnten oder sonstwie kaputten Typen erschossen werden? Sie sind bei uns doch sowieso nur Freiwild. Autofahrer sollten wie in den USA aussteigen und die Hände aufs Dach legen müssen, ich würde das gerne tun für die Sicherheit der Polizisten!
- Am besten gleich schießen und danach kontrollieren
- Sicher ist es für Polizisten nicht unbedingt ungefährlich, eine Verkehrskontrolle durchzuführen, aber so gefährlich, dass eine gezückte Waffe notwendig wäre, ist es wohl kaum. Und dass eine gezückte Waffe es leichter macht, auch mal vorschnell abzudrücken, liegt doch auf der Hand...das senkt nur die Hemmschwelle...auf beiden Seiten.
- In vielen anderen Ländern wird es zu Schutz der Polizisten praktiziert. Es kann nicht angehen das die Menschen die uns beschützen sollen (z.B. vor alkoholisierten Autofahrern) dafür mit ihrem Leben pokern müssen. Unsere Polizei hat, obwohl ich sie nicht besonders mag, auch ein Recht darauf, unversehrt aus einem Einsatz nach Hause zur Familie kommen zu können.
- Unglaublich die Vorstellung, ein Polizist kommt mit gezogener Waffe auf mein Auto zu, in dem ich mit meiner 10-jährigen Tochter sitze! Es tut mir um jeden Polizisten leid, der bei einer Verkehrskontrolle ums Leben kam, aber das kann nicht die Reaktion darauf sein!

Ziel dieser Studie ist, in einer ersten explorativen Studie detailliert zu erfahren, wie Bürger über verschiedene Eigensicherungs- und Kontrollmaßnahmen denken und mit welchen psychologischen Faktoren dies im Zusammenhang steht. Erkenntnisse hierbei könnten Teil der Grundlage sein, die Sicherheit von Polizeibeamten durch entsprechende polizeiliche Eigensicherung zu gewährleisten und gleichzeitig Zufriedenheit des Bürgers im Sinne von Bürgerfreundlichkeit zu erreichen.

## Theorie

### Einstellung

Einstellung stellt ein Merkmal einer Person dar, das eine überdauernde Ausrichtung auf Objekte meint und sich in dessen Gedanken, Gefühlen, Bewertungen und Verhaltensabsichten bzw. Intentionen auf ein soziales Objekt richtet (Mummendey, 1995, S. 30). Einstellungen können entsprechend als Neigung angesehen werden, auf eine bestimmte Klasse von Reizen mit einer bestimmten Klasse von Reaktionen zu antworten (Rosenberg & Hovland, 1960 [zitiert nach Stahlberg, Frey, 1992, S. 145f.]). Dabei werden die Reaktionsklassen in affektiv,

kognitiv und verhaltensmäßig spezifiziert. Die Einstellung vermittelt zwischen vorangehenden externen Reizen und darauf bezogenem Verhalten (Stahlberg & Frey, 1992, S. 145). Dabei zählen Six und Eckes (1996) zu den Funktionen von Einstellungen die selektive Aufnahme und Verarbeitung von Informationen, die kognitive Orientierung in einer komplexen sozialen Umwelt, die Selbstbewertung und Selbstinterpretation, die Vorhersage und Änderung von Verhalten und die Regulation der Beziehung zwischen Individuum und Gesellschaft. Die Konsistenz zwischen Einstellung und Verhalten wird heftig diskutiert. Einerseits nehmen Kritiker des Einstellungsbegriffs eine seltene Konsistenz von Einstellung und Verhalten zum Anlass, eine Verhaltensprognose auf Grund von Einstellungen für unmöglich zu halten (vgl. Thomas, 1991, S. 161f.). Andererseits kommen verschiedene Meta-Analysen zum Zusammenhang von Einstellung und Verhalten zu einer starken Beziehung zwischen Einstellung und Verhalten (Six & Eckes, 1994) und halten sogar das Konzept der Einstellung für den wichtigsten Prädiktor von Verhalten (Six & Eckes, 1996). Dabei wird Verhalten nicht ausschließlich durch Einstellungen determiniert, vielmehr spielen verschiedene weitere Faktoren auch eine Rolle (vgl. Mummendey, 1988; Six, 1998).

Einstellungen des Bürgers zur Polizei wurden seit 1980 ausgehend von verschiedenen Fragestellungen mehrfach erhoben. Naether (1980) fand bei seiner Befragung von über 1000 Nordrhein-Westfalen allgemein eine neutral bis eher positiv wertende Haltung im Sinne einer rein gefühlsmäßigen Sympathie, die er als Ausdruck der Ambivalenz gegenüber der Polizei beurteilt. Dabei stimmten die Befragten der Aussage zu, dass Polizisten einen sehr gefährlichen und risikoreichen Beruf hätten. Schüller (1990) fand in seiner Befragung von 452 Personen eine Ambivalenz gegenüber der Polizei. Während eine allgemeine Kontrolle der Bürger durch die Polizei überwiegend abgelehnt wurde, solle diese verstärkt zur Bekämpfung organisierter und politischer Verbrechen eingesetzt werden. Also „mehr Kontrolle, ohne kontrolliert werden zu wollen“ (Schüller, 1990, S. 295). Das EMNID-Institut (vgl. Emnid, 1988) stellte in der Zusammenfassung einer repräsentativen Befragung fest, das 78% der Bürger der Auffassung sind, die Polizei werde ihrer Aufgabe „die Bürger zu schützen und die öffentliche Sicherheit und Ordnung aufrechtzuerhalten“ gerecht (EMNID, 1988, S. 203). Murck (1989, S.28) stellte bei der überwiegenden Mehrheit derselben Befragung von 1987 (EMNID, 1988) fest, das die Einstellung zur Polizei „alles in allem“ positiv ausfällt und sieht darin „die bekannt positive Grundeinstellung zur Polizei beim weit überwiegenden Teil der Bevölkerung“ (Murck, 1989, S. 29) bestätigt. 1991 bestätigte sich die überwiegend positive Einstellung in den alten Bundesländern, auch wenn sie eine Niveauverschiebung zu verzeichnen hatte (Murck, 1992). Hingegen zeigten zu diesem Zeitpunkt die neuen Länder eine deutlich schlechtere Bewertung ihrer Polizei (gefragt wurde 1991 nach der Polizei des entsprechenden Bundeslandes). Hermanutz (1995) fand in seiner Befragung zur Zufriedenheit von 109 Bürgern, die einen Polizeikontakt hatten, dass das „Ansehen der Polizei allgemein“ überwiegend als durchschnittlich bis gut beurteilt wurde.

## Reaktanz

Die auf Brehm (1966) zurückgehende Reaktanztheorie bezieht sich im Wesentlichen auf freitliches Handeln von Personen und die Folgen, wenn dieser Freiraum eine (subjektive) Einschränkung erfährt. Es wird davon ausgegangen, dass Personen über eine gewisse Freiheit verfügen, Verhalten selbstbestimmend auszuführen. Empfindet eine Person die Wahlfreiheit bzgl. ihres Handelns als gefährdet oder sogar beschränkt, erwächst in ihr die Motivation, diese (potenzielle) Einengung aufzuheben. Dieser motivationale Zustand wird als psychologische Reaktanz bezeichnet und zielt auf die Wiederherstellung der verlorenen oder bedrohten Freiheit ab, was sich in unterschiedlichen Verhaltensweisen äußern kann. Entscheidend für die Reaktanzgenese ist dabei, welche Bedeutung der Entscheidungsfreiheit beigemessen wird, in

welchem Umfang die Freiheit eingeschränkt wird und mit welcher Intensität die freie Handlungswahl beeinträchtigt wird.

Als wesentliche Bedingung für das Auftreten von Reaktanz sehen Dickenberger, Kniech & Grabitz (1993, S. 244), dass die subjektiv eingeschränkte Person, die Freiheitseinschränkung als illegitim ansieht und sich ihr ausgeliefert fühlt. Nicht hingegen soll Reaktanz einsetzen, wenn die Einschränkung aus sozialen Gründen akzeptiert wird oder der einschränkenden Instanz diese Maßnahme zugebilligt wird. Über die momentane Freiheitsbedrohung hinaus kann eine Generalisierung dieser Einschränkung auf zukünftige und weitere Situationen zu einer Verstärkung der Reaktanz führen. Dabei scheint ebenso die Reaktanz zu verstärken, wenn Personen bzgl. einer Freiheitseinschränkung informiert oder vorgewarnt werden (Dickenberger et al. 1993, S. 247).

Als Folge von psychologischer Reaktanz kann die Person versuchen, ihre Freiheit durch subjektive Reaktionen und Verhalten (vgl. Brehm, 1972 [zitiert nach Dickenberger et al., 1993, S. 247]) wiederherzustellen. Während sich subjektive Reaktionen nicht in offenem Verhalten äußern, sondern u.a. in kognitiven Umstrukturierungen und Einstellungsänderungen bestehen, beinhalten die Verhaltenseffekte direkte Aktionen, wie z.B. Widerstand, aggressive Attacken oder Verlassen der Situation. Verhaltens-Effekte werden dabei häufig auf Grund sozialer Normen, gesellschaftlicher Ordnungen, persönlicher Gegebenheiten und natürlich aus Angst vor negativen Konsequenzen - insbesondere in Gegenwart „sanktionsmächtiger Zeugen“ (Dickenberger et al., 1993, S. 252) - unterlassen. Entsprechend finden die Reaktionen auf Freiheitseinschränkungen wohl häufiger in kognitiven Handlungen Niederschlag oder werden auf Zeitpunkte geschoben, die eine Freiheitswiederherstellung ermöglichen.

In polizeilichen Kontrollsituationen kann Reaktanz auf der Seite des Bürgers vor allem durch zwei Komponenten hervorgerufen werden: Erstens durch die Unterbrechung der Handlung des zu Kontrollierenden durch die Kontrolle selbst; und zweitens durch verschiedene polizeiliche Maßnahmen im Rahmen der Kontrolle, die als Verbote oder Vorschriften erlebt werden. Bedeutsam ist hierbei, dass es sich um erlebte Freiheitsbedrohungen oder Einschränkungen handelt und nicht um objektive Tatbestände. Relevant ist also, wie der Bürger Kontroll- und Eigensicherungsmaßnahmen beurteilt und nicht welche rechtlichen Zulässigkeiten vorliegen. Erlebt eine Person, die kontrolliert wird, diese Situation als bzgl. ihrer Handlungsfreiheit hinreichend einschränkend und meint sie, dass diese Einschränkung nicht gerechtfertigt ist, kann Reaktanz auftreten, die die Person motiviert, gegen die einschränkenden Bedingungen vorzugehen. Die Stärke der Reaktanz, kann durch die eventuell totale Handlungseinschränkung (z.B. exakte Verhaltensvorschriften, die keine Wahl lassen) und Versagen jeglicher Alternativen ausgeprägt sein. Ebenso trägt dazu bei, dass natürlich eine Generalisierung auf zukünftige Situationen und weitere Polizeikontakte wahrscheinlich ist. Entsprechend den sozialen Normen, Gesetzen und der Sanktionsmacht der Polizisten werden als Reaktanzeffekte von Kontrollsituationen zwar wahrscheinlich eher subjektive Reaktionen statt Verhaltenseffekte in Form von Angriffen und Widerstandshandlungen auftreten, jedoch sind beide Arten möglich und jeweils unerwünscht. Denn auch kognitive Reaktanzreaktionen und (zeitlich) verschobene Verhaltenseffekte dienen nicht einer kooperativen Beziehung zwischen Polizei und Bürger. Auch sind reaktante Reaktionen nicht im Sinne des Konzeptes „Bürgerfreundlichkeit“.

## **Methode**

Fragebogen

Der für diese Studie konstruierte Fragebogen enthielt allgemeine Fragen zur Person (Geschlecht, Alter, Wohnort-Charakter, Bildung), zu Erfahrungen mit der Polizei (bekannt mit Polizisten, Beurteilung des letzten Kontaktes, Kontrollerfahrung), zur subjektiven Gewaltsituation (Veränderung der Gewalt im Allgemeinen in Deutschland sowie gegen Polizisten), zur Beurteilung der Dienstpistole der Polizei und zur Akzeptanz unterschiedlicher Einsatzmittel. Neben diesen Items war ein Katalog von Eigensicherungs- und Kontrollmaßnahmen zu beurteilen sowie ein Fragebogen zur Reaktanz und zur Einstellung gegenüber der Polizei zu beantworten.

### Eigensicherungs- und Kontrollmaßnahmen

Die Teilnehmer wurden gebeten, sich vorzustellen, sie seien mit dem Auto unterwegs zu einem wichtigen Termin. Dabei würden sie angehalten und sollten kontrolliert werden. Verschiedene Eigensicherungs- und Kontrollmaßnahmen (z.B. „Ein Polizist hat bei der Kontrolle die Hand an seiner Waffe“), die dabei möglicherweise ihnen gegenüber zur Anwendung gebracht werden könnten, sollten sie auf einer 5-stufigen Skala einschätzen (von „ist völlig übertrieben“ bis „ist sehr angebracht“). Analog wurde eine Situation zur Kontrolle der Person als Fußgänger gegeben, wobei Maßnahmen weggelassen wurden, die spezifisch für eine KFZ-Kontrolle sind (z.B. Aufforderung aus dem Auto auszusteigen oder die Hände auf das Lenkrad zu legen).

### Einstellung von Bürgern zur Polizei

Die Messung der Einstellung von Bürgern zur Polizei im Allgemeinen erfolgte durch die Skala Polizei (FRPO) des Frankfurter Rechtsinventars von Deusinger (n.d.).<sup>1</sup> Ausgehend von einem klassischen Einstellungskonzept, das emotionale, kognitive und konative Komponenten umschließt, soll die Einstellung gegenüber der Polizei als rechtspflegende Instanz Verhalten und Entscheidung in entsprechenden Situationen unter Berücksichtigung situativer Faktoren beeinflussen (Deusinger, 1997). Entsprechend erfasst die Skala FRPO mittels 20 Items, die eine 6-fach skalierte Zustimmung (von „trifft sehr zu“ bis „trifft gar nicht zu“) zu allgemeinen Aussagen über die Polizei (z.B. „Polizisten setzen ihre Gesundheit und ihr Leben aufs Spiel, um unser Zusammenleben zu ordnen“) erfordern, die Einstellung multidimensional. Der Einstellungswert wird über einen Summenscore gebildet, der eine Polung der Items berücksichtigt. Es ergibt sich eine Score-Spanne von 20 bis 120. Hohe Testwerte (ab 80) lassen eine positive Einstellung zur Polizei erkennen. Niedrige Testwerte (bis 60) stehen für eine negative Einstellung zur Polizei. Erste Erhebungen an nicht repräsentativen Stichproben zeigen für die Skala Polizei u.a. signifikante Unterschiede zwischen männlichen Delinquenten und Männern der Normalpopulation sowie statistisch sehr signifikante Zusammenhänge mit Einstellungen zu Recht und rechtspflegenden Instanzen (Deusinger, 1995).

### Reaktanz

Die Ermittlung der Neigung auf Freiheitseinschränkung eine Reaktanz-Motivation zu entwickeln, übernahm der Reaktanzfragebogen von Metz (1983), bei dem davon ausgegangen wird, dass sich mittels direkter Befragung eine generalisierte Bereitschaft, psychologische Reaktanz zu mobilisieren, erfassen lässt. Das Instrument enthält 18 Statements (z.B. „Vorschriften und Verpflichtungen erwecken in mir starke Widerstände“), die Reaktanzsituationen und entsprechende Reaktionen darauf beschreiben. Der Befragte hat zu beurteilen, inwieweit diese Aussagen auf ihn persönlich zutreffen (6-fach gestufte Antwortmöglichkeiten von „trifft gar nicht zu“ bis „trifft vollständig zu“). Ein aus den Antworten gebildeter Summenwert (Score-Spanne 18-108) gibt Aufschluss über die Bereitschaft, psychologische Reaktanz zu mobilisieren. Metz (1983, S. 78ff.) gibt an, dass der Fragebogen gut zwischen verschiedenen

Merkmalsausprägungen zu differenzieren vermag, über eine recht befriedigende Reliabilität verfügt und Hinweise Anlass geben, Validität anzunehmen.

## Befragung

Die Studie war auf Grund des explorativen Charakters nicht auf Repräsentativität der Stichprobe ausgerichtet. Es wurden im Januar 2001 in Hessen Bürger an verschiedenen öffentlichen und nicht-öffentlichen Orten zufällig ausgewählt und zur freiwilligen Teilnahme aufgefordert. Deshalb ist von einer Selektivität der Stichprobe auszugehen. Eine Generalisierung der Ergebnisse ist deshalb nicht möglich. Ebenso ist bei einer Interpretation der Ergebnisse zu berücksichtigen, dass durch eine hohe Medienpräsenz der Thematik und durch soziale sowie politische Umstände nicht ausgeschlossen werden kann, dass die Antworten momentanen Tendenzen unterliegen und weniger latente, zeitlich relativ stabile Dispositionen reflektieren.

## Teilnehmer

Insgesamt konnten 431 Personen für eine Teilnahme gewonnen werden, die bei einigen jedoch nur unvollständig blieb. Der Anteil der Frauen betrug 46.9%. Die Mehrheit entstammte einer eher jüngeren Bevölkerungsschicht bis 30 Jahre (54%). Mit nur 9.6% war der Anteil der über 50-jährigen Personen deutlich unterrepräsentiert. Zu jeweils ca. 1/3 hielten die Teilnehmer ihre Wohngegend für ländlich, kleinstädtisch oder großstädtisch. Sehr wenige der Befragten waren ohne Schulabschluss (2,6%). Ca. je 1/3 hatten Abitur oder Realschulbildung. Die Übrigen hatten die Hauptschule abgeschlossen (21.3%) oder einen (Fach-) Hochschulabschluss (12.8%). Während bei über 2/3 die entfernte Verwandtschaft oder der Bekanntenkreis einen Polizisten einschloss, hatten 9.8% einen Polizisten als nahen oder nächsten Verwandten (Lebenspartner, Eltern, Geschwister). Die Beurteilung des letzten Polizeikontaktes fiel bei fast der Hälfte der Teilnehmer neutral aus, während 1/4 diesen als „eher angenehm“ bis „sehr angenehm“ bewertete. Mindestens einmal von der Polizei kontrolliert wurden 64.7% aller Teilnehmer als Autofahrer und 16.9% als Fußgänger.

## Ergebnisse

### Subjektive Gewaltentwicklung, Beurteilung der Dienstwaffe und Akzeptanz polizeilicher Einsatzmittel

Die überwiegende Mehrheit der Befragten meinte, dass Gewalt im Allgemeinen in Deutschland leicht oder sogar stark zugenommen habe (83.7%). Nur etwa jeder Zehnte war der Ansicht, dass diese unverändert sei. Die Zunahme von Gewalt gegen Polizeibeamte wurde statistisch signifikant minimal geringer eingeschätzt (Wilcoxon-Test  $Z=-2.233^*$ ).

Die Majorität der Befragten (63.8%) gab an, dass der Anblick einer Dienstpistole als Ausrüstungsgegenstand bei einem uniformierten Polizisten allgemein keine besonderen Empfindungen bei ihnen auslöse. ¼ aller Befragten fühlten sich durch diese Bewaffnung sogar geschützt. Nur etwa jeder Zehnte gab an, negative Wirkungen durch den Anblick der Waffe zu verspüren. Die Antworten bei der hier vorliegenden Studie unterscheiden sich statistisch sehr signifikant zu denen, die Murck (1989) bei derselben Fragestellung erhalten hatte ( $\chi^2$  nach Pearson=25.136<sup>\*\*</sup>, df=4, n=1442): Die Anzahl derer, die negativ beeindruckt sind, zeigt sich bei der vorliegenden Befragung halbiert, während deutlich mehr Personen angeben, unberührt vom Anblick der Dienstwaffe zu sein.

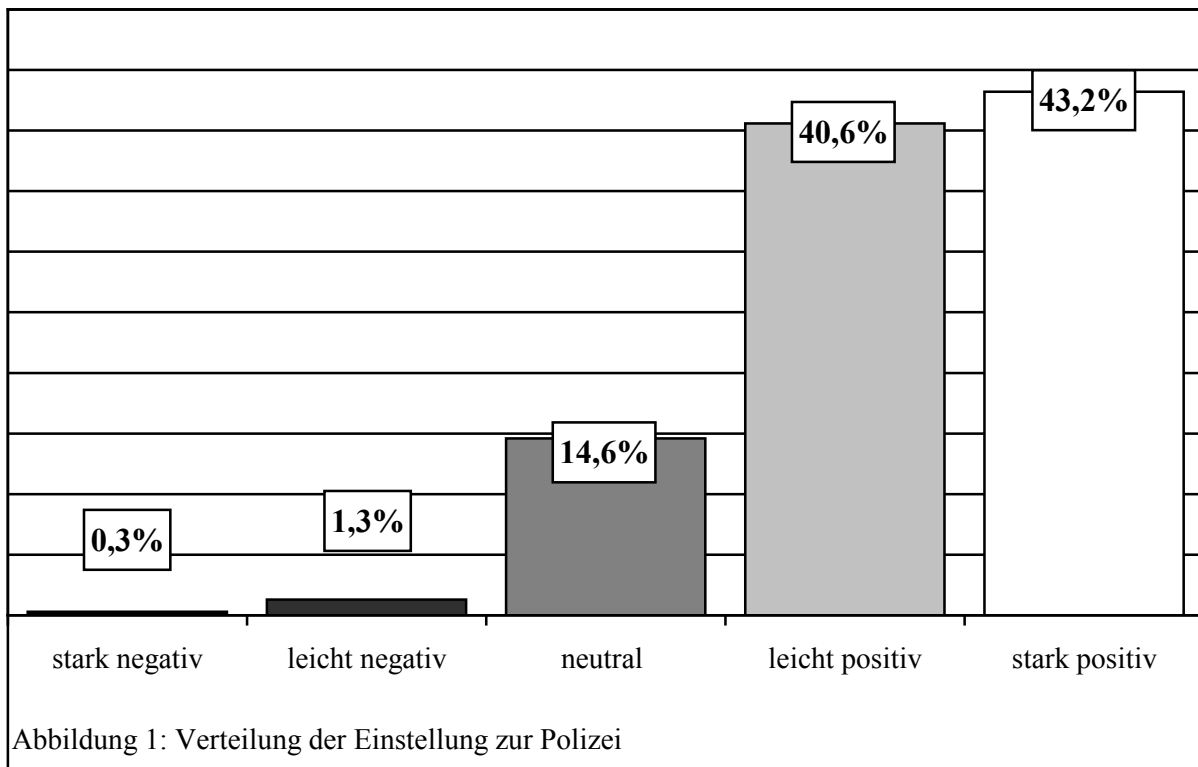


Die Akzeptanz unterschiedlicher polizeilicher Einsatzmittel bei gewaltsamen Auseinandersetzungen differiert meist statistisch signifikant (Wilcoxon-Tests). Mit Abstand am besten akzeptiert ist der Einsatz von Handschellen. Hier halten 96.6% der Befragten den Einsatz für in entsprechenden Situationen für „eher akzeptabel“ oder „völlig akzeptabel“. Weniger einheitlich werden Einsatzmittel beurteilt, die wahrscheinlich vor allem mit Demonstrationsveranstaltungen assoziiert sind: Schlagstock, Tränengas und Wasserwerfer. Sie werden von  $\frac{3}{4}$  der Teilnehmer als „eher akzeptabel“ oder „völlig akzeptabel“ angesehen. Von ca. 65% wird die Schusswaffe in gewaltsamen Auseinandersetzung „eher“ bis „völlig“ akzeptiert. Ein völlig anderes Bild bietet sich hingegen für den Einsatz körperlicher Gewalt: Die Mehrheit hält Schlagen und Treten für „eher nicht akzeptabel“ bis „gar nicht akzeptabel“. Damit liegt dieses Einsatzmittel noch hinter der Schusswaffe. Dabei kann jedoch nicht ausgeschlossen werden, dass für die Beurteilung der Einsatzmittel von unterschiedlichen Einsatzsituation ausgegangen wurde. Vielmehr kann es sein, dass jeweils entsprechende Situationen, die sich möglicherweise im Gewaltniveau unterscheiden, als Bewertungsgrundlage verwendet wurden und keine gemeinsame Referenz Anwendung fand.

Hinsichtlich eines Vergleichs der vorliegenden Befragung mit der von 1987 (Murck, 1989) ergeben sich deutliche Unterschiede ( $\chi^2$  nach Pearson jeweils  $> 42.077^{**}$ ,  $df=3$ ,  $n>1451$ ): Die Einsatzmittel Schlagstock, Tränengas und Wasserwerfer konnten eine deutliche Verbesserung der Akzeptanz verbuchen (die Übrigen wurden 1987 nicht erfragt).

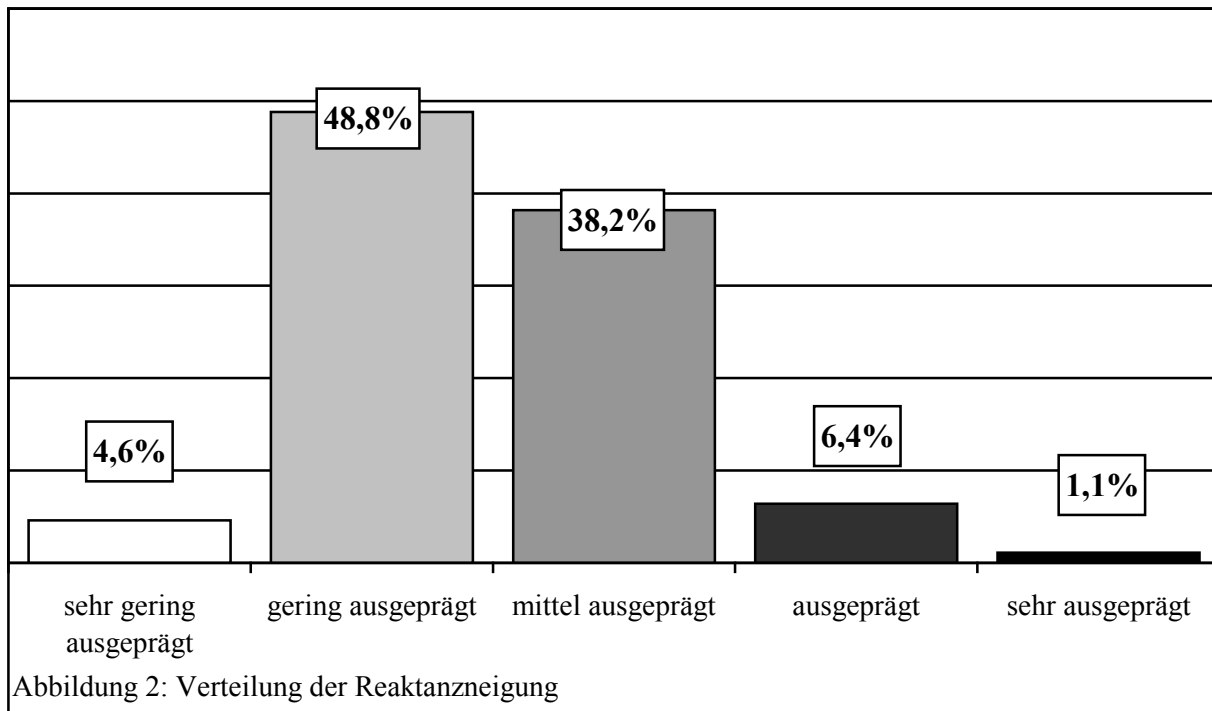
### Einstellung

Die Befragung zur Einstellung gegenüber der Polizei ergab, dass die befragten Personen mehrheitlich (83.8%) positiv zur Polizei eingestellt sind (siehe Abbildung 1). Im Mittel erreichen sie einen Einstellungswert von 94.32 ( $sd=14.15$ ,  $n=370$ ,  $min.=24$ ,  $max.=120$ ). Während 14.6% eher im neutralen Einstellungsbereich antworteten, konnte bei nur 1.6% eine negative Einstellung zur Polizei gefunden werden. Die Einstellung zur Polizei korreliert mit der subjektiven Entwicklung der Gewalt gegen Polizeibeamte (Spearman-Rho= $.330^{**}$ ) und der Beurteilung der Schusswaffe als polizeiliches Einsatzmittel (Spearman-Rho= $-.321^{**}$ ) mindestens gering.



## Reaktanz

Im Mittel erreichen die Befragten einen Reaktanzwert von 54.49 ( $n=390$ ,  $sd=11,79$ ,  $min.=23$ ,  $max.=103$ ), der als mittlere Ausprägung der latenten Reaktanz mit Tendenz zu geringer Reaktanzneigung angesehen werden kann (siehe Abbildung 2). Dabei weisen Befragte durchaus sehr niedrige Reaktanzwerte auf, während andere eine sehr hohe Bereitschaft zur Reaktanzmobilisierung zeigen. Letztendlich finden sich in den Kategorien mit extremer Ausprägung jedoch die wenigsten Personen. Fast 90% neigen nur gering oder mittelmäßig zur Reaktanz. Die Reaktanzneigung korreliert mit der Einstellung zur Polizei (Spearman's-Rho=-.221\*\*) und der Beurteilung der Schusswaffe als polizeiliches Einsatzmittel (Spearman's-Rho=.173\*\*) nur gering. Die subjektive Entwicklung der Gewalt gegen Polizeibeamte steht mit der Reaktanz in keinem statistischen Zusammenhang (Spearman's-Rho=-.099).



### Eigensicherungs- und Kontrollmaßnahmen

Die Bewertungen unterschiedlicher Maßnahmen fallen sehr unterschiedlich aus (siehe Tabelle 2). Einerseits wird bei allen Maßnahmen die gesamte Breite der Antwortmöglichkeiten genutzt, was für eine deutliche interindividuelle Variation der Urteile spricht. Andererseits fallen auch die mittleren Bewertungen für verschiedene Maßnahmen sehr unterschiedlich aus und reichen von „eher angebracht“ mit Tendenz zu „sehr angebracht“ bis hin zu „leicht übertrieben“ mit Tendenz zu „völlig übertrieben“. Es erstrecken sich also auch die mittleren Bewertungen fast über das gesamte Antwortspektrum. Dabei schneiden die Vorgehensweisen durchgängig (bis auf die Maßnahme „ein Polizist will Sie durchsuchen“) bei Fußgängerkontrollen negativer ab als bei Fahrzeugüberprüfungen.

Es finden sich zwischen der Bewertung der Kontroll- und Eigensicherungsverfahren und der Einstellung zur Polizei mehrheitlich signifikante Korrelationen, die jedoch nur auf geringe Zusammenhänge hindeuten. Ebensolches lässt sich für die subjektive Zunahme von Gewalt gegen Polizeibeamte feststellen. Für die Reaktanzneigung konnte nur zweimal ein statistisch bedeutsamer Zusammenhang mit der Maßnahmenbewertung gefunden werden. Deren Größenordnung ist als geringfügig einzustufen. Für die waffenbezogenen Maßnahmen (Hand an der Waffe, Waffe gezogen und Maschinenpistole) konnten Korrelationen mit der allgemeinen Akzeptanz der Schusswaffe als polizeiliches Einsatzmittel gefunden werden, die wie die Übrigen auch nur eine geringe Größe besitzen.

Tabelle 2: Ergebnisse der mittleren Bewertungen von Eigensicherungs- und Kontrollmaßnahmen

Maßnahme	Mittelwert <sup>a)</sup>	SD	Korrelationen <sup>b)</sup>				Wilcoxon Z
			Einstellung	Reaktanz	Gewalt	Schusswaffe <sup>c)</sup>	
Beide Polizisten haben Schutzwesten an	4.27	.98	.379**		.264**	-4.394**	
	4.12	1.10	.393**		.261**		
Einer der Polizisten will Ihren Personalausweis sehen	4.03	.93	.385**		.212**	-4.362**	
	3.89	.97	.362**		.226**		
Ein Polizist fordert Sie auf, im Auto sitzen zu bleiben	3.54	.89	.193**		.105*		

Ein Polizist fordert Sie auf, aus dem Auto auszusteigen	3.38	.97	.327**	.223**		
Ein Polizist sagt zu Ihnen: "Zeigen Sie mir Ihre Hände!"	2.94	1.19	.271**	.193**		-4.704**
	2.74	1.15	.203**	.207**		
Ein Polizist sagt zu Ihnen: "Legen Sie die Hände auf das Lenkrad!"	2.79	1.28	.198**	.203**		
Ein Polizist hat bei der Kontrolle die Hand an seiner Waffe	2.74	1.32	.271**	.253**	-.300**	-6.276**
	2.43	1.20	.253**	.266**	-.313**	
Die Kontrolle wird mit einer Videokamera gefilmt	2.67	1.25	.199**	.136**		
Ein Polizist gibt Ihnen Anweisungen über die Lautsprecheranlage des Streifenwagens	2.55	1.21	.112**	.158**		
Ein Polizist will Ihr Auto durchsuchen.	2.55	1.06	.205*			
Einer der Polizisten will Sie durchsuchen.	2.34	1.05	.163*			-1.719
	2.28	1.03		.119*		
Einer der Polizisten hat bei der Kontrolle seine Waffe gezogen	2.21	1.20	.151**	.165**	-.257**	-7.716**
	1.86	1.08		-.115*	.149**	
Einer der Polizisten hält bei der Kontrolle eine Maschinenpistole	2.11	1.21	.224**	.188**	-.304**	-4.784**
	1.92	1.14	.135**	.173**	-.240**	
Ein Polizist sagt zu Ihnen: „Machen Sie keine falsche Bewegung!“	2.09	1.14	.140**	.165**		-2.767**
	1.98	1.06		.157**		
Ein Polizist sagt zu Ihnen: "Hände hoch!"	1.82	1.04		.124*		-2.2947**
	1.72	.93		.105*		
Ein Polizist will, dass Sie sich mit gespreizten Armen und Beinen an Ihr Auto/an ein Haus stellen	1.82	1.00		.157**		-3.986**
	1.71	.96		-.142**	.112*	
<p>In der oberen Zeile befinden sich jeweils die Parameter für eine KFZ-Kontrolle, in der unteren Zeile stehen die Ergebnisse für eine Fußgängerkontrolle.</p> <p>a) Antwortkategorien: 1= „ist völlig übertrieben“; 2= „ist leicht übertrieben“; 3= „ist mir egal“; 4= „ist eher angebracht“; 5= „ist sehr angebracht“.</p> <p>b) Es sind ausschließlich statistisch signifikante Korrelationen (Spearman-Rho) aufgeführt.</p> <p>c) Der Zusammenhang wurde nur für waffenbezogene Maßnahmen berechnet.</p> <p>*) <math>p &lt; .05</math>; **) <math>p &lt; .01</math>.</p>						

## Faktorenanalyse der Bewertungen von Eigensicherungs- und Kontrollmaßnahmen

Es ist offensichtlich, dass die Eigensicherungs- und Kontrollmaßnahmen mehr oder minder ähnlich zueinander sind bzw. gemeinsame Anteile besitzen. Aus diesem Grund wurden sie einer explorativen Faktorenanalyse unterzogen, die quasi Gemeinsamkeiten verschiedener Maßnahmen zum Vorschein bringen sollte.<sup>2</sup> Eine Hauptkomponentenanalyse mit dem Kriterium „Eigenwerte $>1$ “ extrahiert fünf Faktoren, die gemeinsam eine Varianz von 65.714% erklären. Die Anfangslösung offenbart einen sehr starken ersten Faktor und vier weitere Faktoren, die mehr oder minder augenscheinlich ähnliche Maßnahmen zusammenfassen. Die Interpretation der durch Varimax-Rotation erstellten Komponentenmatrix (siehe Tabelle 3) legt folgende fünf Faktoren nahe:

- Waffenbezogene Eigensicherungsmaßnahmen
- distanzierte Maßnahmen und Verhaltensanweisungen
- „alltägliche“ Maßnahmen und passive Sicherheit
- Durchsuchung
- Aufforderungen bzgl. der Hände

Tabelle 3: Rotierte Komponentenmatrix

Maßnahme	Komponente				
	1	2	3	4	5
Einer der Polizisten hält bei der Kontrolle eine Maschinenpistole*	.805				
Einer der Polizisten hält bei der Kontrolle eine Maschinenpistole	.789				
Einer der Polizisten hat bei der Kontrolle seine Waffe gezogen*	.772				
Ein Polizist hat bei der Kontrolle die Hand an seiner Waffe*	.713				
Einer der Polizisten hat bei der Kontrolle seine Waffe gezogen	.701				
Ein Polizist hat bei der Kontrolle die Hand an seiner Waffe	.695				
Ein Polizist sagt zu Ihnen: "Hände hoch!"	.775				
Ein Polizist sagt zu Ihnen: "Hände hoch!"*	.709				
Ein Polizist sagt zu Ihnen: "Machen Sie keine falsche Bewegung!"	.704				
Ein Polizist gibt Ihnen Anweisungen über die Lautsprecheranlage des Streifenwagens*	.633				
Ein Polizist sagt zu Ihnen: "Machen Sie keine falsche Bewegung!"*	.612				
Ein Polizist will, dass Sie sich mit gespreizten Armen und Beinen an eine Hauswand stellen	.594			.512	
Die Kontrolle wird mit einer Videokamera gefilmt*	.570				
Einer der Polizisten will Ihren Personalausweis sehen*			.843		
Einer der Polizisten will Ihren Personalausweis sehen			.790		
Ein Polizist fordert Sie auf, im Auto sitzen zu bleiben*			.678		
Ein Polizist fordert Sie auf, aus dem Auto auszusteigen*			.653		
Beide Polizisten haben Schutzwesten an*			.574		
Beide Polizisten haben Schutzwesten an			.557		
Einer der Polizisten will Sie durchsuchen*				.798	
Einer der Polizisten will Sie durchsuchen				.763	
Ein Polizist will Ihr Auto durchsuchen*				.641	
Ein Polizist will, dass Sie sich mit gespreizten Armen und Beinen an Ihr Auto stellen*				.551	
Ein Polizist sagt zu Ihnen: "Zeigen Sie mir Ihre Hände"*					.694
Ein Polizist sagt zu Ihnen: "Zeigen Sie mir Ihre Hände"					.609
Ein Polizist sagt zu Ihnen: "Legen Sie die Hände auf das Lenkrad"*					.537
* = KFZ-Kontrolle.					
Ladungen < .5 werden nicht dargestellt.					

## Diskussion

Wie bei allen bewerteten Eigensicherungs- und Kontrollmaßnahmen ersichtlich ist, gehen die Meinungen bei den befragten Bürgern deutlich auseinander, welches polizeiliche Verhalten bei Kontrollen angebracht ist und welches nicht: Die Beurteilungen streuen jeweils über alle Antwortkategorien. Es ist deutlich geworden, dass verschiedene Maßnahmen allgemein eher akzeptiert werden als andere. Dabei finden Maßnahmen, wie „Schutzwesten zu tragen“ überwiegend Zustimmung, während z.B. „das Ziehen der Dienstwaffe“, verbale Verhaltenskommandos wie "Hände hoch!" oder "Machen Sie keine falsche Bewegung!" und

durchsucht zu werden (inkl. der dabei einzunehmenden Position) als übertrieben abgelehnt werden. Die hier zur Beurteilung gebrachten Eigensicherungs- und Kontrollmaßnahmen stoßen bei den Teilnehmern der Befragung überwiegend eher auf Ablehnung. Zu erwähnen ist hierbei unbedingt, dass die befragten Bürger überwiegend positiv oder sogar sehr positiv zur Polizei eingestellt sind und wenig bis mittelmäßig zur Reaktanz neigten. Es handelte sich also wahrscheinlich um eine Stichprobe, die eher kooperativ und polizeifreundlich in Bürger-Polizei-Begegnungen ist. Entsprechend fällt auf, dass die Einstellung zur Polizei nur gering, die Reaktanz sogar quasi gar nicht mit den Bewertungen der Maßnahmen zusammenhängt. Ebenso scheint die Einschätzung der Veränderung der Gewalt gegen Polizeibeamte nur minimal mit der Beurteilung der Maßnahmen einherzugehen. Dies kann darauf hindeuten, dass eine Ablehnung von Maßnahmen nicht unbedingt damit einhergeht, dass man Polizei allgemein ablehnt bzw. dieser gegenüber negativ eingestellt ist. Umgekehrt tolerieren anscheinend Personen, die sehr positiv zur Polizei eingestellt sind, ebensowenig automatisch jegliches polizeiliches Verhalten bei Kontrollen.

Dass fast nie ein Zusammenhang zwischen Reaktanzneigung und Beurteilung der Maßnahmen zu finden war, unterstützt die zuvor schon geäußerten Vermutungen, dass die Befragten nicht pauschal auf Grund von Dispositionen antworten, sondern differenziert die Kontrollverfahren bewerten. Es scheint also, dass eine Maßnahme nicht deshalb abgelehnt werde, weil man sich generell nichts sagen lassen möchte.

Die Akzeptanz von Eigensicherungs- und Kontrollmaßnahmen zeigte sich nur schwach zusammenhängend mit der Einschätzung der Veränderung von Gewalt gegen Polizeibeamte. Zwar scheint grundsätzlich die Einsicht in die Gefahrensituation der Polizei mit der Akzeptanz von Sicherheit spendenden Verhaltensweisen einherzugehen, doch weisen die nur gering ausgeprägten Korrelationen darauf hin, dass Mitgefühl nicht hinreichend ist für Akzeptanz. Hier wird eventuell ein Konflikt zwischen eigener Perspektive und Verständnis für die polizeiliche Lage zu eigenen Gunsten gelöst.

Die Akzeptanz der Schusswaffe als Einsatzmittel hängt zwar mit der Beurteilung von Kontrollmaßnahmen zusammen, erreicht aber keinen hohen Wert. Dies deutet darauf hin, dass die Personen, die allgemein den Einsatz von Schusswaffen in gewalttätigen Situationen durch die Polizei akzeptieren, nicht unbedingt diese Schusswaffe auch bei einer Kontrolle ihrer Person eingesetzt (im Sinne des Konzeptes von Lorei, 1999) sehen wollen. D. h. die Ablehnung, mit gezogener Waffe kontrolliert zu werden, steht nur in geringem Zusammenhang damit, dass man generell eine Dienstwaffe als polizeiliches Einsatzmittel ablehnt.

Es kann vermutet werden, dass die befragten Personen genau zwischen der Polizei allgemein und spezifischen Einsatzmaßnahmen, bzw. Gruppen von polizeilichen Verhaltensweisen (siehe Faktorenanalyse) differenzieren. Dies entspricht Befunden von Tyler & Folger (1980): Sie fanden in ihrer Studie Anzeichen dafür, dass Beurteilungen des Polizei-Bürger-Kontaktes in Hinsicht auf Fairness und Zufriedenheit nicht einfach nur Ergebnisse einer Einstellung oder Voreingenommenheit darstellen, sondern vor allem Resultate der persönlichen Erfahrung einer spezifischen Situation sind. Vergleichbar dazu scheinen die hier gefundenen nur geringen Korrelationen Ausdruck dessen zu sein, dass Eigensicherungs- und Kontrollmaßnahmen nicht auf Grund einer negativen Einstellung zur Polizei abgelehnt werden, sondern die befragten Personen sich diesen Maßnahmen nicht ausgesetzt sehen wollen. Dabei ist zu beachten, dass die statistischen Zusammenhänge keine Aussage über Wirkungszusammenhänge zulassen. Weiterhin ist zu berücksichtigen, dass auf Grund der geringen Streuung der Einstellungswerte sowie der Bewertungen Korrelationen tendenziell geringer ausfallen können. Die Größe der Korrelationen zwischen Bewertungen und der

Einstellung kann somit auch methodisches Artefakt sein, da durch geringe Varianz keine ausgeprägte gemeinsame Variation möglich war.

Auffällig ist, dass alle Maßnahmen bei einer Fußgängerkontrolle weniger Akzeptanz finden als bei einer Fahrzeugkontrolle. Das unterstützt die Annahme, dass die polizeilichen Maßnahmen nicht pauschal und auf Grund einer eher allgemeinen Haltung beurteilt und akzeptiert werden, sondern unter Berücksichtigung spezifischer Situationsmerkmale Bewertung finden.

## **Schlussfolgerungen**

Allgemein kann nicht davon ausgegangen werden, dass polizeiliches Verhalten in Kontrollsituationen einfach hingenommen wird oder sogar begrüßt wird. Vielmehr spricht einiges dafür, dass einige polizeiliche Kontrollverfahren schon bei „polizeifreundlichen“ Personen auf Ablehnung stoßen. Es ist deshalb unumgänglich, die Akzeptanz zu fördern, um sowohl die Sicherheit der Polizisten zu gewährleisten als auch eine Bürgerzufriedenheit zu erreichen.

Mittlerweile sind verschiedene Informationskampagnen zur Eigensicherung gestartet worden. Sie setzen vor allem darauf, den Bürger über die (KFZ-) Kontrollmaßnahmen zu informieren, indem sie u.a. Verhaltensregeln aufstellen und auf die Gefahren für die Polizei hinweisen (siehe z.B. Innenministerium des Landes Nordrhein-Westfalen, 2001). Es bleibt abzuwarten, welche Konsequenzen diese Maßnahmen haben. Aus den Ergebnissen dieser Studie scheint fraglich, ob Hinweise auf Gefahren für Polizeibeamte ausreichen. Wie oben dargestellt, sehen die Bürger der Stichprobe Gewalt gegen Polizeibeamte zwar, doch hängt dies nur wenig mit den Bewertungen der Maßnahmen zusammen. Es ist deshalb zu vermuten, dass weitere Hinweise auf Gewalt gegen Polizeibeamte nicht zu mehr Akzeptanz führen werden. Eben solches kann für das „Image“ der Polizei vermutet werden, da die Einstellung zur Polizei nur gering mit der Akzeptanz der polizeilichen Maßnahmen zusammenhängt.

Sollten die Informationen lediglich Anweisungsscharakter haben, scheint kritisch, dass Reaktanz verstärkt wird, wenn Personen bzgl. einer Freiheitseinschränkung informiert oder vorgewarnt werden (Dickenberger et al., 1993, S. 247). Dies bedeutet, dass ein Hinweis auf Verhaltensvorschriften (oder -erfordernisse) nicht ausreichend ist, sondern womöglich Widerstand hervorrufen könnte.

## **Literatur**

- Brehm, J. W. (1966). *A theory of psychological reactance*. New York: Academic Press.
- Brehm, J. W. (1972). *Responses to loss of freedom. A theory of psychological reactance*. Morristown: General Learning Press.
- Deusinger, I. M. (1995). Polizeiforschung. Betrachtungen zum Forschungsanliegen und erste Ergebnisse einer systematischen Untersuchung zu Einstellungen der Bürger zur Polizei. *Kriminalistik*, 11, 705-710.
- Deusinger, I. M. (1997). Einstellungen zum Gesetz und Zeugenaussagen. In L. Greuel, T. Fabian & M. Stadler (Hrsg.), *Psychologie der Zeugenaussage. Ergebnisse der rechtspsychologischen Forschung* (S. 271-285). Weinheim: Psychologie Verlags Union.
- Deusinger, I. M. (n.d.). *Frankfurter Rechtsinventar*. Göttingen: Hogrefe (in Vorbereitung).

- Dickenberger, D., Gniech, G. & Grabitz, H.-J. (1993). Die Theorie der psychologischen Reaktanz. In D. Frey & M. Irle (Hrsg.) *Theorien der Sozialpsychologie. Band 1: Kognitive Theorie* (S. 243-273). Bern: Huber.
- EMNID (1988). Die persönliche Einstellung zu aktuellen Fragen der Inneren Sicherheit. *Die Polizei*, 7, 193-204.
- Hasselmann, J. (2000). Sicherheit für Polizisten. Ordnungshüter soll schneller zur Pistole greifen. *Tagesspiegel*, 27.6.2000.
- Hermanutz, M. (1995). Prügelknaben der Nation oder Freund und Helfer: Die Zufriedenheit von Bürgern mit den Umgangsformen der Polizei nach einem persönlichen Polizeikontakt – eine empirische Untersuchung. *Die Polizei*, 10, 281-287 .
- Innenministerium des Landes Nordrhein-Westfalen (2001). *Sicherheit braucht Kontrollen*. Faltblatt zur Aktion "Sicherheit braucht sichere Kontrollen".
- Lorei, C. (1999). *Der Schußwaffeneinsatz bei der Polizei. Eine empirisch-psychologische Analyse*. Berlin: Wissenschaftlicher Verlag Berlin.
- Merz, J. (1983). Fragebogen zur Messung der psychologischen Reaktanz. *Diagnostica*, 29 (1), 75-82.
- Mummendey, H. D. (1988). Die Beziehung zwischen Verhalten und Einstellung. In H. D. Mummendey (Hrsg.), *Verhalten und Einstellung* (S. 1 – 26). Berlin: Springer.
- Mummendey, H. D. (1995). *Die Fragebogen-Methode: Grundlagen und Anwendung in Persönlichkeits-, Einstellungs- und Selbstkonzeptforschung* (2., korrigierte Aufl.). Göttingen: Hogrefe.
- Naether, E.-A. (1980). Informations- und Wissensstand der Bevölkerung Nordrhein-Westfalens in bezug auf die Polizei und den Polizeiberuf. *Schriftenreihe der Polizei-Führungsakademie*, 1/80, 66-89.
- Rosenberg, M. J. & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In C. I. Hovland & M. J. Rosenberg (eds.), *Attitude Organization and Change*. New Haven: Yale University Press [zitiert nach Stahlberg & Frey, 1992, S. 145f.].
- Schmalzl, H. P. (1996). Bürgerfreundlich. In M. Herrmanutz, C. Ludwig & H. P. Schmalzl (Hrsg.), *Moderne Polizeipsychologie in Schlüsselbegriffen* (S. 40-47). Stuttgart: Boorberg.
- Six, B. & Eckes, T. (1994). Fakten und Fiktionen in der Einstellungs-Verhaltens-Forschung: Eine Metaanalyse. *Zeitschrift für Sozialpsychologie*, 25 (4), 253 – 271.
- Six, B. & Eckes, T. (1996). Metaanalysen in der Einstellungs-Verhaltens-Forschung. *Zeitschrift für Sozialpsychologie*, 27 (1), 7 – 17.
- Six, B. (1998). Moderatoren der Einstellungs-Verhaltens-Relation. In E. H. Witte (Hrsg.), *Sozialpsychologie der Kognition: Soziale Repräsentationen, subjektive Theorien, soziale Einstellungen: Beiträge des 13. Hamburger Symposions zur Methodologie der Sozialpsychologie* (S. 206 – 228). Lengerich: Pabst.
- Stahlberg, D. & Frey, D. (1992). Einstellung I: Struktur, Messung und Funktionen. In W. Stroebe, M. Hewstone, J.-P. Codol & G. M. Stephenson (Hrsg.), *Sozialpsychologie: Eine Einführung* (S. 144 – 170). Berlin: Springer.
- Tagesspiegel (3.3.2001). *Pro & Contra*. <http://www.tagesspiegel.de>.
- Thomas, A. (1991). *Grundriß der Sozialpsychologie: Band 1: Grundlegende Begriffe und Prozesse*. Göttingen: Verlag für Psychologie.
- Tyler, T. R. & Folger, R. (1980). Distributional and Procedural Aspects of Satisfaction With Citizen-Police Encounters. *Basic and Applied Social Psychology*, 1 (4), 281-292.

<sup>1</sup> Ich danke Frau Prof. Dr. Deusinger für die freundliche Genehmigung, die Skala Polizei des Frankfurter Rechtsinventars in dieser Studie verwenden zu dürfen.

<sup>2</sup> Die strengen Voraussetzungen einer Faktorenanalyse wurden nicht überprüft. Im Rahmen einer explorativen Betrachtung mit vorsichtiger Interpretation erscheint jedoch die Anwendung dieses Verfahrens vertretbar. Die Faktorenanalyse wurde mittels SPSS 8.0 für Windows berechnet.



## Explorationstechnik – Einhaltung von Vernehmungsprinzipien zur Gewinnung valider Daten

*Charlotte Mohrbach, PsychFor\* Gesellschaft für Forensische Aussagepsychologie, Bonn*

Mit der Entscheidung vom 30.07.1999 hat der BGH die aussagepsychologische Begutachtung als wissenschaftliche Methode anerkannt und dabei klare Richtlinien für methodische und formale Standards vorgegeben.

Die aussagepsychologische Begutachtung ist ein mehrdimensionales System, dessen zentrales Element die kriterienorientierte Aussageanalyse ist. Diese kann aber nur dann die Methode der Wahl sein, wenn sichergestellt ist, daß die Zeugenangaben ein ausreichendes Maß an Eigenständigkeit aufweisen und nicht etwa durch Einwirkung Dritter wesentlich beeinflußt oder gar erst entstanden sind.

Ein wichtiger Schritt innerhalb des aussagepsychologischen Begutachtungsprozesses ist deshalb die Analyse der Entstehung und weiterer Entwicklungsbedingungen der Aussage – bis hin zum Zeitpunkt der Begutachtung, wobei die Ergebnisse der Suggestionforschung entsprechend zu berücksichtigen sind. Dabei gerät aus dem Blickfeld, daß auch die aussagepsychologische Exploration zur Sache einen wesentlichen Schritt innerhalb der Entwicklungsgeschichte der Aussage darstellt. Begreift man die Aussage in ihren verschiedenen Stadien als "Lernprozeß", bei dem der Befragte "lernt", worauf es ankommt, und bei dem auch die Voreinstellung des Befragers Einfluß hat, dann ist zu fordern, daß auch GutachterInnen die Regeln, die bei der Gewinnung verwertbarer Aussagen gelten, nicht nur kennen, sondern sie auch anwenden. Die Exploration zur Sache ist eine "Beweisaufnahme ohne Gericht" (Nack) und Gutachter haben deshalb die hohe Verpflichtung, mit Sachkunde zu explorieren.

In diesem Beitrag geht es darum, wesentliche Prinzipien ins Zentrum der Aufmerksamkeit zu rücken, deren Einhaltung bei der Gewinnung einer verwertbaren Aussage unabdingbar sind.

Grundlagen der Erörterungen sind Erfahrungen der Autorin aus aussagepsychologischer Gutachtertätigkeit, aus speziellen Vernehmungstrainings bei Kriminalbeamten sowie aus der Kenntnis von Fremdgutachten, die nach der zitierten BGH-Entscheidung erstellt worden sind. Letztere beinhalten Fehlerquellen im Bereich der Befragungstechnik, die unter dem Aspekt der Invalidierung von Aussagen zu diskutieren sind.

Die wissenschaftlich indizierte Vorgehensweise beim Begutachtungsprozeß erfordert, daß schon bei der Vorbereitung der Exploration auf der Grundlage des Aktenmaterials Hypothesen gebildet werden, die dann im Rahmen der Begutachtung zu prüfen sind. Erfahrungsgemäß ergeben sich aber im Verlauf der Exploration häufig Veränderungen oder auch völlig neue Hypothesen – wir haben es hier demnach mit einer dynamischen Entwicklung zu tun. Dies setzt – selbstverständlich - voraus, daß GutachterInnen flexibel und für alle neuen Aspekte offen bleiben.

Die wichtigsten Fehlerquellen:

1. Oftmals fehlt bereits der erste diagnostische Schritt der aussagepsychologischen Exploration (Steller & Volbert, 1999, Greuel et al., 1998, Arntzen, 1989): Es wird

kein freier Bericht angeregt. Statt dessen bewegt sich die Exploration von Anfang an auf einem Frage – Antwort – Abtausch.

2. Die Befragungen sind (und bleiben) vielfach auf einseitig vorgefaßte Sichtweisen (im Sinne der vorab gebildeten Hypothesen) eingeeengt. Es wird versäumt, die Zeugenangaben in psychologisch relevante Hypothesen zu übersetzen und entsprechend zu prüfen.

3. Im Explorationsverlauf werden erwartungs- (hypothesen-) konforme Antworten verstärkt, wobei dem Probanden teilweise erkennbar gemacht wird, welche Aspekte und Inhalte seiner Bekundungen besonders hoch eingeschätzt werden - etwa durch selektive Verstärkung einer Antwort oder Kommentierungen hinsichtlich ihrer Relevanz. Es wird vielfach konsequent und vorhaltend in eine Richtung gefragt, bis die offensichtlich gewünschte Antwort erfolgt.

4. Nicht selten sind Befragungen auch erkennbar auf eine Produktion derjenigen Aussagebesonderheiten ausgerichtet, die bei der späteren Aussageanalyse im Sinne von Glaubhaftigkeitsmerkmalen zur Stützung oder Falsifizierung der zu prüfenden Hypothesen beitragen können. „Realkennzeichen“ werden somit durch gezielte, oft vorhaltende Fragen evoziert, selektiv verstärkt und im Sinne einer zirkulären Beweisführung im Befund des Gutachtens als Beleg für die vermutete Erlebnisbegründung der Aussage herangezogen. In solchen Fällen kann man demnach von einer ergebnisorientierten Steuerung des Explorationsverlaufs sprechen.

Somit erfolgt bei Begutachtungen in vielen Fällen genau das, was die Forschung zur Suggestionproblematik als Fehlerquellen und Risikofaktoren herausgearbeitet hat. Keineswegs nur bei Aussagen von jungen Kindern, sondern auch bei Erwachsenen sind Suggestionseffekte nachgewiesen worden (Loftus 1974, 1975).

Der so dargestellte Befragungsstil ist nicht mit einer differentiellen diagnostischen Strategie zu verwechseln, bei der einzelne Vorhalte gezielt eingesetzt werden. Es handelt sich um ein viel komplexeres, prozeßhaftes Geschehen, bei dem die Beeinflussung auf eine viel subtilere und zugleich intensivere Art stattfindet als bei einfachen Vorhaltfragen.

Es ist nicht Aufgabe von GutachterInnen, Realkennzeichen „abzufragen“ oder neue Tatbestände zu ermitteln, sondern mittels trichterförmiger Befragung zu einer möglichst klaren, detaillierten und präzisen Erläuterung der Inhalte zu kommen, die vom Befragten eigenständig genannt worden sind, d.h. „in die Tiefe explorieren“. Es stellt sich deshalb die Frage: wie sind solche Fehlerquellen zu vermeiden?

1. Nicht nur durch eine fundierte Kenntnis der Vernehmungsprinzipien (die selbstverständlich sein sollte),

2. sondern auch durch intensives Training der Explorationstechnik

3. sowie durch ständige Selbstkontrolle des eigenen Befragungsverhaltens.

Diese Aspekte erscheinen uns als die wesentlichen Voraussetzungen für eine adäquate und valide Datenerhebung. Die Aufzeichnung auf Tonträger stellt sich dabei als optimale Grundlage zur Umsetzung dieser Voraussetzungen dar.

Zu 1. Eine gute Befragungstechnik ist die Kunst, eine valide Aussage zu gewinnen, indem man den Befragten dazu bringt, eine möglichst eigenständige, hinreichend klare und konkrete Sachverhaltsschilderung abzugeben. Dies wird in den seltensten Fällen schon im Rahmen des freien Berichts erreicht werden – aber dieser bietet die Grundlage, an der sich die anschließende Befragung zu orientieren hat.

Dabei muß sich der Befrager auf eine große Variationsbreite entwicklungspsychologischer und kognitiver Bedingungen einstellen können – dies sind im Grunde selbstverständliche, jedoch erfahrungsgemäß vielfach nicht erfüllte Voraussetzungen. Das bedeutet, der Befrager muß Sprache, Denk- und Wahrnehmungsstil sowie spezielle Persönlichkeitseigenarten des Befragten nicht nur erfassen, sondern in seinen eigenen Frageformulierungen berücksichtigen. Dies wiederum erfordert Flexibilität und Anpassungsfähigkeit seitens des Befragers. Die Anpassungsfähigkeit darf allerdings nicht zu einer Identifizierung mit dem Probanden führen, indem diesem z.B. ein emotionales Mitschwingen des Befragers vermittelt wird oder indem sogar sprachlich formuliert wird: "Ich glaube dir ja!". Dies würde eine Bewertung des Aussageinhaltes mit den bekannten suggestionsfördernden Implikationen darstellen. Es ist auf eine ausgewogene Balance zwischen Anpassung und kritischer Distanz zu achten. Ebenso sind bei Befragungen die aktuellen Erkenntnisse der Gedächtnispsychologie zu berücksichtigen. Versteht man gegenwärtig unser Gedächtnis als Netzwerk von Verbindungen, in dem Ereignisse nicht etwa als komplette Einheiten abgespeichert werden, so ist das Erinnern ein komplexer Vorgang, bei dem die unterschiedlichsten Bereiche des Gedächtnisses aktualisiert werden müssen, um einen einigermaßen detaillierten Sachverhalt abzurufen. Dementsprechend müssen in einer Befragung statt stereotyper Fragewiederholungen vielfach "Umwege" eingeschlagen und unterschiedliche Stichworte (cues) gegeben werden, um bestimmte Erinnerungsinhalte beim Befragten zu aktivieren. Das erfordert vom Befrager Kreativität, damit Compliance – Effekte vermieden werden. Das bedeutet: Je einfacher strukturiert, je unbeweglicher, je weniger sprachgewandt der Befragte ist, desto höher sind die Anforderungen an den Einfallsreichtum des Befragers, dessen Aufgabe darin besteht, mit verständlichen aber nicht vorhaltenden Fragen die Wege zu den potentiellen Erinnerungsinhalten zu bahnen. Schließlich erfordert eine Exploration zur Sache oft auch ein gehöriges Maß an Geduld. Ungeduldiges und voreiliges Nachfragen, zwangsläufig vermehrt in Form von Vorhalten, birgt die Gefahr eines einseitigen und hypothesenkonformen "Ermittelns". Ein guter Interviewer zeichnet sich dadurch aus, daß er Pausen zulassen und aushalten kann. Keinesfalls darf es in schwierigen Befragungssituationen nur die Alternativen "Suggestion" oder Abbruch der Exploration geben – so wie dies in Fremdgutachten formuliert und als exakte wissenschaftliche Methode vertreten wurde. Zu berücksichtigen ist schließlich, daß die Gefahr suggestiver Einflüsse mit dem Grad der Ermüdung und nachlassenden Konzentration des Probanden ansteigt. Daher sind ungeduldige Fragen am Schluß einer ermüdend langen Befragung besonders suggestionsträchtig.

Zu 2. Daß die aufgezeigten Anforderungen an die Befragungstechnik im Rahmen eines intensiven Trainings relativ schnell und erfolgreich umgesetzt werden können, hat sich bei entsprechenden Lehrgängen mit Kriminalbeamten gezeigt (PFI Neuss). In einwöchigen Lehrgängen mit jeweils 12 Teilnehmern (Kriminalbeamte mit Erfahrung auf dem Gebiet von Anhörungen und Vernehmungen) wurden zunächst die theoretischen Grundlagen der Vernehmungstechnik sowie entwicklungs-, gedächtnis- und kommunikationspsychologische Erkenntnisse vermittelt. Sodann hatte jeder Teilnehmer eine Befragung durchzuführen, die mittels Videotechnik aufgezeichnet wurde. Probanden waren Kinder im Alter von 4-12 Jahren, die über ein authentisches Erlebnis berichteten. Im Anschluß an jede Befragung wurde die Videographie einer eingehenden Analyse unterzogen.

Dabei hat sich gezeigt:

Obwohl für viele der praxiserfahrenen Lehrgangsteilnehmer das Erheben eines freien Berichtes keinesfalls eine Selbstverständlichkeit darstellte, bereitete es durchgehend keinerlei Schwierigkeiten, die befragten Kinder zu einem solchen freien Bericht zu veranlassen und daraus dann die notwendigen inhaltlichen Aspekte und sprachlichen Formulierungen für die weiterführende Befragung zu entnehmen. Als schnell erlernbar erwies sich auch die Strategie, am Ende des freien Berichtes noch einmal zum Nachdenken und Überlegen anzuregen, und dabei dem Befragten Ruhe und Geduld zu signalisieren.

Schwieriger gestalteten sich dann die unterschiedlichen Formen der trichterförmigen Befragung. Dabei bereitete die Formulierung von offenen Fragen vielfach größere Schwierigkeiten als die Gestaltung von Wahlvorhalten, was sicherlich damit zusammenhängt, daß Wahlvorhalte wesentlich konkreter sein können als offene Fragen, die ohne inhaltliche Vorgabe dennoch gezielt Sachverhalte ansteuern und teilweise auch die erwähnten "Umwege" im Denken erfordern. Auch die Formulierung von Stichwortfragen gelang zunehmend sicher und variantenreich. Weitgehend vermieden werden konnten schließlich Vorhaltfragen und Erwartungsfragen. Die Einstellung auf die kognitive Ebene des befragten Kindes gelang zunehmend sicher.

Wie oben dargelegt, erfordert eine solche Befragungstechnik Kreativität und Flexibilität vom Interviewer. Bei den genannten Lehrgängen wurde als weiterer Trainingseffekt erkennbar, daß auch diese Eigenschaften mit zunehmender Übungsdauer deutlich anstiegen, wobei das Lernen am Beispiel sich als besonders effektiv erwies.

Sehr gut umgesetzt wurde durchgehend der Abbau des sogenannten "Autoritätsgefälles"; bis auf wenige Ausnahmen konnte dabei gleichzeitig eine sachliche Distanz zwischen Befragter und Befragtem durchgehalten werden. Die Lehrgangsteilnehmer erlernten, ihren Befragungsstil und ihr eigenes Verhalten in der Befragung kritisch zu analysieren.

Zu 3. Diese Erfahrungen zeigen, daß **intensives Training** und **ständige Selbstkontrolle** innerhalb kurzer Zeit zu guten Erfolgen führen und somit eine gute Grundlage für eine adäquate und valide Datenerhebung darstellen. Im Rahmen der aussagepsychologischen Begutachtung ermöglicht und fördert die Aufzeichnung auf Tonträger eine kritische Selbstkontrolle und sollte deshalb auch unbedingt dazu genutzt werden. Das Ziel ist eine **ausgefeilte Befragungstechnik** – die Kunst nämlich, eine valide Aussage zu gewinnen, indem man den Befragten dazu bringt, eine möglichst eigenständige, ausreichend klare und konkrete Sachverhaltsschilderung abzugeben.

Die wesentlichen Grundprinzipien einer effizienten Befragungstechnik finden nach unserer Erfahrung auch nach der BGH-Entscheidung vom 30.07.99 zu wenig Beachtung in der gutachterlichen Praxis. In den eingangs zitierten Fremdgutachten, die an den vom BGH geforderten Standards ausgerichtet waren, zeigten sich Mängel in der Explorationstechnik. Um Fehlentscheidungen zu vermeiden, ist die Einhaltung der Grundprinzipien unerlässlich, um mittels aussagepsychologischer Explorationen Daten zu gewinnen, die eine gültige Basis zur Durchführung der kriterienorientierten Aussageanalyse darstellen.

## Literatur

- Arntzen, F. (1989). *Vernehmungspsychologie*. München: Beck.
- Bender, R. & Nack, A. (1995). *Tatsachenfeststellung vor Gericht. Bd. II: Vernehmungslehre*. München: Beck.
- Greuel, L., Offe, S., Fabian, A., Wetzels, P., Fabian, T., Offe, H. & Stadler, M. (1998). *Glaubhaftigkeit der Zeugenaussage. Theorie und Praxis der forensisch-psychologischen Begutachtung*. Weinheim: Psychologie Verlags Union.
- Loftus, E. F. & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13, 585-589.
- Loftus, E. F. & Zanni, G. (1975). Eyewitness testimony: The influence of the wording of a question. *Bulletin of the Psychonomic Society*, 5, 86-88.
- Steller, M. & Volbert, R. (1999). Forensisch-aussagepsychologische Begutachtung (Glaubwürdigkeitsbegutachtung). Wissenschaftliches Gutachten für den Bundesgerichtshof. *Praxis der Rechtspsychologie*, 9 (2), 46-112.



# Lay-Perspectives on criminal deviance and retributive justice<sup>1</sup>

*Margit E. Oswald, Stefan C. Klug, Jörg Hupfeld & Ute Gabriel*  
*Department of Psychology, University of Bern, Switzerland*

Laymen's thinking about how to react on criminal deviance involves many variables which are usually discussed separately in social justice research. To analyze the structure of those reactions, we created a questionnaire. The questionnaire was mailed to a random sample drawn from the telephone directory in Bern, Switzerland. 357 persons responded to items measuring constructs like relativism (Forsyth, 1985), perceived threat to society (Tyler & Boeckmann, 1997), goals of sentencing (e.g. Endres, 1992), punitivity (e.g. Oswald, 1994), or target of justice considerations (e.g. Miller & Vidmar, 1981), after reading one of three short stories on a specific crime (fraud, physical injury, assault). Results of this study do not only unveil laymen's assessment of degree and disparity of punishment but also, and above all, the underlying dimensional structure of explanatory constructs.

Behavior deviating in a negative way from social or legal norms seems to be very spontaneously associated with moral, emotional, and intentional reactions, both in victims as well as in neutral observers. These reactions have been the subject of several theoretical approaches, which are, however, discussed disconnectedly in social justice research. We will concentrate on three of them in more detail. Although these approaches deal with completely different constructs like 'sentencing goal', 'perceived threat to society', and 'relativism of moral values', we will study how they are related to each other. This seems particularly interesting since all three approaches are more or less supposed to explain why people differ in their attitudes towards punishment. In other words, they claim to provide an explanation why, not seldom, opinions are divided on the issue of whether an offender should receive parole or should be kept in prison, or whether instead of a conventional sentence an alternative procedure of restitution should be applied.

In the discussion of the individual theoretical approaches, we will have to ask critically whether it is sufficient to assume a single dimension of punitivity to systematically conceive these constructs, or whether several dimensions are needed to represent the structure of relations.

*(1) Goals of sentencing:* There exists a long tradition of research in criminology and penal law focussing on the categorization of the goals aimed at by the reactions to criminal deviance (cf. Streng, 1984; Endres, 1992). These so called sentencing goals may be differentiated roughly into 'deterrence of the actual offender', 'deterrence of potential offenders', 'rehabilitation', 'positive general prevention' (confirmation of legal norms in the public), and 'retribution of guilt'. However, this listing is not final since nowadays, e.g., we find often the goal of 'restitution of harm' and, in particular in American papers, the goal of 'incapacitation'.

Although it has been tried to align single goals of sentencing along a unique dimension of punitivity (cf. Hassemer, 1983; Kim, 1987), this has been shown to be difficult (cf. Clancy, Bartolomeo, Richardson & Wellford, 1981). Even if it made sense to assign the goal 'rehabilitation' to the lenient end, and 'deterrence of the offender' to the hard end of a dimension of punitivity, then this assignment would by no means be clear for goals of sentencing like 'positive general prevention', 'retribution', or 'restitution of harm'. Many

observations point to the fact that at least two dimensions are necessary to represent them all. Thus, the sentencing goals mentioned above may be widely separated on a second dimension but very close together or even inseparable on the dimension of 'punitivity' such that it would be difficult to rank order them according to punitivity. Looking closer at how the individual sentencing goals are differentiated, we find that they are either related to how to deal with the actual offender, or to what should be done considering a larger social entity, e. g., society as such. In this respect, several authors mention the variable 'target of justice considerations', and differentiate whether the sentencing goals are mainly directed at the offender's concern, or at that of society in general (cf. Millar & Vidmar, 1981; Karpardis, 1985; Oswald, 1994; Hupfeld, 1995). The importance of differentiating between the various 'targets of justice considerations' becomes particularly comprehensible if we focus on the tension which is experienced by persons who have to decide upon the kind and amount of punishment of an offender – Zipf (1973), e. g., talks about 'antinomy of sentencing goals'. We can only seldom assume that a particular punishment is accepted by all people concerned, and thus, the choice of a particular punishment means, in most cases, also a decision on which target is weighted heavier. Thus, the decision in favor of a lengthy therapy will serve more the needs of the offender but less those of society because of the high costs associated with it. By the way, the distinction between different 'targets' of sentencing goals made here is clearly related to the differentiation of sentences discussed recently about whether sentences are uttered from a micro or macro perspective (Ellsworth, 1978; Tyler et al., 1997). Sentences of the micro perspective are focussed, above all, on actual participants in the event under consideration. However, if the event is looked at from a 'higher' perspective, e. g., from a societal perspective, the sentence is created from a macro perspective. Dependent on the perspective taken, very different judgments on justice may occur.

However, it is by no means clear whether the micro/macro perspective or the target of justice considerations might be a second dimension independent of the punitivity, and whether the goals of sentencing could be meaningfully placed in this space.

(2) *Perceived threat to society*: Emotional reactions on criminal deviance might be e.g. fear of becoming a victim oneself (cf. Miller & Widmar, 1981) or fear that the cohesion of society will be threatened (cf. Tyler et al., 1997). Because criminal deviance indicates an erosion of values that is offensive and threatening to the normative system that people in groups rely upon, punishment maintains the important symbolic function to restore the structure of society and the positive social characteristics of the group (Tyler et al., 1997, p. 200). Contrary to this perspective, fear of crime is seen as instrumentally related to punishment. Punishment is considered as a direct consequence of perceived threat to self-interest because people who are afraid of becoming a victim seek punishment out of a desire to control the future behavior of the rule breaker and other potential rule breakers. However, according to Tyler & Boeckmann (1997), fear of becoming a victim (fear of crime) is significantly less suited in predicting punitivity than perceived threat to society.

The importance of the symbolic function of punishment, outlined by Boeckmann & Tyler (1996), should not only result in a close connection between perceived threat to society and punitivity but also in a strong relation between punitivity and preferences of those sentencing goals, which emphasize the symbolic rather than the instrumental function of punishment. Furthermore, according to Tyler & Boeckmann (1997, p. 132),



people who consider the problem of criminal deviance from a societal perspective (macro perspective) should be more punitive in order to symbolically reassert the status of the violated rules.

However, whether people are more punitive who perceive criminal deviance on a macro rather than on a micro perspective and whether people are more punitive who prefer the symbolic and expressive function (e.g. the sentencing goal 'positive general prevention') rather than the instrumental function of punishment (e.g. the goals of deterrence) remains an open empirical question.

(3) *Relativism of norms*: Other theories attempt to find principles on which the moral evaluation of deviant behavior may be based. E. g., Schlenker & Forsyth (1977) and Forsyth (1980, 1985) assume that people differ in the degree they consider rules as universally valid in their evaluation of moral issues, or that they are to a different degree willing to revise their moral judgment by taking into account the particularity of the specific situation and of the person acting in it. This dimension has been called 'relativism' by the authors. Also, more recent interpretations of the distinction between 'ethics of care' and 'ethics of justice' introduced by Gilligan (1982) seem to be related to a relativism of norms. According to Baier (1995) and Daly (1989) we may consider ethics of justice as an inclination to resort to formal rules in case of conflict, while ethics of care imply an inclination to see the weakness and impotence of the individual who could not (yet) act with equal rights in the competition for keeping legal rules, and whose particularities and needs should be respected, therefore.

McHoskey (1996) demonstrated in a study that the preference for universal moral rules correlates positively with a punitive attitude towards offenders. This result, however, may also be due to the specific operationalization of the construct of 'relativism'. Based upon theoretical considerations, this relation is by no means mandatory since the use of universally accepted moral rules does not a priori imply anything about the resulting punitivity: The use of a universally accepted rule would be strongly dependent on what kind of rule is applied, and an application of the rule dependent on person and situation would let us expect a considerable variability of the measure of punishment over different cases since, dependent on the specific person and situation, the punishment might be high in one case, e.g. a sexual abuse, and low in the other, e.g. a property offence. Also, the relation between punitivity on one hand, and the differentiation between ethics of care and ethics of justice on the other, is not unambiguous since formal rules may imply differently harsh reactions, and ethics of care could result in differently hard punishments, dependent on the status and power potential of the culprit (cf. Hupfeld, 2000). Based on theoretical considerations, it is possible that the dimension of 'relativism' might be almost independent of the dimension of punitivity. The present paper analyzes the central structures of the different reactions elicited by criminal deviance. Above all, we have to look for the systematic relations of the variables discussed in respect to punitivity, and which dimensions independent of punitivity might be appropriate to structure these relations. Further dimensions could be, e. g., the 'target' of justice considerations, or the readiness to apply rules generally compulsive or dependent on the respective person and/or situation.

## Method

*Sampling and Population:* A random sample was drawn from the telephone directory in Bern, Switzerland. Questionnaires were sent on agreement after a short personal contact by phone. The telephone contact also served the need to pre-exclude persons who either had not reached the voting age of eighteen, were not of Swiss nationality, did not speak German sufficiently, or did not live in the predefined area. Participants received a small amount of money for their participation (10 Swiss francs). 54 % of the persons agreed to fill out the questionnaire. 84 % of those sent the questionnaire back. Relating the amount of questionnaires that were sent back to the total of persons who either disagreed, agreed and those who could not be contacted (persons not reached within three weeks) still gives a response rate of 33.8 %.

After a final selection (six cases had to be dismissed) our sample consists of 357 persons (39.5% male) with a mean age of 46.3 years (SD = 17.77).

*The Questionnaire:* The questionnaire consisted of one of three short stories of a crime (fraud, physical injury or assault) which were pretested according to their perceived seriousness in relation to other crimes.<sup>2</sup> The story was followed by a series of questions to measure the following constructs (see below). All constructs were pre-tested in two different ways. First, an early draft of the questionnaire was filled in by 54 pupils recruited in a school in Bern (mean age = 23.33, SD = 4.48). Classical item-analysis was executed on this data. Second, the items were rated by twelve experts (advanced students of psychology) according to their quality to measure these constructs. Information gained by these two methods provided the possibility for a high quality item selection.

*Constructs:* Punitivity has been conceived both as a trait (cf. Table 2) and as an actual suggestion of a sentence referring to the respective case. People had to rate a given sentence according to its leniency or harshness by marking a cross on a 120 mm line (M = 61.28, SD = 24.36). To differentiate these two measures of punitivity, we called the trait variable 'punitivity', and the actually suggested punishment 'harshness of penalty'. Six aims of punishment were chosen -- those most often quoted in the literature -- and briefly explained to the subjects (cf. Table 1). In order to measure the individual degree of preference for these aims, participants had to rank-order the importance of different sentencing goals according to their personal importance.

Table 1: List of sentencing goals and their descriptions

<b>Sentencing goal</b>	<b>Description of the goal</b>
Retribution of guilt	<i>The offender should expiate the crime proportionately to his guilt.</i>
Deterrence of the actual offender	<i>The offender should be kept from committing other crimes in the future.</i>
Rehabilitation	<i>The offender should be treated and re-educated about proper attitudes and behavior.</i>
Deterrence of the potential offenders	<i>Other people should be deterred from committing a similar crime.</i>
General positive prevention	<i>The punishment of the offender should strengthen the legal consciousness in the public.</i>
Restitution of harm	<i>The offender should make a restitution for the damage done.</i>

In measuring the variable 'target of justice considerations' we started from the theoretical consideration that each reaction to a norm-deviant behavior has consequences not only to the offender, but also to the society and to the victim. E.g., a lifelong prison sentence implies for the society that a high amount of tax money will be used up, and for the victim it means that he or she will not be endangered by this individual offender any more (which does not mean that the victim's harm was restituted), and for the offender it means the loss of any social integration. In considering possible reactions in a certain case, a person can now be oriented more or less to one of the three targets (victim, society or offender). To measure this construct, people were asked to distribute 100 points according to their personal opinion about the importance of the three targets in sentence decision making.

All other variables have been measured by means of scales which were particularly designed for this study (cf. below). Except for the items of the punitivity scale, all other scales represent case specific items (see Table 2). They refer to the offender of the different crime scenarios by name (Beat N.).

Table 2: Variables measured by means of scales in this study

<i>Variables</i>	<i>Examples</i>	<i>Cronbach's Alpha (No. of items)</i>
<b>Punitivity</b>		0.83 (4)
<i>High punitivity</i>	<i>Offences against laws and norms in our society should be punished as severely as possible.</i>	0.73 (2)
<i>Low punitivity</i>	<i>In my opinion one should not comply with the public call for harsher penalties.</i>	0.74 (2)
<b>Relativism</b>	<i>Depending on the situation, one will refer to different standards in order to evaluate Beat N.'s behavior. There are no absolute moral principles.</i>	0.42 (3)
<b>Absolutism</b> (=low relativism)	<i>To evaluate Beat N's behavior I solely have to consider which general moral principles it contravenes or possibly corresponds to.</i>	0.45 (3)
<b>Ethics of Justice</b>	<i>The evaluation of Beat N.'s behavior should be based on a "higher order" justice, i.e. that one should draw upon rules of judgment that are transferable to other cases.</i>	0.22 (3)
<b>Ethics of Care</b>	<i>In the trial of Beat N.'s case, one must find a specific solution which takes into account the needs, interests, concerns and troubles of all people directly involved, also those of the victim and the offender.</i>	0.63 (3)
<b>Perceived Threat to Society</b>		0.73 (6)
<i>High perceived threat to society</i>	<i>Because of people like Beat N., social intercourse becomes rougher and more irreverent and therefore, living together becomes worse.</i>	0.70 (3)
<i>Low perceived threat to society</i>	<i>Offences against social rules like the one by Beat N. have always occurred. So far, society has not been damaged by this and neither will it in the future.</i>	0.78 (3)

*Statistical Analysis:* Most of the items had to be answered on a 5 point likert-scale ranging from 'agree' to 'disagree'. All dimensional analyses were conducted by multidimensional scaling analysis (MDS) based on Pearson correlation coefficients using the Kruskal (mono) algorithm included in SYSTAT 8.0. Items for the two theoretical end-points of the constructs 'punitivity', and 'perceived threat to society' were treated as separate scales (see Table 2).

## Results

*Dimensional structure of sentencing goals:* Let us consider first the relations between individual goals of sentencing and the potential inclination of the respondents to punish an offender severely, both as his or her general tendency, and as related to the respective case. An MDS analysis unveils that a two-dimensional view explains a rather large portion of the variance (RSQ = .96), both absolutely and as compared with a one-dimensional analysis (RSQ = .77). Figure 1 shows that case-related harshness of penalty and individual punitivity are highly correlated, as expected, and that the goals of sentencing 'rehabilitation' and 'deterrence of the actual offender' are almost identical with the respective ends of the dimension of punitivity. Almost orthogonal to this, i. e., independent of punitivity we find, opposing each other, the goals of sentencing 'restitution of harm' and 'general positive prevention'.

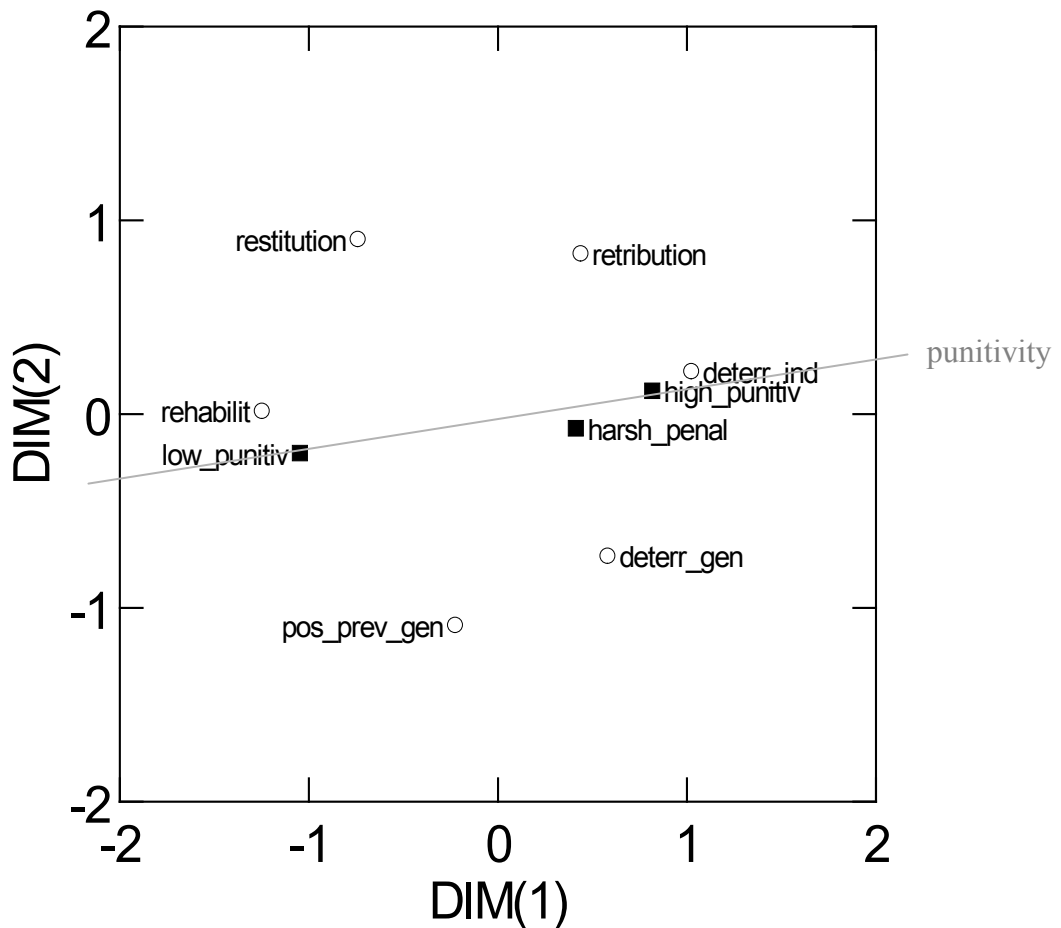


Figure 1: Relation between sentencing goals and punitivity (Stress = 0.07, RSQ = 0.96; rehabilit = rehabilitation, deterr\_ind = deterrence of the actual offender, deterr\_gen = deterrence of the potential offenders, pos\_prev\_gen = general positive prevention, retribution = retribution of guilt, restitution = restitution of harm, low\_punitiv = low punitivity, high\_punitiv = high punitivity, harsh\_penal = harshness of penalty)

The results confirm our assumption that the different goals of sentencing cannot be aligned on one dimension. Attempting to interpret the second dimension we notice that the goals of sentencing in the lower half of the figure refer to the society afflicted by the deviant behavior (general positive and deterrence of the potential offenders), whereas the goals of sentencing in the upper half refer to the concrete individuals involved in criminal acts. To test the interpretation that this might be a dimension 'considerations of justice on a micro versus macro level', we extended the analysis by the variable 'target of justice considerations'.

Indeed now it becomes clear that a strong focusing on the need and interests of the society (macro perspective) coincides with a clear preference for the goal 'general positive prevention' whereas the concern of the society moves to the background in case of a preference for the goal 'restitution of harm' (cf. Figure 2). Instead, here the concern of actual persons, i. e., of the offender and/or of the victim are in the focus of penal elaborations (micro perspective). Furthermore we notice that the offender's concern is so much lesser important the higher the degree of punitivity. According to these results we have to assume that the degree to which the offender's concern is respected seems to be very important for the assessment of the harshness of the penalty. A 'marginalizing' of the offender's concern seems to be related to a high inclination to coercive measures. Independent of this degree of offender orientation or punitivity, people assume either a micro- or a macro perspective in their elaborations of punishment, i.e., they are either more concerned about the restitution of justice for the actually concerned persons, or about a clarification of justice on a societal level. It is interesting to note that victim-orientation is more closely related to retribution than to restitution. Maybe people think that state-regulated retribution includes some form of restitution, which is usually not the case in penal law.

In spite of the rather clear result that there exists a systematic relation between goals of sentencing, and weighting of targets of justice considerations, we still have to keep in mind that statements on weights are not to be considered as absolute measures but rather referring to the respective average weighting. In the offences studied here, we find, on the average, a very high weighting of the victims concern such that even in case of a low weighting of the victim's concern, as well as in case of a preference of the goal of re-socialization, the victim's concern still got the highest weight. Nevertheless, we have to notice that there are considerable and statistically significant differences in the weighting of 'targets', depending on the respective goal of sentencing.

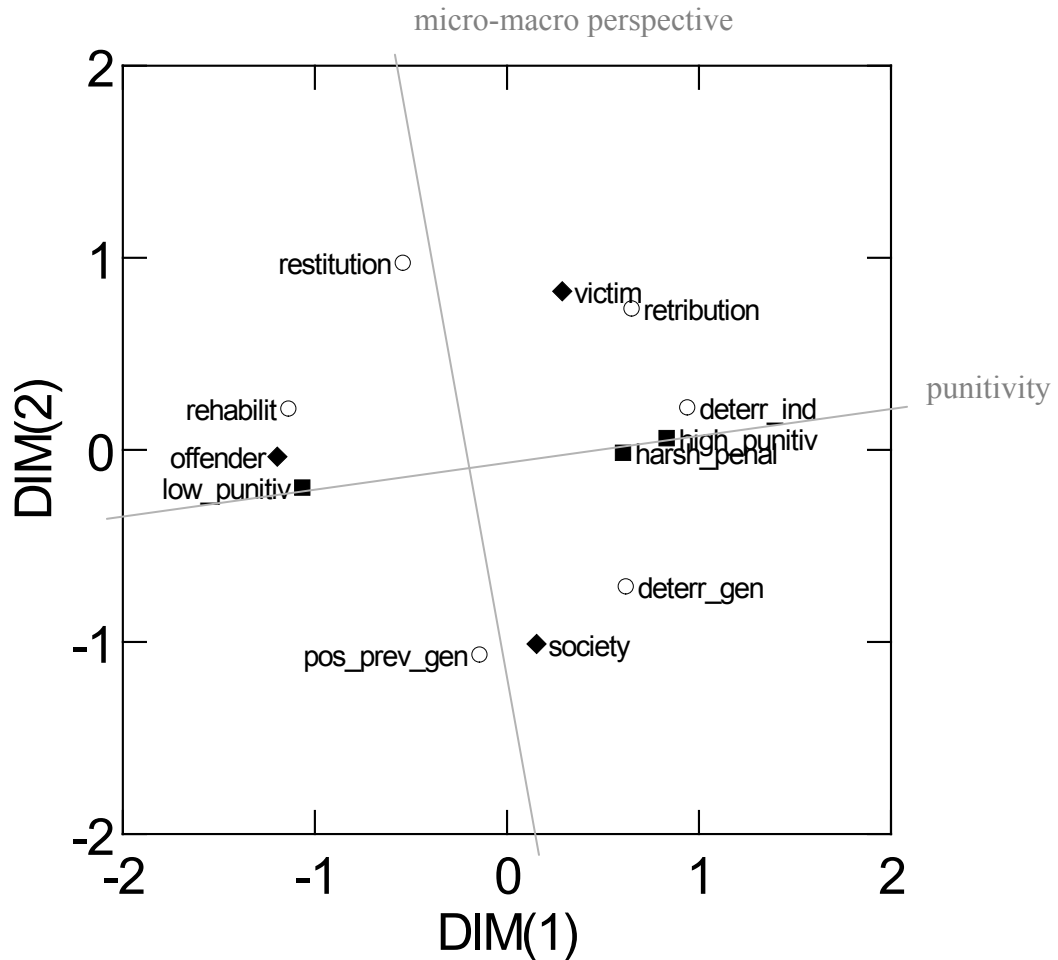


Figure 2: Configuration of the sentencing goals on the two dimensions punitivity and micro-macro perspective (Stress = 0.10, RSQ = 0.93; rehabilit = rehabilitation, deterr\_ind = deterrence of the actual offender, deterr\_gen = deterrence of the potential offenders, pos\_prev\_gen = general positive prevention, retribution = retribution of guilt, restitution = restitution of harm, low\_punitiv = low punitivity, high\_punitiv = high punitivity, harsh\_penal = harshness of penalty; offender, victim, society are the three targets of punishment considerations)

*Perceived threat to society and punitivity:* Tyler et al. (1997) assume that, above all, the degree to which a person perceives the moral cohesion of society as threatened by criminal deviance determines his or her need for punishment. Furthermore, they propose a 'relational approach', i.e., that punishment has primarily a symbolic and not an instrumental function. Punishment is mainly applied in order to symbolically reassert the status of the violated rules and less in order to control offenders behavior. The MDS analysis confirms the close relation between 'societal threat' and punitivity as postulated by Tyler et al. (1997) (cf. Figure 3).

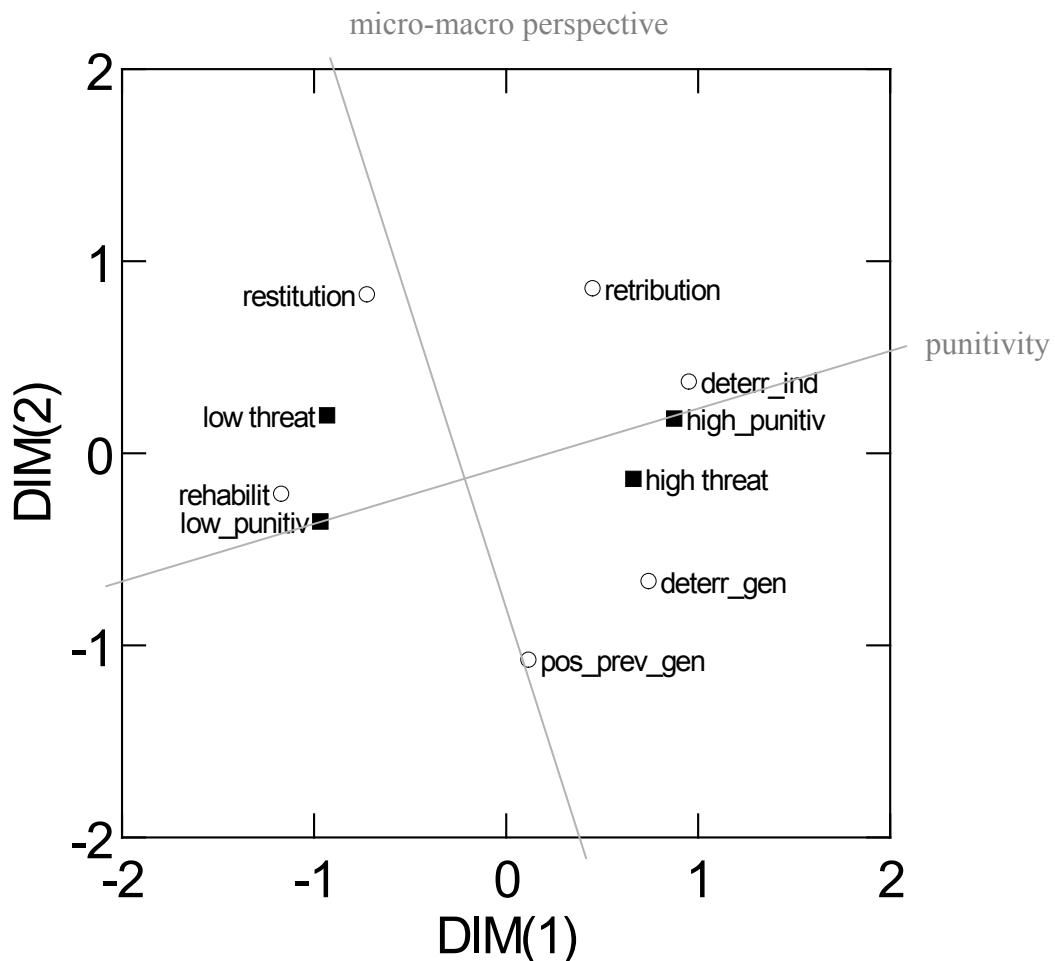


Figure 3: Relation between perceived threat to society, punitivity and sentencing goals (Stress = 0.10, RSQ = 0.93; rehabilit = rehabilitation, deterr\_ind = deterrence of the actual offender, deterr\_gen = deterrence of the potential offenders, pos\_prev\_gen = general positive prevention, retribution = retribution of guilt, restitution = restitution of harm, low\_punitiv = low punitivity, high\_punitiv = high punitivity, high threat = high perceived threat to society, low threat = low perceived threat to society)

However, the theoretical assumption that the need for punishment associated with perceived threat to society should go along with sentencing goals who emphasize the symbolic rather than the instrumental function of punishment has not been confirmed. As can be seen from Figure 3, with high perceived threat to society the punishment obtains increasingly the function of deterring actual or potential offenders, whereas a preference of goals who emphasize the symbolic function of punishment, like 'general positive prevention' or 'retribution' goes along with a much lower need for punishment.

*Moral dimensions and punitivity:* The readiness to apply ethical rules depending on persons and situations ('relativism vs. absolutism') and 'ethics of care versus ethics of justice' are rather closely related to each other, with high relativism being conjoint with ethics of care, and low relativism (absolutism) with 'ethics of justice'. This is not surprising as the definitions of the constructs overlap (cf. above). However, it is interesting that, viewed in two dimensions, these variables do not coincide with the dimension of punitivity but seems to be almost independent of it (cf. Figure 4). However, the increase in variance explained with a two-dimensional solution as



compared to a one-dimensional solution, though, is very small (RSQ = .97 as compared to RSQ = .95). For that reason, our theoretically derived assumption of independence of relativism from punitivity is not completely confirmed. Yet, we have to respect that the kind of cases constructed by us as well as the facts that the respondents had to refer to only one case, respectively, provide only a small chance to discover a potential variability, i. e., a punitivity varying from case to case with high relativism.

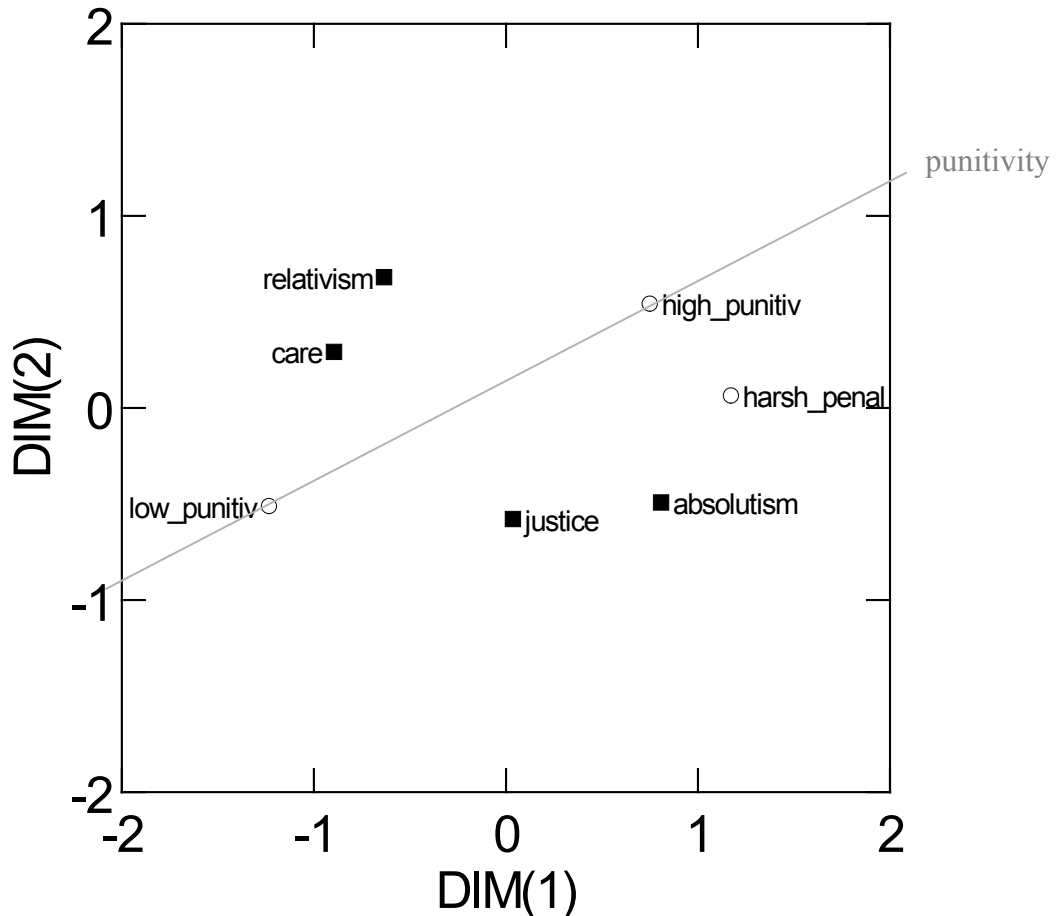


Figure 4: Relation between relativism and punitivity (Stress = 0.06, RSQ = 0.97; low\_punitiv = low punitivity, high\_punitiv = high punitivity, harsh\_penal = harshness of penalty, care = ethics of care, justice = ethics of justice)

*Adequacy of a two- or three-dimensional space:* After having unveiled, in two separate analyses, in one case a very clear cue, and in the other a rather weak one, that there exists a dimension independent of the dimension of punitivity, namely 'justice considerations on a micro- versus macro perspective', and 'relativism', we finally have to ask how these dimensions relate to each other if we analyze all variables simultaneously. With a two-dimensional solution, we obtain, as compared to previous findings, a relatively small portion of variance explained (RSQ = 0.88), and the variable 'relativism' is almost congruent with punitivity (cf. Figure 5).

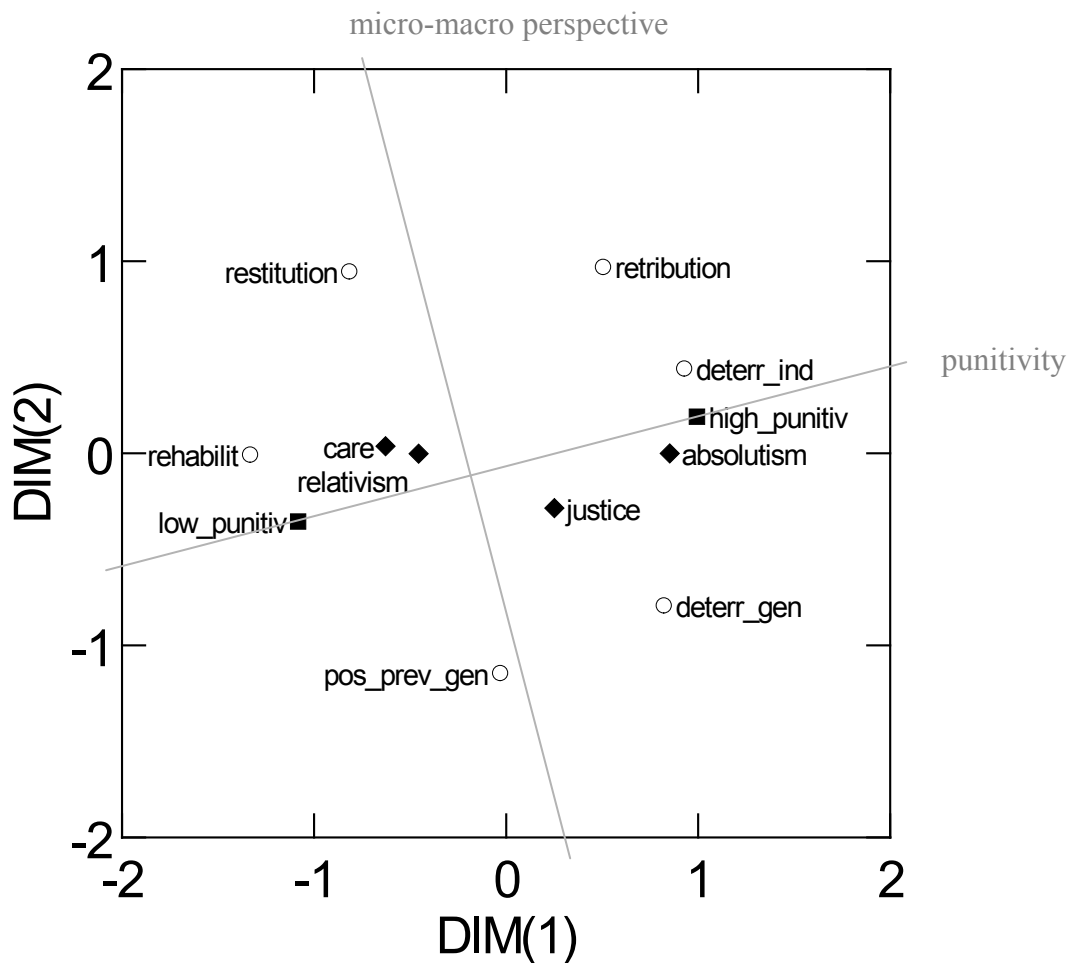


Figure 5: Relation between relativism, punitivity and micro-macro perspective – a two dimensional solution (Stress = 0.14, RSQ = 0.88; rehabilit = rehabilitation, deterr\_ind = deterrence of the actual offender, deterr\_gen = deterrence of the potential offenders, pos\_prev\_gen = general positive prevention, retribution = retribution of guilt, restitution = restitution of harm, low\_punitiv = low punitivity, high\_punitiv = high punitivity, harsh\_penal = harshness of penalty, care = ethics of care, justice = ethics of justice)

If we consider a three-dimensional solution the proportion of variance explained does not increase largely, but at any rate by 8 %. For better clearness, we show in a three-dimensional figure only the central variables which are respectively positioned at the ends of the dimensions postulated by us (cf. Figures 2 and 3). However, the coordinates of the variables in Figure 6 were taken from the three-dimensional analysis of the variables presented in Figure 5. We see very clearly that the dimensions 'punitivity' and 'justice considerations on a micro versus macro level' persist. The variables 'relativism' and 'ethics of care versus ethics of justice' seem to form the vertical third dimension although they are not completely independent of the dimension of punitivity.

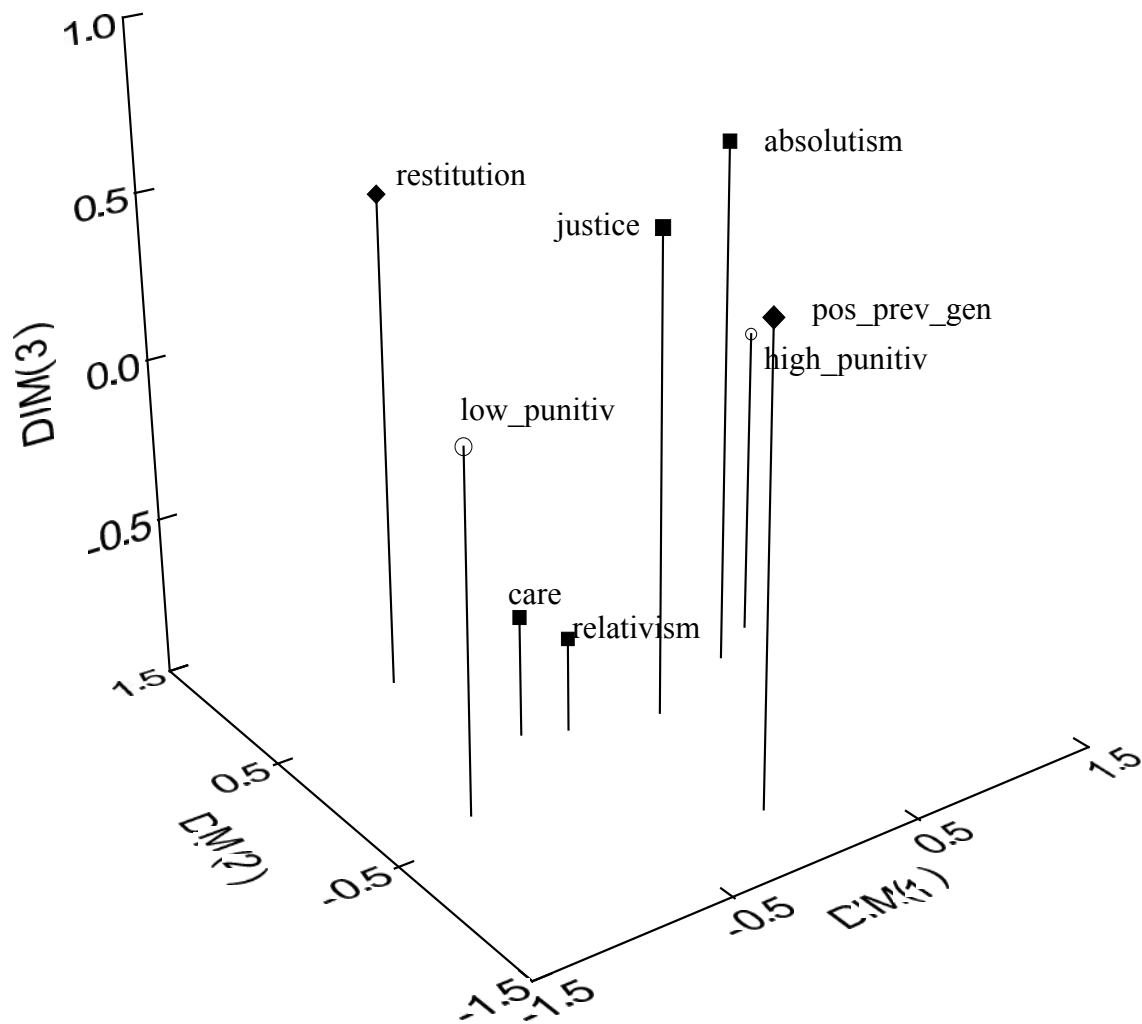


Figure 6: Relation between relativism, punitivity and micro-macro perspective – a three dimensional solution (Stress = 0.05, RSQ = 0.97; low\_punitiv = low punitivity, high\_punitiv = high punitivity, care = ethics of care, justice = ethics of justice, pos\_prev\_gen = positive general prevention, restitution = restitution of harm; not all variables presented here)

*Stability of the results:* We would like to point at the fact that these results are not specific to our sample; the same study run on a random sample (N = 334) in Erlangen (Germany) revealed the very similar results. MDS analyses on the basis of the data of the second sample revealed variable configurations almost identical to those described before (presented above). Even jointly fitting the matrices of both cities in a common space by means of INDSCAL (individual differences scaling) resulted in the same structural results.

### Final remarks

The purpose of this paper was to study the dimensional structure of laymen's reactions to deviant behavior. Particularly, we were interested in the relations between punitivity and several other relevant variables that have been discussed in such different fields like moral philosophy, sociology, criminology, and social psychology. The empirical findings of the present research show, as well as most of the other

research conducted in this field, that punitivity is a very dominant dimension. This dominance might in our case be partly due to some statistical reasons, especially the use of positive as well as negative poles for punitivity and some other highly associated constructs (e.g. integration/exclusion; leniency/harshness of penalty) in MDS. The respective positive and negative indicators show the most extreme negative correlations and therefore might dominate the results of the multidimensional analyses. But even if we cannot answer the question whether punitivity is the most important dimension that is underlying laymen's justice considerations, it surely is an important one.

Some of the sentencing goals can be ordered according to this dimension. On the lenient side we find the goals of rehabilitation and positive individual prevention. On the harsh side we find different forms of deterrence as well as retribution. But laymen's judgments are obviously also guided by another dimension which seems to be almost independent from punitivity. On the one side of this second dimension we find the sentencing goal of restitution that might be understood as the assurance of the norm validity by restoration of equity at the micro level (see Tyler et al., 1997, p. 104). The macro counterpart affirms norm validity to the public as mentioned by the goal of positive general prevention.

And even a third interesting dimension shows up in the data. On one side of this ethical dimension we find relativistic thoughts, showing the willingness to consider i.e. time, circumstances, and the wishes and needs of the involved people. On the other side we find moral absolutism, emphasizing abstract norms and rules. In order to prove whether this ethical dimension is related to punitivity future research should include some specially constructed crime vignettes. As already mentioned these vignettes should include for example different combinations of status and power with respect to the offender and the victim.

Despite some open questions, the present study is an important step toward an integration of the variety of different approaches to the field of retributive justice. One conclusion that can be drawn from our results is that retributive justice considerations can not be reduced to a single punitivity dimension, but have a lot of faces.

## References

- Baier, A. C. (1995). *Moral Prejudice. Essays on Ethics*. Cambridge: Harvard University Press.
- Clancy, K., Bartolomeo, J., Richardson, D. & Wellford, C. (1981). Sentence decisionmaking: The logic of sentence decisions and the extent and sources of sentence disparity. *The Journal of Criminal Law & Criminology*, 72 (2), 524-554.
- Daly, K. (1989). Criminal Justice Ideologies and Practices in Different Voices. Some Feminist Questions about Justice. *International Journal of the Sociology of Law*, 17, 1-18.
- Ellsworth, P. (1978). Attitudes towards capital punishment: From application to theory. Paper presented at the Annual Meeting of the Society for Experimental Social Psychology, Stanford, CA.
- Endres, J. (1992). *Sanktionszweckstellungen im Rechtsbewusstsein von Laien [Sentencing goals in the awareness of justice in laymen]*. Frankfurt a. M.: Peter Lang.
- Forsyth, D. R. (1980). A Taxonomy of Ethical Ideologies. *Journal of Personality and Social Psychology*, 39, 175-184.

- Forsyth, D. R. (1985). Individual differences in information integration during moral judgement. *Journal of Personality and Social Psychology*, 49, 264-272.
- Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge: Harvard University Press.
- Hassemer, W. (1983). Alternativen zum Schuldprinzip [Alternatives to the principle of guilt]. In H. M. Baumgartner & A. Eser (Hrsg.), *Schuld und Verantwortung* (S. 89 – 107). Tübingen: Mohr.
- Hupfeld, J. (1995). The structure of goal-directed cognitions in judges of juvenile courts. In J. J. Hox, G. J. Mellenbergh & P. G. Swanborn (eds.), *Facet Theory: Analysis and Design* (pp. 145-152). Zeist: Setos.
- Hupfeld, J. (2000). Frauen, eine Gefahr für das Strafrechtssystem? [Women, a danger for the legal system]? In B. Lischetti (ed.), *Frauen im Recht* (pp. 43-59). Bern: eFeF.
- Kapardis, A. (1985). *Sentencing by English magistrates as a human process*. Nicosia, Cyprus: Asselia Publishers.
- Kim, Y. W. (1987). *Zur Fragwürdigkeit und Notwendigkeit des strafrechtlichen Schuldprinzips* [About doubtfulness and necessity of the legal principle of guilt]. Edelsbach: Gremer.
- McHoskey, J. W. (1996). Authoritarianism and ethical ideology. *The Journal of Social Psychology*, 136, 709-717.
- Miller, D. T. & Vidmar, N. (1981). The social psychology of punishment reactions. In M. J. Lerner & S. C. Lerner (Eds.), *The Justice Motive in Social Behavior* (pp. 145-172). New York: Plenum.
- Oswald, M. E. (1994). *Psychologie des richterlichen Strafens* [The psychology of judges' sentencing]. Stuttgart: Enke.
- Schlenker, B. R. & Forsyth, D. R. (1977). On the Ethics of Psychological Research. *Journal of Experimental Social Psychology*, 13, 369-396.
- Streng, F. (1984). *Strafzumessung und relative Gerechtigkeit* [Sentencing and relative justice]. Heidelberg: Decker.
- Tyler, T. R. & Boeckmann, R. J. (1997). Three strikes and you are out, but why? The psychology of public support for punishing rule breakers. *Law & Society Review*, 31, 237-265.
- Tyler, T. R., Boeckmann, R. J., Smith, H. J. & Huo, Y. J. (1997). *Social justice in a diverse society*. Boulder, Colorado: Westview Press.
- Zipf, H. (1973). *Kriminalpolitik. Eine Einführung in die Grundlagen*. [Crime policy. An introduction]. Karlsruhe: Müller.

<sup>1</sup> The authors wish to express their thanks to Dirk Wendt, Marianne Aeberhard and Eliane Schneider as well as Franz Streng, head of the parallel study in Erlangen (Germany), for their valuable contributions. This research was supported by the Volkswagen-Stiftung (Germany).

<sup>2</sup> Results of analysis using the independent variable "story" will not be presented here, while it is still important to notice, that the main part of the questions referred to the specific case (state) and do therefore not measure solely personality traits.



# Emotionen im Prozeß der Glaubwürdigkeitsbeurteilung

*Kerstin Panhey, Günter Köhnken & Frank Eggert  
Institut für Psychologie der Universität Kiel*

## **Hintergrund und Fragestellung**

Im Prozeß der Glaubwürdigkeitsbeurteilung spielen Emotionen eine große Rolle. Dabei konzentriert sich die Aufmerksamkeit hauptsächlich auf den Beitrag von Emotionen zur Entdeckung von Lügen, und es wurde in vielfältigen Untersuchungen (Überblicke z.B. bei Köhnken, 1990; Vrij, 2000; Ekman & O'Sullivan, 1991; Ekman, 1992) gezeigt, daß Emotionen und emotionales Ausdrucksverhalten – insbesondere nonverbaler Art – einen Beitrag zur Unterscheidung wahrer und falscher Aussagen leisten können. Auch die Anwendung von Polygraphen im forensischen Kontext beruht auf diesem Prinzip (siehe etwa Steller, 1987).

Nach dieser Ansicht stellen Emotionen in Kommunikationssituationen einen potentiell verräterischen Aspekt dar, so daß der Sender bemüht sein dürfte, insbesondere im Falle einer Lüge den Emotionsausdruck zu kontrollieren und dem bei einer wahren Aussage anzugleichen. Der jeweilige Empfänger wird – sofern er eine Lüge zumindest für möglich hält – entsprechend gut auf den Emotionsausdruck seines Gegenübers achten und ggf. versuchen, Gewißheit zu erlangen.

Im Gegensatz zu diesen Annahmen gibt es aber auch Gründe für die Annahme einer strategischen Rolle von Emotionen (siehe etwa Andersen & Guerrero, 1998) bei der Übermittlung wahrer und falscher Informationen. Dabei werden Emotionen nicht als unerwünschtes und zu kontrollierendes „Nebenprodukt“, sondern als funktional eingesetztes Kommunikationselement betrachtet. Ausgehend von dem Ansatz der kognitiven Kapazität von Ellis und Ashbrook (1989), nach dem die vorhandene kognitive Kapazität durch Emotionen gezielt verringert werden kann, haben wir ein Modell (siehe Abbildung 1) entwickelt, nach dem ein Sender Emotionen gezielt zur Tarnung einer Lüge einsetzen kann. Dabei wird der Empfänger mittels emotionaler Inhalte oder emotionaler Darstellung der Informationen beschäftigt und seine kognitive Kapazität zur Überprüfung und Enttarnung der Lüge – wie etwa die Analyse interner Inkonsistenzen der Aussage – verringert. Somit wird die Wahrscheinlichkeit der Aufdeckung einer Lüge herabgesetzt.

Nach diesem Modell wäre somit zu erwarten, daß ein Sender bei einer Lüge den Emotionsausdruck nicht verringert, sondern betont. Diese Emotionen werden entsprechend vom Empfänger wahrgenommen und können sich auf ihn übertragen – bei einer lustigen Begebenheit steigt seine Stimmung, bei einer unerfreulichen ärgert er sich mit dem Erzähler, empfindet Mitleid etc. All diesen verschiedenen Emotionen beim Empfänger ist gemeinsam, daß dieser durch sie kognitiv beansprucht wird und eine Lüge des Senders nicht so sicher bemerkt, wie er es in aufmerksamer, aber emotional neutraler Stimmung könnte.

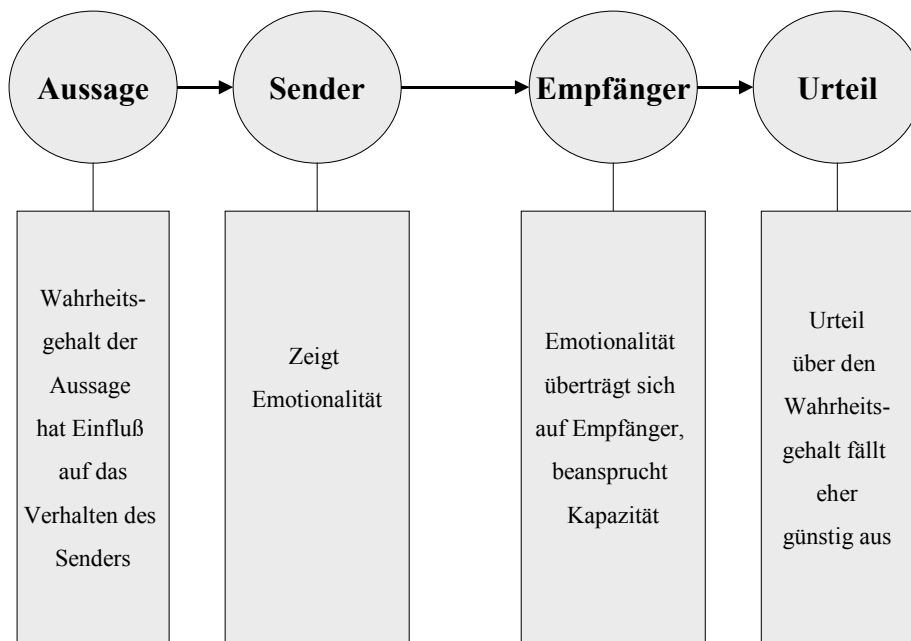


Abbildung 1: Ein Modell zur strategischen Rolle von Emotionen bei Glaubwürdigkeitsentscheidungen

Dieses Modell ist keineswegs auf den forensischen Kontext beschränkt, obwohl seine Gültigkeit auch für Situationen mit polizeilicher oder richterlicher Befragung plausibel ist. Ein Kind, das länger als erlaubt bei einem Spielkameraden war, kann bei einer dramatischen Geschichte von einer entlaufenen Katze (sofern sie nicht zu übertrieben ist) ebenso auf eine weniger genaue Prüfung der Umstände hoffen, wie ein Verdächtiger mit einer dramatischen Begründung für seine Anwesenheit an einem bestimmten Ort. Ob die betreffenden Personen letztendlich mit ihrer Lüge Erfolg haben, hängt von vielen Faktoren ab, aber unter sonst gleichen Umständen ist es anzunehmen, daß ein emotional beteiligter Zuhörer in stärkerem Ausmaß bereit sein wird, dieser Aussage Glauben zu schenken. Das Modell macht keine spezifischen Angaben über die Art der Emotionen oder ihrer Übermittlung. Generell ist der Einsatz positiver wie negativer Emotionen ebenso denkbar wie auch die Übermittlung über nonverbales Verhalten oder inhaltliche Unterschiede des Emotionsgehalts.

Die von uns verwendete Konzeption von Emotionen und damit auch die Erfassung der Emotionen in unserer Untersuchung orientiert sich an einem Emotionsmodell von Rolls (1990, 1999), welches in Abbildung 2 dargestellt ist. Dieses Modell ist neben seiner Anschaulichkeit insbesondere durch seine lerntheoretische und neurophysiologische Fundierung ausgezeichnet.



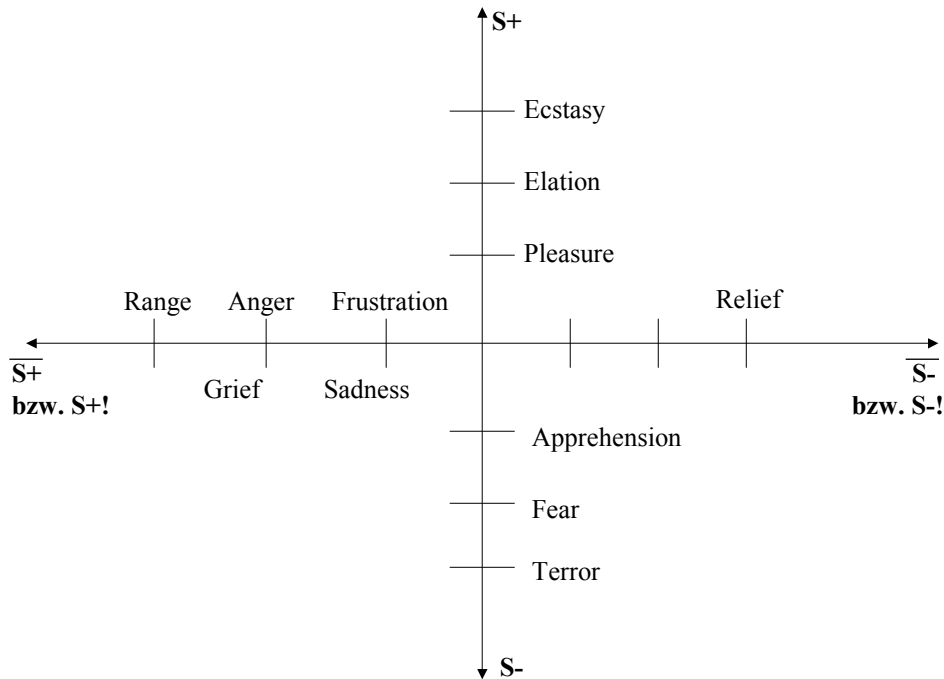


Abbildung 2: Eine positive Konsequenz (S+) führt zu positiven Gefühlen wie Freude, eine negative Konsequenz (S-) zu Besorgnis und Angst. Das Ausbleiben oder Ende einer negativen Konsequenz ( $\bar{S}$ - bzw. S-!) führt zu Erleichterung, das Ausbleiben oder Ende einer positiven Konsequenz ( $\bar{S}$ + bzw. S+!) zu Trauer oder Ärger (Rolls, 1990, p. 163).

Anhand des oben beschriebenen Modells zur strategischen Rolle von Emotionen wurden mehrere Hypothesen über den Mechanismus entwickelt. Einige davon wurden im Rahmen der folgenden Fragestellungen untersucht:

1. Unterscheidet sich die Darstellung von wahren und falschen Geschichten in emotionaler Hinsicht?
2. Zeigt der Empfänger beim Anhören der Geschichten ein ähnliches emotionales Bild wie der Sender?
3. Zeigt die Emotionalität des Empfängers einen Zusammenhang mit dem Urteil über den Wahrheitsgehalt der Geschichten?

## Methode

Die Datenerhebung fand im Sommer 1999 an der Dalhousie University in Halifax, Nova Scotia, Kanada und im Frühjahr 2000 an der Christian-Albrechts-Universität in Kiel statt. Als Reizmaterial wurden 12 auf Video aufgenommene Geschichten verwendet, die von Studentinnen der Christian-Albrechts-Universität erzählt wurden. Die Geschichten waren jeweils etwa 5 Minuten lang. Die Hälfte beschrieb wahre – d.h. tatsächlich erlebte – Begebenheiten, während die andere Hälfte der Geschichten ausgedacht war, aber wie eine tatsächlich erlebte Begebenheit erzählt werden sollte.

Die Versuchspersonen wurden in mehrere Gruppen eingeteilt, von denen eine die Videobänder sowohl sah als auch hörte (vollständige Informationen, N = 48), eine weitere Gruppe sah die Erzählerinnen nur, ohne sie zu hören (nur Bild, N = 10) während eine dritte Gruppe aus kanadischen Studenten ohne Deutschkenntnisse bestand, die die Geschichten sah und hörte, jedoch den Inhalt nicht verstehen konnte (keine Kenntnis über den Inhalt der Geschichten, N = 17).

Die Versuchspersonen sahen die Videos in kleinen Gruppen von etwa acht Personen und füllten nach jeder Geschichte einen Fragebogen aus, der sich entweder auf die Emotionen der Erzählerin oder aber auf die Emotionen der Versuchspersonen bezog. Es wurden neben einem globalen Emotionsmaß differenzierte Angaben zu den vier Emotionsdimensionen gemäß Rolls (1999) erhoben. Darüber hinaus wurden noch weitere Fragen z.B. zur Person (Alter, Geschlecht) und zur Situation (wie interessant war die Geschichte, wie sympathisch die Erzählerin) gestellt.

Um die Ergebnisse nicht durch implizite Theorien zu verfälschen, wurden die Versuchspersonen nur jeweils entweder zu den Emotionen des Senders (d.h. der Erzählerin der Geschichten) oder aber zu ihren eigenen Emotionen während der Präsentation der Geschichte befragt. Letztere wurden darüber hinaus zur Glaubwürdigkeit der Geschichte befragt.

## Ergebnisse

Die Emotionalität der Sender zeigt zwischen wahren und falschen Geschichten im globalen Emotionalitätsmaß keine Unterschiede. Auf dem differenzierteren Niveau der vier Emotionsdimensionen zeigt sich ein Unterschied zwischen wahren und falschen Geschichten in der Einschätzung gezeigter Angst.

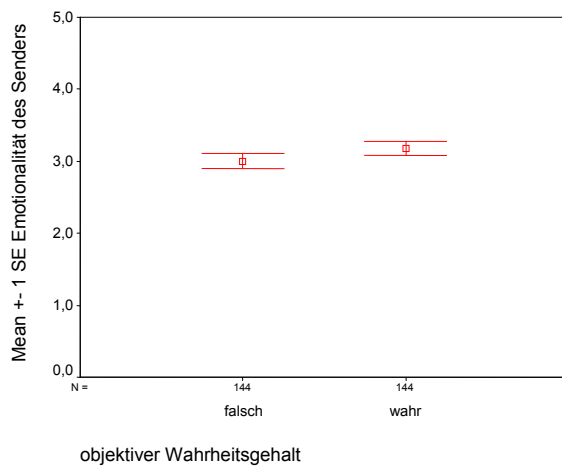


Abbildung 3: globale Emotionalität der Sender bei wahren und falschen Geschichten (n.s.)

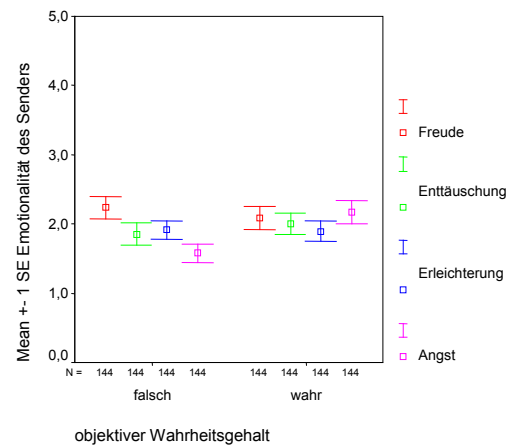


Abbildung 4: Emotionalität der Sender bei wahren und falschen Geschichten in den vier Emotionalitätsdimensionen (nur bei „Angst“  $t < -2.8$ ;  $p < 0.01$ )

Sender und Empfänger weisen bei den verschiedenen Geschichten ein jeweils stark übereinstimmendes Bild der Emotionen auf. Dies gilt sowohl für das globale Emotionalitätsmaß als auch für die differenziertere Analyse der verschiedenen Emotionsdimensionen.

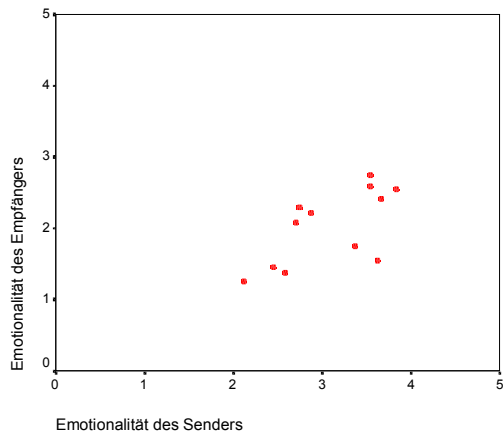


Abbildung 5: Emotionalitäten der Sender und Empfänger (global)  
 $(r = 0.65; p < 0.05)$

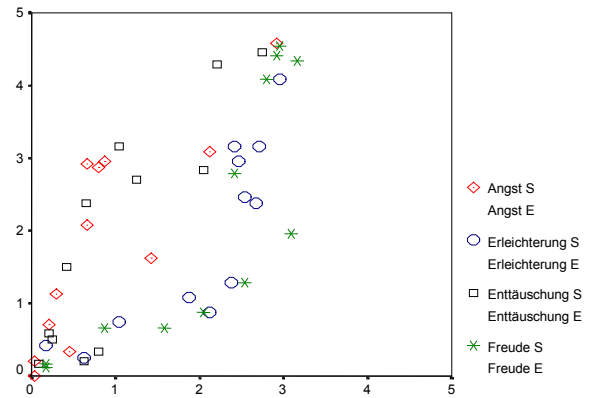


Abbildung 6: Emotionalitäten der Sender und Empfänger in den vier Dimensionen  
 $(\chi^2 = 48,9; p < 0.01)$

Es konnte mittels t-Test für abhängige Stichproben nachgewiesen werden, daß die Empfänger auf dem globalem Niveau der Emotionalität signifikant höhere Werte bei für wahr gehaltenen Aussagen erreichen.

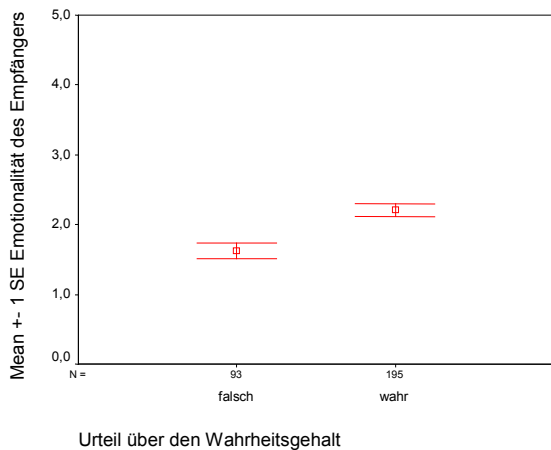


Abbildung 7: Urteil über den Wahrheitsgehalt und Emotionalität des Empfängers (global)  
 $(t = -4.1; p < 0.01)$

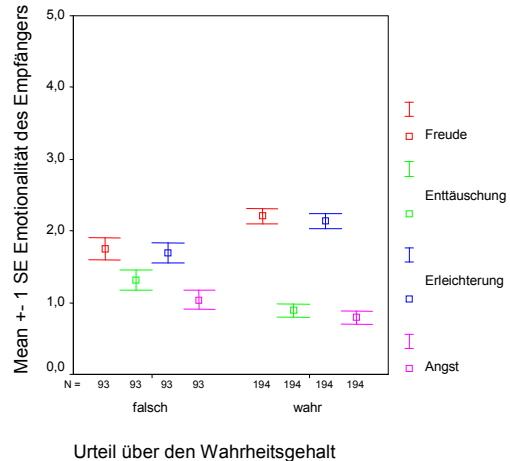


Abbildung 8: Urteil über den Wahrheitsgehalt und Emotionalität des Empfängers in den vier Dimensionen (bis auf „Angst“ alle  $t > 2.4; p < 0.05$ )

In der Stichprobe der kanadischen Studenten zeigte sich darüber hinaus besonders deutlich, daß den als emotionaler eingeschätzten Geschichten mehr Glauben geschenkt wird.

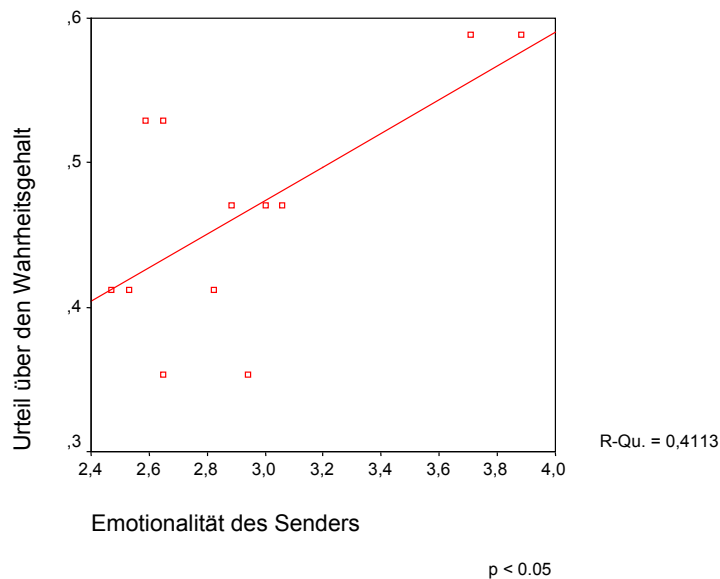


Abbildung 9: Regression des Urteils über den Wahrheitsgehalt auf die Emotionalität des Senders (global)

## Diskussion

Die Ergebnisse belegen einen deutlichen Einfluß von Emotionen bei der Übermittlung und Bewertung wahrer und falscher Aussagen.

Wahre Aussagen werden vom Sender – entgegen der Erwartung – mit höheren statt niedrigeren Emotionen (als falsche Aussagen) begleitet. Diese setzen sich dann – und dies erwartungsgemäß – beim Empfänger fort und führen bei diesem – ebenfalls entsprechend der Erwartung – zur Tendenz, die Aussage für wahr zu halten. Der letzte Effekt trat besonders stark bei den kanadischen Studenten auf, welche sich mangels inhaltlichen Verständnisses vermutlich vorrangig auf das Ausdrucksverhalten des Senders bezogen.

Für den der Annahme widersprechenden Befund, daß die Sender bei falschen Aussagen keine höhere Emotionalität zeigten, lassen sich mehrere Erklärungen finden:

1. Zum einen könnte es sein, daß sich die Sender (möglicherweise aufgrund zu geringer Motivation) weniger strategisch verhalten als angenommen wurde, und die Emotionen *nicht* zur Tarnung der Lügen einsetzten, obwohl das Verhalten der Empfänger zeigt, daß eine solche Strategie Erfolg verspräche.
2. Ferner wäre es denkbar, daß – entsprechend dem Content Complexity Approach – die Sender während der gelogenen Geschichten zu beschäftigt waren, um intensivere Emotionen zu zeigen.
3. Eine weitere Erklärung könnte darin liegen, daß die Sender bei den Geschichten, mit denen sie reale Erinnerungen verbinden, emotional involvierter waren und dies von den Empfängern bemerkt wurde. In diesem Fall wäre es denkbar, daß die Sender Emotionen strategisch einsetzten, dies aber nicht deutlich wurde, da die Emotionen bei den wahren Geschichten noch größer waren.

Welche Erklärung zutrifft, läßt sich derzeit noch nicht abschätzen, dennoch lassen die dafür durchzuführenden Untersuchungen in jedem Fall interessante Implikationen für die Glaubwürdigkeitsbeurteilung erwarten.

## Literatur

- Andersen, P. A. & Guerrero, L. K. (1998). *Handbook of Communication and Emotion*. San Diego: Academic Press.
- Ekman, P. & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist*, 46 (9), 913-920.
- Ekman, P. (1992). *Telling Lies (2 ed.)*. New York: Norton.
- Ellis, H. C. & Ashbrook, P. W. (1989). The "state" of mood and memory research: a selective review. *Journal of Social Behavior and Personality*, 4, 1-21.
- Köhnken, G. (1990). *Glaubwürdigkeit: Untersuchungen zu einem psychologischen Konstrukt. (Fortschritte der psychologischen Forschung 5 ed.)*. München: Psychologie-Verlags-Union.
- Rolls, E. T. (1990). A theory of emotion, and its application to understanding the neural basis of emotion. *Cognition and Emotion*, 4 (3), 161-190.
- Rolls, E. T. (1999). *The Brain and Emotion*. New York: Oxford University Press.
- Steller, M. (1987). *Psychophysiologische Aussagebeurteilung: Wissenschaftliche Grundlagen und Anwendungsmöglichkeiten der "Lügendetektion"*. Göttingen: Hogrefe.
- Steller, M. & Köhnken, G. (1989). Criteria-based statement analysis. In D. C. Raskin (Ed.), *Psychological Methods in Criminal Investigation and Evidence* (pp. 217-245). New York: Springer.
- Vrij, A. (2000). *Detecting Lies and Deceit*. Chichester: John Wiley & Sons Ltd.



# Multimethodale Diagnostik bei forensisch-psychologischen Begutachtungen

*Ernst Plaum*

*Katholische Universität Eichstätt*

Ungeachtet einer praktischen Realisierung wird in der Klinischen Psychologie seit langem eine multimethodale (und „multimodale“) Einzelfalldiagnostik gefordert (hierzu Plaum, 1992, 1993). Bei forensischen Begutachtungen überwiegen aber offenbar Vorgehensweisen, die den Schwerpunkt auf ganz bestimmte methodische Zugänge legen, im Extremfall sich auf nur einen Ansatz – das psychodiagnostische Interview – beschränken. Gebräuchlich sind auch Beurteilungsinstrumente (wie etwa SVR 20, HCR 20, PFI oder die Hare Psychopathy Checklist) bzw. Kriterienlisten, die Prognosen ermöglichen sollen. Dieser Beitrag befaßt sich nur mit einer Täterdiagnostik; die Untersuchung von angeblichen oder tatsächlichen Opfern wird hier ausgeklammert (hierzu beispielsweise Plaum, 1998).

Die herangezogenen Verfahren beruhen in weitaus den meisten Fällen auf Selbstaussagen der Probanden, allenfalls noch Verhaltensbeobachtungen. Die hinreichend bekannten Verfälschungstendenzen bzw. Beurteilungsfehler bei diesen Methoden bedürfen an dieser Stelle keiner weiteren Erläuterung. Sogenannte Offenheitsskalen sind grundsätzlich von zweifelhaftem Wert, umso mehr als ein Straftäter nicht allzu viel Cleverneß aufbringen muß, um zu erkennen, daß es höchst ungeschickt wäre, zu allem Überfluß auch noch Items mit der Beschreibung derart läppischer „Verfehlungen“ verfälschend zu beantworten. Vielfach hofft man jedoch, durch trickreiche Befragungstechniken und/oder auf Grund einer Selbsteinschätzung als qualifizierter (erfahrener) Beurteiler, einen validen Eindruck von der Ehrlichkeit der Probanden (Pbn) zu gewinnen. Da objektive Validierungskriterien aber meist nicht zur Verfügung stehen, beruhen solche subjektiven Überzeugungen bezüglich der eigenen Fähigkeiten auf fragwürdigen Voraussetzungen. Es mag wohl in manchen Fällen – etwa bei vorliegender Minderbegabung oder unverblühten Äußerungen der Delinquenten (beispielsweise: „Wenn ich rauskomme, möchte ich Bandenking werden“) - relativ leicht sein, zu realitätsentsprechenden Urteilen zu kommen, man trifft aber zweifellos Strafgefangene, die recht gut zu täuschen vermögen. Gutachter sollten sich ruhig eingestehen, daß sie in so mancher Hinsicht ihren Pbn überlegen sein mögen, aber nicht unbedingt hinsichtlich der Fähigkeit, unehrliche Menschen zu „durchschauen“. Die Vorstellung, gute Psychologen müßten auf Anhieb in der Lage sein, eine fremde Person richtig einzuschätzen, sollten wir uninformatierten Laien überlassen – bekanntlich trifft eher das Gegenteil zu.

Nicht zuletzt deshalb wäre es unverantwortlich, Untersuchungen mit dem Ziel einer Begutachtung innerhalb eines sehr knapp bemessenen Zeitraumes vorzunehmen. Manche Auftraggeber können durchaus akzeptieren, daß eine hinreichend gründliche Diagnostik etwa zwei volle Tage in Anspruch nehmen muß. Schon allein dieser Zeitaufwand bietet eine gewisse Gewähr dafür, daß die Pbn systematische Verfälschungen nicht konsistent durchzuhalten vermögen. So entglitt selbst einem äußerst geschickt vorgehenden, recht intelligenten Studenten am Ende des zweiten Untersuchungstages die Verhaltenssteuerung während einer ausführlichen Exploration: Der Verfasser hatte ihn gefragt, mit welchem Tier er sich selbst am ehesten vergleichen würde und die Antwort lautete, daß dies „so ein kleiner bunter Kolibri“ wäre. An dieser Stelle griff der anwesende protokollierende junge Kollege ein und sagte ganz unvermittelt: „Aber das ist doch sicher Ihre Idealvorstellung – sagen Sie

einmal, wie es real ist!“ Zu unserer Verblüffung kam nun ein ganz anderer Vergleich: „Ja, da würde ich sagen – Wolf im Schafspelz“.

Es reicht nun gewiß nicht aus, einfach lange genug zu untersuchen und dabei zu hoffen, daß die Pbn irgendwann einmal zwischendurch „ihre Maske fallen lassen“. Vielmehr schließt eine multimethodale diagnostische Strategie ganz gezielt recht unterschiedliche Verfahren ein. Von besonderem Interesse ist bei Begutachtungen von Straftätern zumeist die Gefährlichkeitsprognose, auf der Basis einer Aggressionsdiagnostik. Die folgenden Darlegungen sollen hierauf beschränkt bleiben (vgl. hierzu etwa Rehder, 1990). Es erscheint keineswegs sinnlos, mit Persönlichkeitsfragebögen, auch solchen, die spezifische Aggressionsskalen enthalten, zu arbeiten. Natürlich werden Strafgefangene hierbei – nicht immer, aber in der Regel – zu verfälschen versuchen. Erzielt ein Proband (Pb) nun einen so niedrigen Aggressionswert, wie er auch in einer Normalpopulation sehr selten vorkommt, dann darf man dies, zunächst hypothetisch, als Hinweis auf Reaktionen in Richtung sozialer Erwünschtheit werten. Wie aber lassen sich nun positive Anhaltspunkte für dennoch vorhandene Aggressionstendenzen finden? Hier wäre an weniger durchschaubare Verfahren zu denken, in erster Linie die sogenannten projektiven Techniken (siehe Schaipp & Plaum, 1995; Schaipp, 2001a u. b).

Bevor man nun diesbezüglich das geläufige „Mainstream-Protestgeschrei“ anstimmt oder sich besserwisserisch abwendet, sollten empirische Resultate zur Kenntnis genommen werden, die für die Anwendung dieser umstrittenen diagnostischen Verfahren sprechen. So konnten Plaum & Biedermann (1999) Schülerinnen und Schüler zweier Hauptschulklassen mit zwei einschlägigen Fragebogentests (PFK9-14 und der Hostilitätsfragebogen) sowie zwei „projektiven“ Techniken (Foto-Hand-Test, Wunschprobe nach Wilde – zu diesen Methoden siehe Plaum & Biedermann, 1999) untersuchen und feststellen, daß vor allem bei den männlichen Pbn Resultate der beiden letztgenannten Verfahren mit einem einfachen Außenkriterium (Einschätzung als „aggressiv“ vs. „nicht aggressiv“ durch einen Pädagogen) signifikant korrelierten, während Skalenwerte der erstgenannten Instrumente solche Zusammenhänge vermissen ließen (siehe Tabelle 1, biseriale Rangkorrelationen). Daß die Ergebnisse dieser Untersuchung keineswegs nur singuläre Zufallsresultate darstellen, haben im übrigen Arbeiten mit Gymnasiasten (N = 39) und Jugendlichen in Heimen (N = 13) gezeigt (Plaum, Chau & Müller, 1999).

Tabelle 1: Signifikant mit dem Außenkriterium (aggressiv vs. nicht aggressiv) korrelierende Variablen der „projektiven“ Techniken (\*\* = signifikant auf dem 1%-Niveau) bei 11 männlichen Hauptschülern

---

Foto-Hand-Test, „aggressive“ Antworten:	r =	.76**
Foto-Hand-Test, „direktive“ Antworten:	r =	-.73**
Foto-Hand-Test, Abhängigkeit bezeichnende Antworten:	r =	.63*
Wunschprobe, „destruktive“ Thematik (bei Distanzierung von derselben):	r =	.63*

---

Sieht man einmal von dem interpretationsbedürftigen negativen Vorzeichen eines Koeffizienten (Tab. 1, zweiter Wert von oben) ab, so ist der in Tabelle 1 wiedergegebene Zusammenhang bezüglich der Wunschprobe höchst interessant, und zwar im Hinblick auf Verfälschungstendenzen: Dieses Verfahren verlangt von den Pbn eine Identifikation mit (in der hier verwendeten Version 30) vorgegebenen nichtmenschlichen Existenzformen, wobei jeweils angegeben werden soll, ob man eine solche für sich selbst ablehnen oder befürworten würde, und zwar mit einer entsprechenden Begründung. Entscheidend ist dabei nun das Anführen einer Destruktionsthematik, wobei diese meistens eine Ablehnung erfährt (z.B.



„Bombe möchte ich nicht sein, weil die zerstört“). Je häufiger nun eine solche Thematik verbalisiert wurde – trotz der Kundgabe einer ablehnenden Haltung! -, desto eher waren die Schüler der Kategorie „aggressiv“ zugeordnet worden. Dies läßt vermuten, daß das bloße Auftreten „destruktiver“ Assoziationen bereits auf Aggressionstendenzen hinweist, auch wenn diesbezüglich sozial erwünscht reagiert wird. Nebenbei bemerkt fanden wir bezüglich einer in der Aggressionsdiagnostik häufig verwendeten „projektiven“ Technik, des Rosenzweig P-F Tests, keine bedeutsamen Resultate (Plaum, Chau & Müller, 1999).

Unsere übrigen Ergebnisse zur Brauchbarkeit „projektiver“ Verfahren ließen sich nun auch, neben anderen interessanten Befunden, bei einer Gruppe von Strafgefangenen (N = 30) im wesentlichen bestätigen. Es handelt sich dabei um eine Arbeit von Thiel (1999; Plaum & Thiel, 2001) mit einer Standardform (51 Items) der eben erwähnten Wunschprobe. Außerdem kamen drei Fragebogenverfahren – Freiburger Persönlichkeitsinventar (FPI-R), Fragebogen zur Erfassung von Aggressivitätsfaktoren (FAF) und Trierer Persönlichkeitsfragebogen (TPF) – zur Anwendung. Einige wenige ausgewählte Korrelationen (Kendalls Tau) sollen hier, in Tabelle 2, wiedergegeben werden. Zunächst fällt auf, daß beide Offenheitsskalen sehr signifikant mit der Fragebogen-Aggressivität zusammenhängen, trotz der Fragwürdigkeit der ersteren, d.h. was hier als Verfälschungsabsicht interpretiert werden könnte, steht in enger (negativer) Beziehung zu Selbstaussagen bezüglich der Aggressionstendenzen.

Tabelle 2: Signifikante Korrelationen (\*\* = 1%-Signifikanzniveau) bei 30 Straftätern. Of = Offenheitsskala; Vor = Anzahl der Vorstrafen; Str = jetziges Strafmaß; Ag = Skala „Aggressivität“; Sp = Skala „spontane Aggressivität“; Ve = Verhaltenskontrolle (TPF), Dt = „destruktive“ Thematik insgesamt (Wunschprobe), Dt- = „destruktive“ Thematik bei Distanzierung, Dt+ = „destruktive“ Thematik bei Befürwortung („möchte ich sein ...“).

	Ag (FPI-R)	Sp (FAF)	Ve	Dt	Dt-	Dt+
Of (FAF)	.41**	.40**	-	-.35*	-	-
Of (FPI)	.54**	.49**	-.31*	-.27*	-	-
Vor	.47**	.24*	-.33**	-.34*	-.38**	.26*
Str	-.26*	-.26*	-	-	.25*	-

Dies ist ein Hinweis auf recht massive Neigungen, sich sozial erwünscht zu präsentieren. Weiterhin bestehen Zusammenhänge mit der Anzahl der Vorstrafen (besonders ausgeprägt beim FPI-R), d.h. je größer diese Anzahl ist, desto eher werden Aggressionstendenzen zugegeben; man kann somit vermuten, daß für die Pbn *sehr* deutliche Verfälschungen weniger opportun erscheinen, je stärker die Biographie bereits aktenkundig von Delinquenz geprägt ist. Dem entsprechen auch die Korrelationen zwischen den Offenheitsskalen und den Vorstrafen (.33\*\* bzw. .29\*). Was das Strafmaß bei der jetzigen Inhaftierung betrifft, so zeigt sich jedoch eine gegenläufige Tendenz: Es waren hier Koeffizienten mit negativen Vorzeichen bezüglich der Offenheitsskalen zu finden (-.30\* bzw. -.24\*) und auch zwischen Strafmaß und Aggressivitätsskalen besteht eine noch signifikant negative Korrelation (Tab. 2), d.h. je niedrigere Werte die letzteren ergeben, desto höher ist das Strafmaß.

Sozusagen spiegelbildlich zu einigen der eben dargestellten Zusammenhänge waren die Beziehungen der Skala „Verhaltenskontrolle“ des TPF (Tab. 2): je kontrollierter sich die Pbn selbst schildern, desto geringer stellt sich die „Offenheit“ (des FPI-R) dar und desto geringer ist die Anzahl der Vorstrafen. Das heißt also, auch der Trierer Persönlichkeitsfragebogen vermag anscheinend mit der Skala „Verhaltenskontrolle“ zur Diagnostik möglicher Verfälschungstendenzen beizutragen. Es ist nun, im Hinblick auf eine multimethodale

Strategie unter Einbeziehung nichtpsychometrischer Verfahren bemerkenswert, daß die Korrelationen der Destruktionsthematik in der Wunschprobe (Dt) in die gleiche Richtung weisen wie die Werte der zuletzt genannten Fragebogenskala. Zudem zeigt sich noch in Tabelle 2 eine Differenzierung zwischen einer abgelehnten (Dt-) und einer befürworteten (Dt+) „Destruktivität“; erstere korreliert negativ mit der Anzahl der Vorstrafen, aber noch signifikant positiv mit dem Strafmaß (Letzteres bedeutet, daß eine Ablehnung von Destruktivität eher mit einem höheren Strafmaß einhergeht). Dt+ hingegen hängt ebenso eng mit der Anzahl der Vorstrafen zusammen wie die „spontane Aggressivität“ des FAF (Tab. 2). In diesem Zusammenhang ist noch bemerkenswert, daß die Skala „Verhaltenskontrolle“ positiv mit Dt- korreliert (.49\*\*), mit Dt+ jedoch negativ (-.26\*); die letztgenannte Beziehung bedeutet, daß bei einer Bejahung von Destruktivität (wie z.B. „Orkan möchte ich sein, weil der zerstört“) eher mit einer geringeren Verhaltenskontrolle zu rechnen ist. Dies mag trivial erscheinen; bei der Suche nach Mehrfachbelegen im Sinne des Konvergenz-Divergenz-Prinzips (siehe unten, hierzu Plaum, 1996) kann aber die Vorgabe der „Wunschprobe“ somit sehr sinnvoll sein.

Auf interessante (signifikante) Tendenzen bei der Betrachtung von Subgruppen hinsichtlich der Straftaten wird nicht eingegangen, zumal sie sich nur auf sehr kleine Pbn-Gruppen beziehen. Natürlich besagen die hier präsentierten Korrelationen nicht allzu viel bezüglich des jeweiligen Individuums. Daher dürfte die Skizzierung einiger Gutachtenfälle nicht uninteressant sein, bei denen sich eine deutliche Diskrepanz zwischen den Ergebnissen der Fragebogentests und denen der „projektiven“ Techniken zeigte. Außer den erwähnten Verfahren wurden noch zahlreiche andere vorgegeben. Bei diesen Pbn boten mehrere Fragebogentests *keine* Anhaltspunkte für Aggressionstendenzen:

#### Fall 1

Fragebogenverfahren: FPI-R, FAF, MMPI u.a. Foto-Hand-Test: Bei 11 der 34 Items manifeste Aggressivität (Bedrohen, Zuschlagen, Gewalt); Wunschprobe (51 Items): Bei 7 Items Ablehnung von Destruktivität; Tiger am beliebtesten.

#### Fall 2

Fragebogenverfahren: FPI-R, FAF, Gießentest u.a. Foto-Hand-Test: 3 zweifelsfreie Aggressionsdeutungen, 7 weniger eindeutige (jemanden greifen), bei 13 Items „Selbstbefriedigung“ bzw. sexuelle Handlung. Wunschprobe: Positive Wahlen: Gesetz („ich kann bestimmen“), Berg (gewaltig und groß), Orkan (mächtig und gewaltig), Rakete (mächtig, kann zerstören), Blitz (Töten gehaßter Menschen), Vulkan und Tiger „faszinierend“.

#### Fall 3

Fragebogenverfahren: FPI-R, FAF, MMPI. Foto-Hand-Test: wenig auffällig; Wunschprobe: keine Hinweise auf Aggressivität. TAT (18 Bilder): Bei insgesamt 6 Tafeln „Streit“, „Schläge“, „Erwürgen“, der Frau etwas antun, Fluchtversuch (2 mal).

#### Fall 4

Fragebogenverfahren: FAF u.a. Foto-Hand-Test: 10 Antworten mit aggressiver Tendenz (Schlagen, „Schmeißen“), zwei Mal Krankheit bzw. Tod. Wunschprobe (51 Items): Bei 16 (abgelehnten) Existenzformen Begründung mit Destruktivität als Thema. Hinweis auf soziale Erwünschtheit: „Kirche“ als beliebtestes Item bezeichnet (extrem unwahrscheinliche Präferenz). Bemerkenswerte Reaktion auf „Gesetz“: Lachen, dann: „Zu viele Leute ... z.B. so wie ich, dieses Gesetz muß mir gefallen, diese Gesetze nicht. Für andere Leute nicht gut“; ähnlich bei dem Item „Ordnungssystem“.

Häufig lassen schon einzelne Antworten bestimmte Vermutungen zu, die – entsprechend dem Konvergenz-Divergenz-Prinzip (siehe unten) – selbstverständlich weiterer Bestätigungen bedürfen, so z.B. folgende Begründungen zu einzelnen Items (+ = Befürwortung, - = Ablehnung) bei realen Gutachtenfällen.

- Berg (+): weil ich dort unbesiegbar bin und weil ich der Größte bin.
- Blitz (+): weil ich dann mehr Macht habe, ich könnte dem eins auswischen, der mich verletzt oder ärgert.
- Orkan (+): weil wenn ich mal durchdrehe, kann sich niemand mir in den Weg stellen.
- Piratenschiff (+): weil ich dann immer andere Schiffe überfallen und ausbeuten kann.
- Tiger (+): weil ich dann zeigen kann, wer hier das zu sagen hat.
- Gesetz (-): weil ich mich auch nicht nach dem Gesetz halte.
- Teufelsmaske (-): weil man damit vielleicht eine Bank überfallen könnte.

Zum Abschluß sollen noch einige Bemerkungen zum Konvergenz-Divergenz-Prinzip folgen, da dies häufig mißverstanden wird. Das Konvergenzprinzip nach Bierkens (1968) besagt, daß zahlreiche, für sich genommen nicht hinreichend aussagekräftige diagnostische Hinweise, die alle in die gleiche Richtung weisen (z.B. Aggressionstendenzen betreffend), *insgesamt* betrachtet doch ein relativ hohes Maß an Urteilssicherheit zu begründen vermögen. Nach einem solchen Konvergenzprinzip wird auch in anderen angewandten Wissenschaften gearbeitet, wie etwa der Medizin und der Jurisprudenz. Wenn ein sorgfältig arbeitender Diagnostiker seine Tätigkeit dementsprechend ausrichtet, mag ihm unterstellt werden, er suche lediglich nach Bestätigungen für seine Hypothese(n) und klammere alles Entgegenstehende aus. Es sei dahingestellt, inwieweit hierin eine Projektion der Arbeitsweise dieser Kritiker zu sehen ist; jedenfalls beinhaltet das Konvergenzprinzip ein solches einseitiges Vorgehen gerade nicht. Um derartige Mißverständnisse zu vermeiden, spricht der Verfasser dieses Beitrages lieber vom Konvergenz-*Divergenz*-Prinzip, da es das Bestreben der Vertreter desselben sein sollte, nicht nur nach Konvergenzen zu suchen, sondern ebenso denselben entgegenstehende Befunde – Divergenzen – aufzufinden (Plaum, 1996; siehe auch Biedermann & Plaum, 1999). Treten solche auf, so sind gegebenenfalls Metaperspektiven einzunehmen, auf Grund derer widersprüchlich scheinende diagnostische Resultate vielleicht erklärt werden können. Um bei einer multimethodalen Diagnostik – beispielhaft – zu bleiben: Diskrepanzen zwischen Fragebogenresultaten und Ergebnissen „projektiver“ Techniken ließen sich eventuell als Folge der Tendenz zur „sozialen Erwünschtheit“ bei den erstgenannten Verfahren auffassen, wobei eine Betrachtung unter diesem Aspekt als Metaperspektive anzusehen wäre. Es erübrigt sich, hinzuzufügen, daß eine *einzig*e „projektive“ Technik – nicht zuletzt wegen der Mehrdeutigkeit zahlreicher Reaktionen – grundsätzlich nicht mehr als Hypothesen zu liefern vermag (Plaum, 1992, 1996).

## Literatur

- Biedermann, Th. & Plaum, E. (1999). *Aggressive Jugendliche. Fakten, Theorien, Hintergründe und methodische Zugangsweisen*. Wiesbaden: Deutscher Universitäts-Verlag.
- Bierkens, P. B. (1968). *Die Urteilsbildung in der Psychodiagnostik*. München: Barth.
- Plaum, E. (1992). *Psychologische Einzelfallarbeit*. Stuttgart: Lucius & Lucius.
- Plaum, E. (1993). Einzelfalldiagnostik in der Klinischen Psychologie. In F. Baumgärtel & F.-W. Wilker (Hrsg.), *Klinische Psychologie im Spiegel ihrer Praxis* (S. 13-27). Bonn: Deutscher Psychologen-Verlag.

- Plaum, E. (1996). *Einführung in die Psychodiagnostik*. Darmstadt: Wissenschaftl. Buchgesellsch./Primus.
- Plaum, E. (1998). Ein „Wahrheitstest“ für Glaubhaftigkeitsgutachten? Zur Psychodiagnostik bei Opfern von Sexualstraftaten. *Kriminalistik*, 52 (8-9), 549-557.
- Plaum, E. & Biedermann, Th. (1999). Die Erkennbarkeit destruktiver Verhaltensdispositionen als Voraussetzung von Interventionen zur Gewaltprävention. *Forensische Psychiatrie und Psychotherapie. Werkstattsschriften*, 6, 2, 23-38.
- Plaum, E., Chau, H.-V. & Müller, St. (1999). „Projektive“ Techniken als Hilfsmittel zur Erfassung sozial unerwünschter Verhaltensdispositionen? *Zeitschrift für Politische Psychologie*, 7, SH/99; 175-194.
- Plaum, E. & Thiel, M. (2001). Zum Problem der Verfälschung bei Äußerungen von Straftätern zur eigenen Person: Körperverletzung, Vergewaltigung und sexueller Mißbrauch an Kindern im Spiegel psychodiagnostischer Untersuchungsergebnisse. Unveröff. Manuskript.
- Rehder, U. (1990). *Aggressive Sexualdelinquenten*. Kriminalpädagogische Praxis, Bd. 10. Lingen/Ems: Kriminalpädagogischer Verlag.
- Schaipp, Ch. (2001a). Neuere Entwicklungen bei den „projektiven“ Techniken. *Report Psychologie*, 26/2, 94-102.
- Schaipp, Ch. (2001b). *Validität und diagnostische Brauchbarkeit ausgewählter indirekter und direkter Befragungsmethoden zur Diagnostik von Aggressivität, Neurotizismus bzw. psychischer Stabilität*. München: Herbert Utz.
- Schaipp, Ch. & Plaum, E. (1995). „Projektive Techniken“: Unseriöse „Tests“ oder wertvolle qualitative Methoden? Bonn: Deutscher Psychologen Verlag.
- Thiel, M. (1999). *Aggressivität und deren Zusammenhang mit weiteren Persönlichkeitsmerkmalen sowie Verfälschungstendenzen bei Strafgefangenen*. Unveröff. Diplomarbeit. Philosophisch-Pädagogische Fakultät der Katholischen Universität Eichstätt, Fach Psychologie.

# Kriminalität - Ursachenzuschreibung und Strafhärte: Eine Untersuchung aus der Sicht juristischer Laien

*Anne Reichert & Wolfgang Bilsky*

Ursachenzuschreibung und Strafzumessungspräferenzen juristischer Laien wurden im Rahmen einer Interviewstudie deliktspezifisch untersucht. Ausgehend von einer Sekundäranalyse der von Görden (1996) mit einem umfangreichen Fragebogen erhobenen kriminalitätsbezogenen Ursachenzuschreibungen wurde zunächst ein Instrument entwickelt, das es gestattet, für Diebstahls- und Körperverletzungsdelikte zwischen person-, interaktions- und systembezogenen Attributionen zu differenzieren (Bilsky, Oswald & Marciszewski, 1989). In der anschließenden Hauptuntersuchung mit insgesamt 211 Befragten unterschiedlichen Alters und unterschiedlicher Berufszugehörigkeit wurden im ersten Teil auf der Grundlage vorgegebener Sanktionsalternativen individuelle Strafzumessungspräferenzen erhoben, die sich auf kurze Beschreibungen eines Diebstahls- und eines Körperverletzungsdeliktes bezogen; zwei Tätermerkmale, Vorstrafe und sozioökonomischer Status, wurden hierbei systematisch variiert. Im zweiten Teil erfolgte die Erhebung kriminalitätsbezogener Ursachenzuschreibungen mit Hilfe des zuvor entwickelten Instrumentes; jeweils die Hälfte der Befragten äußerte sich hierbei zur Entstehung von Diebstahls- und Körperverletzungsdelikten. Im abschließenden dritten Teil bestand die Aufgabe der Befragten darin, die im ersten Teil vorgegebenen Sanktionsalternativen nach ihrer subjektiv wahrgenommenen Strafhärte zu ordnen. Hinsichtlich der Deliktspezifität von Ursachenzuschreibungen ergaben sich für Körperverletzungsdelikte höhere personbezogene Attributionen als für Diebstahlsdelikte. Darüber hinaus konnte gezeigt werden, daß ein deliktspezifischer Zusammenhang zwischen den Ursachenzuschreibungen einer Person und der von ihr je präferierten Strafhärte existiert. Schließlich bestätigte die kovarianzanalytische Datenauswertung sowohl für Diebstahl als auch für Körperverletzung den Einfluß der untersuchten Tätervariablen auf die individuell präferierte Strafhärte.

## **1 Alltagstheorien, Deliktspezifität und Strafhärte**

### *Alltagstheorien*

Für die Orientierung im täglichen Leben sind sogenannte Alltags- oder Laintheorien von zentraler Bedeutung. Sie beziehen sich auf verschiedenste Bereiche, so auch auf die Entstehung und die Persistenz normabweichenden, kriminellen Verhaltens sowie die Angemessenheit unterschiedlicher Interventions- und Reaktionsformen. Juristische Laien verfügen über teilweise sehr elaborierte Vorstellungen darüber, wie Kriminalität entsteht. Furnham (1988) führt als Grund für die intensive Beschäftigung von Laien mit dem Thema Kriminalität die persönliche Betroffenheit der Menschen an. Subjektive Alltagstheorien über Kriminalität findet man jedoch nicht nur bei juristischen Laien. Vielmehr verfügen auch Juristen, neben ihrem Expertenwissen auf dem Gebiet der Rechtswissenschaften, über implizite, ihnen selbst zum Teil nicht bewußte Theorien bezüglich der Entstehung von Kriminalität (Oswald & Bilsky, 1991).

Zahlreiche Studien haben sich mit den von der Allgemeinbevölkerung subjektiv wahrgenommenen Gründen für die Entstehung von Kriminalität beschäftigt (z.B. Furnham & Henderson, 1983; Erskine, 1974). Die Ergebnisse weisen darauf hin, daß in Alltagstheorien in der Regel multiple Ursachen für die Entstehung von Kriminalität

angenommen werden. Die Ursachenzuschreibungen verschiedener Personen unterscheiden sich dabei insofern, als den einzelnen Faktoren unterschiedlich viel Bedeutung beigemessen wird. Kidder und Cohn (1979) haben die Ergebnisse einiger Studien zusammengefaßt und festgestellt, daß sich ein hohes Maß an Übereinstimmung in den wahrgenommenen Kriminalitätsursachen zeigt. Zu den wichtigsten Ursachen gehören nach dieser Studie unter anderem Drogen- und Alkoholmißbrauch, Armut und Arbeitslosigkeit, Mangel an Aktivität und elterlicher Lenkung bei Kindern, das soziale Umfeld und die unzureichende Durchsetzung von Gesetzen. Furnham und Henderson (1983) fanden in einer Untersuchung bei juristischen Laien in Großbritannien faktorenanalytisch sechs Verursachungsbereiche: „gestörte Erziehung“, „mentale Instabilität“, „Versuchung“, „Nervenkitzel“, „Entfremdung“ und „Eltern“. Ferner konnten sie einen Einfluß von Geschlecht, Alter und politischer Einstellung der Befragten auf die Ursachenzuschreibung finden. Es scheint also Unterschiede in der kriminalitätsbezogenen Ursachenzuschreibung zu geben, welche sich zumindest teilweise durch Personenmerkmale der urteilenden Person erklären lassen. Furnham (1988) stellt hierzu fest: „... certain demographic and psychological features of people appear to be systematically related to the explanation or theory they tend to support“ (p. 188).

Bei der Untersuchung kriminalitätsbezogener Attributionen ist man allerdings häufig weniger an den vom Individuum genannten einzelnen Ursachen für Kriminalität interessiert, sondern sucht statt dessen nach allgemeinen Dimensionen, die diesen individuellen Attributionen zugrunde liegen (vgl. Hupfeld, 1996). Oswald und Bilsky (1991) überprüften beispielsweise, ob sich die Differenzierungen Weiners (1986) nach Lokation (internal/external), Stabilität (stabil/instabil), Kontrolle (kontrollierbar/unkontrollierbar) und Intentionalität (absichtlich/unabsichtlich) auf den Bereich kriminalitätsbezogener Attributionen übertragen lassen. Sie legten Strafrichtern und Jurastudenten eine Liste möglicher Ursachen für Kriminalität mit der Instruktion vor, einzuschätzen, welche Bedeutsamkeit jeder einzelnen Ursache bei der Entstehung von Kriminalität zukommt. Um die Dimensionalität der kriminalitätsbezogenen Attributionen zu überprüfen, wurde eine Faktorenanalyse durchgeführt. Die gewählte Vier-Faktoren-Lösung gestattete eine deutliche Trennung der Items gemäß der Lokationsdimension: Auf den beiden als „destruktive Motivation“ und „Plan- und Perspektivlosigkeit“ bezeichneten Faktoren luden solche Items, die a priori als „internal“ eingestuft worden waren. Die externalen Items ließen sich weiter unterteilen in einen distalen Faktor „soziales System“ und in einen proximalen Faktor „soziale Randständigkeit“, der den Einfluß des näheren sozialen Umfeldes des Täters betrifft. Die übrigen Dimensionen konnten von Oswald und Bilsky (1991) demgegenüber nicht identifiziert werden. Bilsky, Oswald und Marciszewski (1989) werteten denselben Datensatz zudem mit Hilfe der Multidimensionalen Skalierung (MDS) aus. Hierbei gelang es, die Lokationsdimension (internal/external) insofern weiter auszudifferenzieren, als eine Trennung der Ursachenfaktoren nach Maßgabe der Involviertheit des Täters möglich war. Analog einer Trennung in Mikro-, Meso- und Makroebene konnten Person-, Interaktions- und System-Aspekte unterschieden werden. Görge (1996) führte eine Untersuchung dazu durch, welche Ursachen- und Motivzuschreibungen Jurastudenten in Bezug auf Kriminalität machen, und wie sich diese Zuschreibungen im Laufe des Studiums verändern. Für die Erfassung kriminalitätsbezogener Ursachenzuschreibungen stellte er, u.a. unter Rückgriff auf bereits vorhandene Instrumente, einen Fragebogen mit 250 Ursachenitems zusammen. Faktorenanalysen des im Rahmen einer Voruntersuchung auf 179 Items verkürzten Inventars ergaben 32 Skalen, die von Görge fünf übergeordneten Verursachungsbereichen zugeordnet wurden: „Täter“, „Opfer“, „mikrosoziale Bedingungen“, „makrosoziale Bedingungen“ und „anthropologische Faktoren“.

### *Deliktspezifität*

In einem Großteil der vorliegenden Studien zu Alltagstheorien von Kriminalität wird von einem „globalen“ Konzept von Kriminalität ausgegangen. Hollin und Howels (1987) haben dieses Vorgehen in Frage gestellt und untersucht, ob Alltagstheorien von Kriminalität sich tatsächlich auf Kriminalität im allgemeinen beziehen, oder ob sie sich je nach Delikt, auf das sie sich beziehen, unterscheiden. Zu diesem Zweck präsentierten Hollin und Howels (1987) ihren Versuchspersonen Fallbeschreibungen für die drei Delikte Einbruch, Raub und sexuelle Nötigung und erfaßten die Ursachenzuschreibungen mit Hilfe des Fragebogens von Furnham und Henderson (1983). Die Autoren gingen davon aus, daß sich die Delikte nach dem Ausmaß der angewendeten Gewalt gegen eine andere Person unterscheiden. Dabei ist Einbruch das Delikt mit der wenigsten und sexuelle Nötigung das Delikt mit der meisten Gewaltanwendung. Die Ergebnisse zeigten deliktspezifische Unterschiede in der Ursachenzuschreibung. Einbruch und Raub wurden hauptsächlich durch „gestörte Erziehung“ und „Eltern“ erklärt, während sexuelle Nötigung vor allem auf „mentale Instabilität“ attribuiert wurde.

Görgen (1996) untersuchte im Rahmen seiner Studie unter anderem die Hypothese, daß Gewalttaten eher auf Störungen und Krankheiten in der Person des Täters sowie auf die Täter-Opfer-Beziehung attribuiert werden als Eigentumsdelikte. Die Ergebnisse seiner Untersuchung bestätigten diese Annahmen aber nur teilweise. Die Faktoren „kriminogene Peer-Einflüsse“ und „gesamtgesellschaftliche Mißstände und Fehlorientierungen“ wurden in hohem Maße für die Erklärung sowohl von Eigentums- als auch von Gewaltdelikten herangezogen. Eigentumskriminalität wurde darüber hinaus vor allem auf „Armut und Arbeitslosigkeit des Täters“, „günstige Tatgelegenheit“ und „Armut und Arbeitslosigkeit als soziales Phänomen“ attribuiert. Gewaltkriminalität wird vor allem durch „Verhaltens- und Problembewältigungsdefizite“, „mangelnde elterliche Liebe und Akzeptanz“, „kriminogene Medieneinflüsse“ und „Sucht und Rausch“ erklärt. Görgen (1996) erfaßte auch Ursachenzuschreibungen für Kriminalität im allgemeinen und stellte fest, daß die Werte für Attributionen, die sich auf Kriminalität im allgemeinen beziehen, in den meisten Fällen zwischen denen für Eigentums- und Gewaltkriminalität lagen.

Wilken (1997) untersuchte ebenfalls die Deliktspezifität subjektiver Kriminalitätstheorien. Zu diesem Zweck verkürzte er die von Oswald und Bilsky (1991) entwickelte Skala zur Erfassung von Ursachen für die Entstehung von Kriminalität und legte sie einer Stichprobe von juristischen Laien vor. Variiert wurde dabei die Art des Deliktes, so daß die Befragten entweder zur Entstehung von Diebstahlsdelikten oder von Körperverletzungsdelikten befragt wurden. Zu diesem Zweck wurde den Versuchspersonen eine alltagsprachliche Definition des jeweiligen Deliktes vorgegeben, die von Kahe und Walther (1994) in enger Anlehnung an das StGB entwickelt worden war. Wilken (1997) untersuchte, ob sich die von Oswald und Bilsky (1991) gefundenen Entstehungsfaktoren in ihrer Bedeutsamkeit für die Ursachenzuschreibung bei Diebstahls- und Körperverletzungsdelikten unterscheiden. Die Ergebnisse zeigten, daß bei der Erklärung von Körperverletzungsdelikten den Entstehungsfaktoren „destruktive Motivation“ und „soziale Randständigkeit“ mehr Bedeutung zugemessen wurde als bei der Erklärung von Diebstahlsdelikten. Für die Erklärung von Diebstahlsdelikten wurde dem Faktor „soziales System“ höhere Bedeutsamkeit beigemessen. Beim Faktor „Plan- und Perspektivlosigkeit“ zeigten sich keine Unterschiede zwischen Diebstahls- und Körperverletzungsdelikten. Bezüglich der von Bilsky et al. (1989) identifizierten Involviertheits-Facette konnte Wilken (1997) zeigen, daß bei Diebstahlsdelikten häufiger auf Items, die das „soziale System“ betreffen, attribuiert wird als bei Körperverletzungsdelikten. Den Entstehungsfaktoren „Person“ und

„Interaktion“ hingegen kam bei der Erklärung von Körperverletzungsdelikten eine größere Bedeutung zu.

Sowohl die Ergebnisse von Hollin und Howells (1987) als auch die von Wilken (1997) geben Hinweise darauf, daß kriminalitätsbezogene Ursachenzuschreibungen insofern deliktspezifisch sind, als internalen Entstehungsfaktoren immer dann besonders viel Bedeutung zugemessen wird, wenn der Täter Gewalt gegen andere Personen anwendet. Wenn kriminalitätsbezogene Ursachenzuschreibungen tatsächlich deliktspezifisch sind, ergibt sich daraus eine eingeschränkte Aussagefähigkeit derjenigen Studien, die Ursachenzuschreibungen für Kriminalität im allgemeinen erfassen. Die Ergebnisse dieser Studien lassen in der Regel keinen Rückschluß darauf zu, ob die Befragten ihre Antworten tatsächlich auf Kriminalität im allgemeinen bezogen haben oder ob sie bei der Beantwortung der Fragen eine bestimmte Deliktart bzw. einen konkreten Fall vor Augen hatten. Görden (1996) konnte zeigen, daß Versuchspersonen, die zu ihren Attributionen in Bezug auf Straftaten ohne nähere Spezifikation befragt werden, häufig an Gewalttaten denken, während die Wahrscheinlichkeit, daß sie an Eigentumsdelikte denken, relativ gering ist. Es scheint daher sinnvoll, die Möglichkeit deliktspezifischer Unterschiede bei der Planung weiterer Untersuchungen zu kriminalitätsbezogenen Ursachenzuschreibungen zu berücksichtigen.

#### *Strafzumessung*

Bei der Untersuchung alltagstheoretischer Annahmen über den angemessenen Umgang mit Kriminalität fand auch der Einfluß von Merkmalen des Straftäters auf die Strafzumessung Beachtung. Unter anderem wurden die Effekte von Rassenzugehörigkeit, Schichtzugehörigkeit, sozio-ökonomischem Status, Geschlechtszugehörigkeit, Alter, Ausbildung, Familienstand, Drogen- und Alkoholabhängigkeit und der kriminellen Karriere analysiert (vgl. Hagan, 1989).

Gabriel und Greve (1996) beispielsweise untersuchten im Rahmen einer bundesweiten Repräsentativbefragung, ob bei Vorgabe unterschiedlicher Delikte die Tätervariablen Alter und Vorstrafe sowie die Wiedergutmachung des verursachten Schadens einen Einfluß auf die Sanktionierung des Täters haben. Den über 9000 Befragten wurden jeweils kurze Fallbeschreibungen vorgelegt mit der Bitte, aus verschiedenen Sanktionsmöglichkeiten eine einzelne Maßnahme oder eine Kombination von Maßnahmen zu wählen. Hupfeld (1996) untersuchte den Einfluß von Vorstrafen auf das jugendrichterliche Sanktionsverhalten, indem er Jugendrichtern und -richterinnen die Schilderung eines typischen Falls von Jugenddelinquenz vorgab und die Anzahl der Vorstrafen des jugendlichen Straftäters experimentell variierte. Auch Oswald (1994) untersuchte im Rahmen ihrer Analyse von Strafzumessungsentscheidungen an deutschen Amtsgerichten den Einfluß der Vorstrafenbelastung des Straftäters.

Der Einfluß der Vorstrafenbelastung auf die Strafzumessung wird in der Strafzumessungsforschung häufig als „rechtlicher“ oder „legitimierter“ Einflußfaktor bezeichnet, weil er - ähnlich wie die Schwere der Tat - in vielen Staaten als strafe erhöhend im Gesetz verankert ist (vgl. Hagan, 1989). Faktoren wie Rassenzugehörigkeit oder Geschlecht des Straftäters hingegen werden als „außerrechtlich“ oder „nicht legitimiert“ bezeichnet, weil ihr Einfluß auf das Strafzumessungsverhalten einer entsprechenden Grundlage entbehrt. Ein häufig untersuchter „nicht legitimierter“ Faktor ist der sozio-ökonomische Status, der je nach Studie auf sehr verschiedene Arten operationalisiert worden ist (vgl. Hagan, 1989; Langer, 1994).



### *Fragestellungen der Untersuchung*

Aus der zuvor skizzierten Forschung zu alltagstheoretischen Annahmen über Kriminalität und zur Angemessenheit unterschiedlicher Reaktionsformen ergaben sich vier Fragestellungen, denen in der hier vorgestellten Untersuchung nachgegangen wurde:

- (1) Läßt sich die bei der alltagstheoretischen Erklärung der Entstehung von Kriminalität gefundene Ausdifferenzierung der Lokationsdimension nach „Person“, „Interaktion“ und „soziales System“ (Bilsky et al., 1989) auch bei Verwendung eines anderen Inventars replizieren?
- (2) Variiert die zugeschriebene Bedeutsamkeit einzelner Entstehungsfaktoren bei der kriminalitätsbezogenen Ursachenzuschreibung in Abhängigkeit von der Art des Deliktes? In Anlehnung an die Arbeit von Wilken (1997) sollte dabei zwischen Diebstahls- und Körperverletzungsdelikten unterschieden werden.
- (3) Gibt es einen Zusammenhang zwischen Ursachenzuschreibungen für Kriminalität und Strafhärte?
- (4) Welchen Einfluß haben Vorstrafe und sozio-ökonomischer Status des Straftäters auf das Strafzumessungsverhalten?

## **2 Methode**

Die empirische Beantwortung der genannten Fragestellungen sollte durch Interviews juristischer Laien anhand eines mehrteiligen standardisierten Fragebogens erfolgen. Für die Erfassung kriminalitätsbezogener Ursachenzuschreibungen war dabei zunächst die Entwicklung eines unter wissenschaftlichen und verfahrensökonomischen Aspekten gleichermaßen brauchbaren Instrumentes erforderlich. Sie erfolgte im Rahmen einer Voruntersuchung, auf die nachfolgend näher eingegangen wird.

### 2.1 Voruntersuchung: Erfassung kriminalitätsbezogener Ursachenzuschreibungen

Grundlage des in unserer Untersuchung einzusetzenden Instrumentes bildete das Inventar von Görge (1996), das von seinem Autor in drei Versionen für Gewaltkriminalität, Eigentumskriminalität und Kriminalität im allgemeinen verwendet worden war. In einem ersten Schritt wurden die Items dieses Inventars aufgrund zuvor festgelegter Kriterien (vgl. Reichert, 1999) danach klassifiziert, ob sie sich primär (a) auf die Person des Täters, die Interaktion des Täters mit seiner Umwelt oder auf das soziale System bzw. die Gesellschaft beziehen und (b) zur Erklärung von Diebstahlsdelikten, von Körperverletzungsdelikten oder beider Deliktarten geeignet erscheinen. An dieser Klassifikationsaufgabe waren zwischen 16 und 21 Versuchspersonen beteiligt. In einem zweiten Schritt wurden die von Görge (1996) erhobenen Daten<sup>1</sup> einer MDS unterzogen, um zu überprüfen, inwieweit sich die a priori-Klassifikationen der Items im Hinblick auf die erweiterte Lokationsdimension empirisch bestätigen läßt. Nach Ausschluß von 11 Items, die sich auf das Kriminalitätsoffer beziehen, ergab sich eine weitgehend erwartungskonforme Partitionierung der verbleibenden Items (17 Fehlplatzierungen). Der dritte Schritt bestand darin, die Itemzahl auf der Grundlage der zuvor durchgeführten MDS sowie zusätzlicher Kriterien auf einen für unsere Untersuchung akzeptablen Umfang zu reduzieren. Zu diesen zusätzlichen Kriterien zählte (a) der Anspruch, Items aus möglichst vielen der 32 von Görge (1996) identifizierten Skalen zu berücksichtigen, um eine breite inhaltliche Varianz zu gewährleisten, (b) der vermutete Erklärungswert für die Entstehung von Diebstahlsdelikten und Körperverletzungsdelikten sowie (c) die von Görge (1996) im Rahmen seiner Untersuchung errechneten Itemkennwerte (Trennschärfe und Itemschwierigkeit). Auf diese

Weise wurde der ursprüngliche Fragebogen in mehreren Schritten auf 39 Items verkürzt; diese Version ist im Anhang (Tabelle 1) wiedergegeben.

## 2.2 Hauptuntersuchung

### *Instrument*

Für die Untersuchung der zuvor erläuterten Fragestellungen im Rahmen von Interviews war es erforderlich, einen standardisierten Fragebogen zu erstellen, der es gestattet, Strafzumessungsurteile und Ursachenzuschreibung der Versuchspersonen jeweils deliktspezifisch zu erfassen. Dieses Instrument besteht aus insgesamt drei Teilen:

In *Teil I* des Fragebogens werden kurze *Fallbeschreibungen* (Vignetten) eines Diebstahlsdelikts und eines Körperverletzungsdelikts vorgegeben, zu denen jeweils eine von mehreren Sanktionsmöglichkeiten auszuwählen ist. Die Deliktbeschreibungen lehnen sich eng an das StGB an. Für beide Delikte sind gleiche Strafrahmen vorgesehen; hierdurch ist die Vorgabe gleicher Sanktionsalternativen möglich. Um den Einfluß der Tätervariablen „Vorstrafe“ und „sozio-ökonomischer Status“ auf die Strafzumessung zu überprüfen, werden sie bei beiden Delikten jeweils zweistufig variiert. Beispiele für die verwendeten Vignetten sind nachfolgend wiedergegeben.

Diebstahlsdelikt: Ein 30jähriger Mann bestiehlt im Flughafengebäude mehrere Reisende und erbeutet dabei Bargeld, Fotoapparate und zwei Lederjacken im Gesamtwert von 6.200,- DM sowie einige Kreditkarten. Der Täter steht in einem festen Arbeitsverhältnis. Er stand bislang noch nie vor Gericht.

Körperverletzungsdelikt: Ein 30jähriger Mann fängt mit einem anderen Besucher einer Gaststätte Streit an und schlägt mit den Fäusten auf ihn ein, bis dieser am Boden liegt. Anschließend versetzt er ihm zwei Tritte an Kopf und Körper. Das Opfer erleidet Prellungen, eine Gehirnerschütterung und eine Platzwunde am Kopf, die genäht werden muß. Der Täter ist arbeitslos. Er ist früher bereits zweimal wegen Körperverletzung verurteilt worden.

Als Reaktionsmöglichkeiten stehen 13 verschiedene Sanktionen zur Auswahl, die sich in vier Sanktionsgruppen zusammenfassen lassen: „Ermahnung mit Einstellung“, „Geldstrafe“, „Freiheitsstrafe mit Bewährung“ und „Freiheitsstrafe ohne Bewährung“. Die auf Kärtchen notierten Sanktionsalternativen werden vor der Versuchsperson so ausgebreitet, daß die einzelnen Maßnahmen einer Sanktionsart (z.B. Geldstrafe) als Gruppe nach Strafhöhe geordnet untereinander liegen; die Sanktionsgruppen liegen nebeneinander. Ferner erhält jeder Befragte eine Vorlage, auf der er die vom Interviewer vorgelesenen Instruktionen und Deliktbeschreibungen mitliest. Tabelle 1 gibt einen Überblick über die Alternativen.

Tabelle 1: Sanktionsmöglichkeiten

Code	Sanktion
E1	Ermahnung mit Einstellung (das Verfahren wird ohne weitere Folgen für den Täter eingestellt)
G1	Geldstrafe, 15 Tagessätze (also das Nettoeinkommen bzw. das Arbeitslosengeld des Täters für einen halben Monat)
G2	Geldstrafe, 30 Tagessätze (also das Nettoeinkommen bzw. das Arbeitslosengeld des Täters für einen Monat)
G3	Geldstrafe, 90 Tagessätze (also das Nettoeinkommen bzw. das Arbeitslosengeld des Täters für drei Monate)
G4	Geldstrafe, 180 Tagessätze (also das Nettoeinkommen bzw. das Arbeitslosengeld des Täters für sechs Monate)
B1	3 Monate Freiheitsstrafe <u>mit</u> Bewährung
B2	6 Monate Freiheitsstrafe <u>mit</u> Bewährung
B3	1 Jahr Freiheitsstrafe <u>mit</u> Bewährung
B4	2 Jahre Freiheitsstrafe <u>mit</u> Bewährung
F1	1 Monat Freiheitsstrafe <u>ohne</u> Bewährung
F2	3 Monate Freiheitsstrafe <u>ohne</u> Bewährung
F3	6 Monate Freiheitsstrafe <u>ohne</u> Bewährung
F4	1 Jahr Freiheitsstrafe <u>ohne</u> Bewährung

In *Teil II* des Fragebogens erfolgt die Erfassung von *Ursachenzuschreibungen für Kriminalität* mit Hilfe des in der Voruntersuchung erstellten Instrumentes (vgl. Anhang). Für die deliktspezifische Erfassung wurden zwei Versionen erstellt, die in der Instruktion zunächst wiederum eine alltagssprachliche Definitionen des betreffenden Delikts enthalten; diese Definitionen wurden von Kahe und Walther (1994; zit. nach Wilken, 1997, S. 42) übernommen:

Diebstahl: „Bitte stellen Sie sich dabei Diebstahlsdelikte vor, bei denen einer Person absichtlich ein Gegenstand entwendet wird, ohne daß dabei eine Waffe im weitesten Sinne verwendet oder in einen geschlossenen Raum eingebrochen wird.“

Körperverletzung: „Bitte stellen Sie sich dabei Körperverletzungsdelikte vor, bei denen eine Person absichtlich geschlagen oder gesundheitlich geschädigt wird, ohne daß dabei eine Waffe im weitesten Sinne verwendet wird oder Schäden hinterbleiben.“

Auf der Grundlage dieser Definition bearbeiten die Versuchspersonen die auf 39 Items verkürzte Form des Görden-Inventars. Dabei werden die Items jeweils deliktspezifisch eingeleitet („Zu Diebstahlsdelikten kommt es, weil ...“ bzw. „Zu Körperverletzungsdelikten kommt es, weil ...“). Auch in diesem Fragebogenteil liest der Interviewer die Instruktion und die einzelnen Items vor, während die Befragten anhand einer Vorlage mitlesen. Anschließend drücken sie den Grad ihrer Zustimmung zu den einzelnen Aussagen auf einer siebenstufigen Skalen von 0 (trifft überhaupt nicht zu) bis 6 (trifft voll und ganz zu) aus; diese liegt ihnen ebenfalls schriftlich vor. Der Interviewer notiert die Antworten im Fragebogen. Es steht den Befragten dabei frei, gegebenenfalls mit „weiß nicht“ zu antworten.

*Teil III* des Fragebogens dient der Erfassung der durch die Versuchspersonen wahrgenommenen *Strafhärte* der einzelnen Sanktionen. Zu diesem Zweck werden alle Versuchspersonen aufgefordert, die Kärtchen mit den Sanktionsvorschlägen, die in Teil I des Fragebogens als Reaktionsmöglichkeiten zur Auswahl standen, nach der von ihnen subjektiv empfundenen Strafhärte zu ordnen. Abschließend werden von den befragten Personen die folgenden biographischen Angaben erhoben: Alter, Geschlecht, Beruf, Wohnortgröße und höchster Schul- bzw. Hochschulabschluß.

### *Plan der Untersuchung*

In Teil I des Fragebogens wurden den Versuchspersonen jeweils eine Variante des Diebstahls- und des Körperverletzungsdeliktes vorgegeben. Zum Ausschluß von Reihenfolgeeffekten wurden die Delikte in unterschiedlicher Abfolge dargeboten. Aus der Variation von Vorstrafe und sozio-ökonomischem Status des Täters und der Rotation der Reihenfolge ergaben sich vier Versionen für Teil I des Fragebogens (Tabelle 2).

Tabelle 2: Rotationsplan für Teil I des Fragebogens

Version	Delikt 1		Delikt 2	
	Deliktart	Tättervariablen	Deliktart	Tättervariablen
1	Diebstahl	festes Arbeitsverhältnis nicht vorbestraft	Körperverletzung	arbeitslos vorbestraft
2	Diebstahl	arbeitslos vorbestraft	Körperverletzung	festes Arbeitsverhältnis nicht vorbestraft
3	Körperverletzung	festes Arbeitsverhältnis vorbestraft	Diebstahl	arbeitslos nicht vorbestraft
4	Körperverletzung	arbeitslos nicht vorbestraft	Diebstahl	festes Arbeitsverhältnis vorbestraft

Die mit Teil II des Fragebogens erhobenen Ursachenzuschreibungen erfolgten alternativ für Diebstahls- oder Körperverletzungsdelikte. Aus der experimentellen Variation der Tätervariablen in Teil I und der alternativen Erfassung von Ursachenzuschreibungen mit Bezug auf Diebstahls- oder Körperverletzungsdelikte in Teil II ergaben sich somit acht Materialsätze, die jeweils aus einem Fragebogen für den Interviewer und entsprechenden Vorlagen für den Befragten bestanden. Angestrebt war die Untersuchung von 200 Versuchspersonen im Alter von 16 oder mehr Jahren.

### *Stichprobe*

Die Datenerhebung erfolgte 1999 in Osnabrück in Räumen des Marktforschungsinstitutes *Produkt + Markt* durch insgesamt sechs Interviewer. Die Mehrzahl der Versuchspersonen wurde in der Fußgängerzone angesprochen; da es sich als schwierig erwies, ältere Menschen auf der Straße für die Durchführung eines Interviews zu gewinnen, wurden zusätzlich 13 der Interviews bei Personen ab 60 Jahren durchgeführt. Die durchschnittliche Interviewdauer betrug etwa 15 Minuten.

Insgesamt nahmen 211 Personen im Alter von 16 bis 79 Jahren an der Befragung teil. Drei Interviews mußten abgebrochen werden, zwei davon wegen Verständigungsschwierigkeiten mit Teilnehmern ausländischer Herkunft und eins auf Wunsch des Befragten. Von den 208 auswertbaren Fragebögen entfallen 101 auf Version D (Diebstahlsdelikte) und 107 auf Version K (Körperverletzungsdelikte). Der Altersdurchschnitt liegt bei 41 Jahren. Tabelle 3 zeigt die Verteilung auf die verschiedenen Altersgruppen.

Tabelle 3: Verteilung der Versuchspersonen auf die Altersgruppen

Altersgruppe	Geschlecht		Σ
	männlich	weiblich	
16 - 24 Jahre	25	26	51
25 - 39 Jahre	27	26	53
40 - 59 Jahre	28	25	53
ab 60 Jahre	25	26	51
Σ	105	103	208

Die am häufigsten vertretenen Berufsgruppen in der Stichprobe sind Angestellte (33%), Studenten (16%), Rentner (14%), Hausfrauen (12%), Selbständige (8%) und Schüler (7%). Als höchsten Schul- bzw. Hochschulabschluß gaben 18% der Befragten einen Haupt- bzw. Volksschulabschluß, 21% einen Realschulabschluß, 7% die Fachhochschulreife, 26% Abitur und 20% ein abgeschlossenes Studium an. Die Stichprobe ist nicht repräsentativ für die Bevölkerung der Bundesrepublik Deutschland, kann aber als hinreichend heterogen angesehen werden, so daß Störeinflüsse durch ein bestimmtes Bildungsniveau oder eine bestimmte soziale Schicht ausgeschlossen werden können.

### 3 Ergebnisse

#### 3.1 Dimensionalität der Ursachenzuschreibungen

Zur Prüfung, ob der vorliegende Datensatz wie erwartet eine Ausdifferenzierung der Lokationsdimension in der in Fragestellung 1 spezifizierten Form ermöglicht, wurden mit den deliktsspezifischen Ursachenzuschreibungen der Gesamtstichprobe zwei- und dreidimensionale Multidimensionale Skalierungen (MDS) nach Guttman gerechnet. Während die zweidimensionale Lösung keine eindeutige Partitionierung im Sinne unserer Erwartungen zuließ, war dies auf den Dimensionen zwei und drei einer dreidimensionalen MDS problemlos möglich (vgl. Abbildung 1 im Anhang). Lediglich zwei Items der Skala „soziales System“ (Items 8 und 31) und ein Item der Skala „Interaktion“ (Item 3) waren nicht erwartungskonform plaziert; sie wurden daher bei der Bildung der drei in den weiteren Analysen eingesetzten Attributionsskalen nicht berücksichtigt.

#### 3.2 Deliktsspezifität von Ursachenzuschreibungen für Kriminalität

Mit Hilfe von t-Tests für unabhängige Stichproben wurde die unterschiedliche Bedeutung der Faktoren „Person“, „Interaktion“ und „soziales System“ bei der Erklärung von Diebstahls- und Körperverletzungsdelikten analysiert. Die jeweiligen Skalenmittelwerte sind in Abbildung 1 wiedergegeben. Während sich für den Entstehungsfaktor „Person“ ein signifikanter Unterschied ergab ( $p < .01$ ), erwiesen sich die Differenzen bei den übrigen Ursachenfaktoren als nicht signifikant.

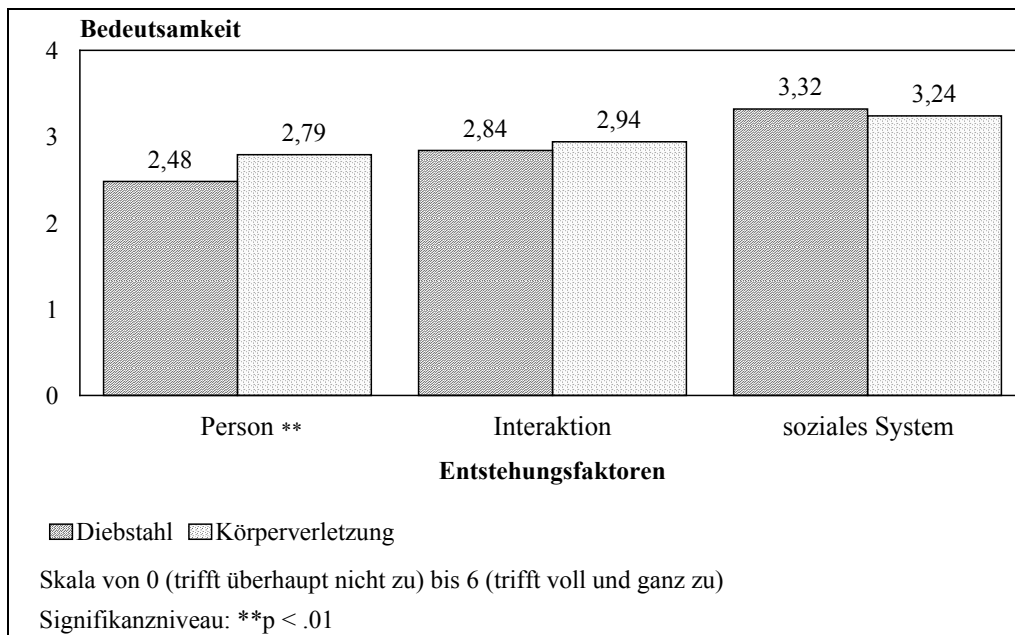


Abbildung 1: Mittelwertvergleich der Attribution auf „Person“, „Interaktion“ und „soziales System“ bei der Ursachenzuschreibung für Diebstahls- und Körperverletzungsdelikte

### 3.3 Zusammenhang zwischen Ursachenzuschreibungen für Kriminalität und Strafhärte

Um den in Fragestellung 3 angesprochenen Zusammenhang zwischen kriminalitätsbezogenen Ursachenzuschreibungen und Strafhärte untersuchen zu können, war es erforderlich, die einzelnen in Teil I des Fragebogens zur Auswahl stehenden Sanktionsmöglichkeiten auf einer Strafhärteskala einzuordnen. Auf Basis der in Teil III des Fragebogens von allen Versuchspersonen erstellten individuellen Rangreihen wurde daher zunächst für jede Sanktionsmöglichkeit ein mittlerer Strafhärte-Wert berechnet (vgl. Anhang, Tabelle 2).

Anschließend wurden je zwei Regressionsanalysen für Diebstahls- und Körperverletzungsdelikte mit Strafhärte als Kriterium und den Ursachenfaktoren „Person“, „Interaktion“ und „soziales System“ als Prädiktoren gerechnet; dabei bezogen sich Kriterium und Prädiktoren in jeweils einer der Analysen auf dasselbe (z.B. Strafhärte bei Diebstahl; Ursachenfaktoren für Diebstahl), in der anderen auf das alternative Delikt (z.B. Strafhärte bei Körperverletzung; Ursachenfaktoren für Körperverletzung).

Die Regression von Strafhärte bei *Diebstahlsdelikten* auf *diebstahlsbezogene Attributionsfaktoren* ermöglichte eine schwache aber überzufällige Vorhersage des Kriteriums ( $R^2=.06$ ,  $N=101$ ). Dabei erwies sich die Attribution auf das "soziale System" als effektiver Prädiktor ( $t=2.29$ ,  $p=.02$ ); der positive Beta-Koeffizient zeigt, daß eine höhere Gewichtung dieses Faktors mit einer höheren Strafzumessung einher geht. Dieser auf den ersten Blick kontraintuitive Zusammenhang wird bei inhaltlicher Sichtung der unter "soziales System" subsumierten Items verständlich, die hier nicht, wie in anderen Studien, eine Relativierung der individuellen Verantwortlichkeit aufgrund gesellschaftlicher Rahmenbedingungen nahelegen, sondern im Sinne eines "Mehr an Recht und Ordnung" zu interpretieren sind (vgl. Anhang). Für die Regression von Strafhärte bei *Körperverletzungsdelikten* auf *Ursachenzuschreibung für Körperverletzungsdelikte* ergab sich eine deutlich bessere Varianzaufklärung ( $R^2=.12$ ,  $N=107$ ). Hier war es die Attribution auf die Person des Täters, die überzufällig zur Vorhersage des Kriteriums beitrug ( $t=3.56$ ,

$p=.001$ ), während die Prädiktoren "Interaktion" und "soziales System" keinen Beitrag zur Varianzaufklärung leisteten.

Zwischen diebstahlsbezogenen Ursachenzuschreibungen und der Strafzumessung bei Körperverletzungsdelikten konnte mit Hilfe der Regressionsanalyse kein Zusammenhang festgestellt werden. Ebenso erlaubte die Ursachenzuschreibung in Bezug auf Körperverletzungsdelikte erwartungsgemäß keine Vorhersage auf das Strafverhalten bei Diebstahlsdelikten.

### 3.4 Der Einfluß von Tätervariablen auf die Strafhärte

Die Überprüfung des Einflusses von Tätervariablen auf die Strafhärte bei Diebstahls- und Körperverletzungsdelikten erfolgte jeweils mittels einer mehrfaktoriellen Kovarianzanalyse. In ihr fungierten das Alter und das Geschlecht des Befragten als Kontrollvariablen, die Vorstrafen sowie der sozioökonomische Status des Täters als unabhängige Variablen.

Für das Diebstahlsdelikt ergab die Analyse eine Varianzaufklärung von rund 25%. Dabei leistete die Regression auf die beiden Kontrollvariablen einen signifikanten Beitrag ( $p < .05$ ). Ferner erwiesen sich beide Haupteffekte als signifikant. Beim Körperverletzungsdelikt konnten durch die Kovarianzanalyse insgesamt 35% der Strafhärtevarianz aufgeklärt werden. Hier leistete die Regression auf die beiden Kontrollvariablen keinen Beitrag, während sich sowohl die beiden Haupteffekte als auch die Interaktion als signifikant erwiesen. Die Ergebnisse beider Analysen sind in den Tabellen 4 und 5 sowie den Abbildungen 2 und 3 zusammengefaßt.

Tabelle 4: Mehrfaktorielle Kovarianzanalyse: Diebstahlsdelikt

Strafhärte (AV)	SS	df	MS	F	Signif. F
Regression	44.14	2	22.07	3.50	.032
Vorstrafe	332.50	1	332.50	52.69	.000
sozio-ökonomischer Status	34.78	1	34.78	5.51	.020
Vorstrafe x sozio-ökonomischer Status	1.04	1	1.04	.16	.686
Fehlervarianz	1274.76	202	6.31		
Total	1704.13	207	8.23		

Determinationskoeffizient:  $R^2 = .252$

Tabelle 5: Mehrfaktorielle Kovarianzanalyse: Körperverletzungsdelikt

Strafhärte (AV)	SS	df	MS	F	Signif. F
Regression	1.57	2	.78	.13	.881
Vorstrafe	468.70	1	468.70	75.65	.000
Arbeit	24.91	1	24.91	4.02	.046
Vorstrafe x sozio-ökonomischer Status	175.19	1	175.19	28.28	.000
Fehlervarianz	1251.54	202	6.20		
Total	1923.92	207	9.29		

Determinationskoeffizient:  $R^2 = .349$

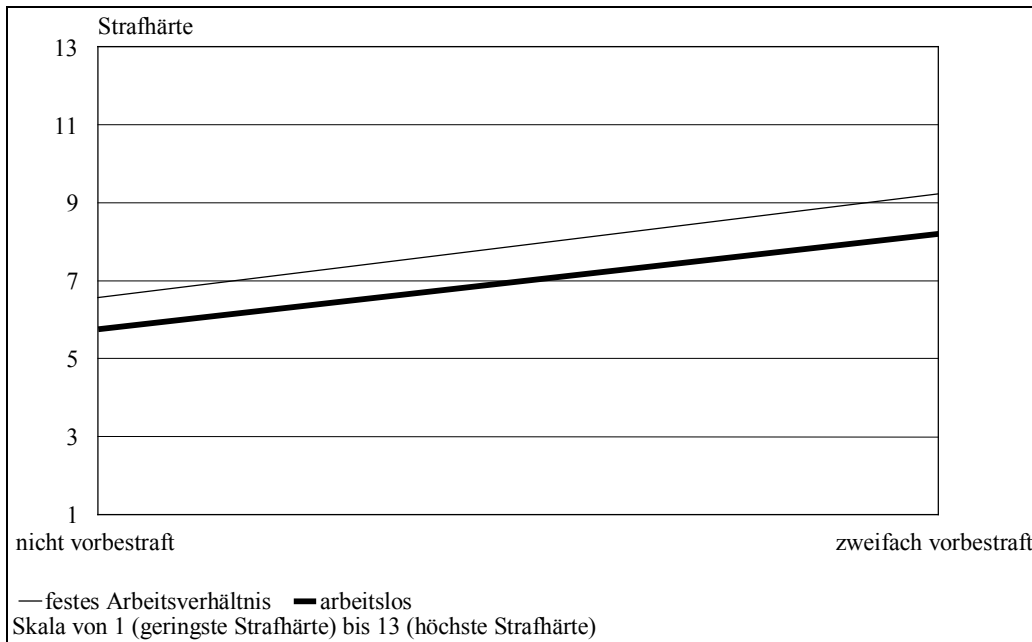


Abbildung 2: Einfluß von Tätervariablen auf die Strafhärte bei Diebstahlsdelikten

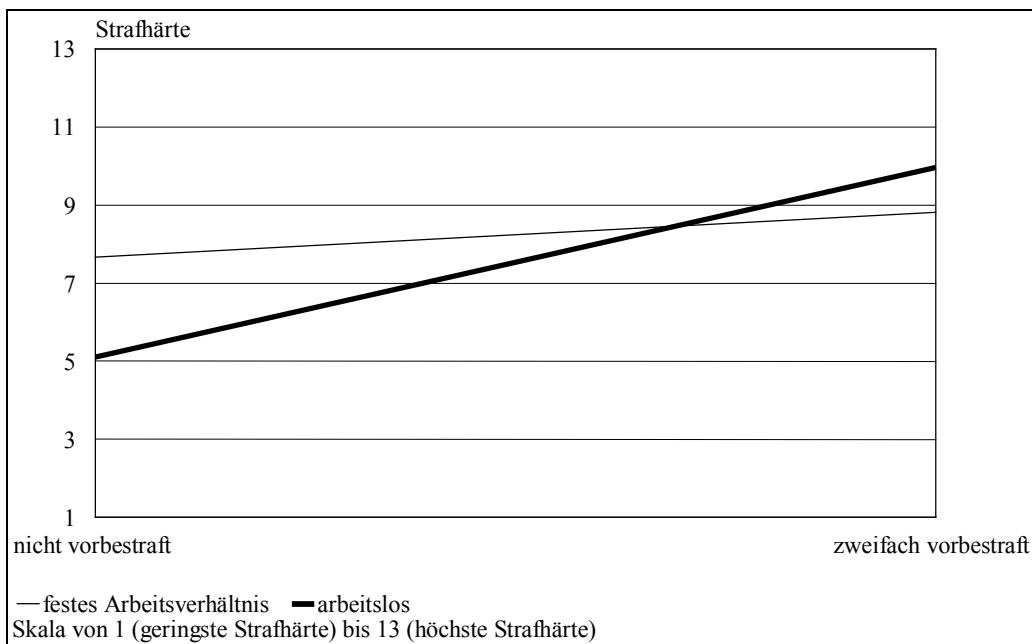


Abbildung 3: Einfluß von Tätervariablen auf die Strafhärte bei Körperverletzungsdelikten

#### 4 Diskussion

Sowohl bei der Reanalyse eines von Görden zur Verfügung gestellten Datensatzes alltags-theoretischer Annahmen über Kriminalität als auch aufgrund der in unserer Hauptuntersuchung erhobenen Daten war es möglich, Faktoren, die als wichtig für die Entstehung von Kriminalität erachtet werden, entsprechend einer Erweiterung der Lokationsdimension von Weiner (1986; Bilsky et al., 1989) nach „Person“, „Interaktion“ und „sozialem System“ zu unterscheiden. Items des Görden-Inventars, in denen eine Täter-Opfer-Beziehung thematisiert wird, waren im Rahmen unserer Voruntersuchung zunächst als „Interaktions“-Items klassifiziert worden, fielen jedoch bei der anschließenden MDS



relativ geschlossen in den Bereich der Person-Faktoren. Dies ist insofern stimmig, als es sich bei den meisten dieser Items um Erklärungen handelt, in denen dem Opfer ein gewisses Maß an Mitschuld an der Straftat zugeschrieben wird (z.B. „Zu Straftaten kommt es, weil die Opfer sich zuvor aggressiv gegenüber den Tätern verhalten haben“; Görge, 1996, S. 276). Da auf die Rolle des Opfers bzw. der Täter-Opfer-Beziehung im Rahmen der vorliegenden Untersuchung nicht weiter eingegangen werden sollte, blieben die entsprechenden Items bei der Bildung der von uns verwendeten Attributionsskalen unberücksichtigt. Es erscheint jedoch sinnvoll, der Rolle von Opfern bei der Analyse laientheoretischer Annahme über die Entstehung von Kriminalität verstärkt Beachtung zu schenken.

Bezüglich der Deliktspezifität von Ursachenzuschreibungen konnten die Ergebnisse von Wilken (1997) nur teilweise repliziert werden. Erwartungsgemäß war die Attribution auf den Faktor „Person“ bei Körperverletzungsdelikten höher als bei Diebstahlsdelikten. Dies ist ein Hinweis dafür, daß bei der körperlichen Schädigung anderer eher auf internale Ursachenfaktoren attribuiert wird, so wie dies auch schon Hollin und Howells (1987) zeigen konnten. Für die Attribution auf die Faktoren „Interaktion“ und „soziales System“ konnten demgegenüber keine signifikanten Unterschiede zwischen Diebstahls- und Körperverletzungsdelikten gefunden werden. Eine mögliche Erklärung dafür könnte sein, daß die Versuchspersonen die Instruktion, ihre Antwort auf Diebstahls- bzw. auf Körperverletzungsdelikte zu beziehen, nicht in gewünschtem Maße umsetzen. Görge (1996) bemerkt dazu: „Subjektiven Deliktfokussierungen kommt ... oft höhere prädiktive Bedeutung zu als den vorgegebenen Kriminalitätsbereichen.“ Gegen diese Erklärung spricht, daß Wilken (1997) die gleiche Definition für Diebstahls- und Körperverletzungsdelikte wie die in der vorliegenden Untersuchung verwendet und Unterschiede im Ausmaß der Attribution auf „Interaktion“ und „soziales System“ gefunden hat.

Unsere Ergebnisse weisen ferner darauf hin, daß ein Zusammenhang zwischen den kriminalitätsbezogenen Ursachenzuschreibungen einer Person und der von ihr präferierten Strafhärte existiert. Dieser Zusammenhang ist deliktspezifisch, d.h. die Ursachenzuschreibungen in Bezug auf Diebstahlsdelikte lassen eine Vorhersage auf das Strafzumessungsverhalten bei Diebstahlsdelikten zu, nicht aber bei Körperverletzungsdelikten. Ebenso lassen die Ursachenzuschreibungen für Körperverletzungsdelikte nur eine Vorhersage auf das Strafzumessungsverhalten bei Körperverletzungsdelikten zu, nicht aber bei Diebstahlsdelikten. Der Einfluß der deliktspezifischen Ursachenzuschreibungen auf die Strafhärte erscheint dabei mit einer Varianzaufklärung von etwa 6% für Diebstahlsdelikte bzw. 12% für Körperverletzungsdelikte relativ gering. In Anbetracht zahlreicher weiterer Variablen, die Einfluß auf die Höhe einer zuzumessenden Strafe haben dürften, kann dies jedoch kaum verwundern. So müssen eine Reihe weiterer Faktoren wie die Deliktschwere, der Tathergang, Opfererfahrungen der Versuchspersonen und weitere allgemeine Strafeinstellungen der Versuchspersonen wie die Strafbegründungspräferenz (vgl. Oswald, 1994) als mögliche Einflußfaktoren in Erwägung gezogen werden.

In Bezug auf Körperverletzungsdelikte existiert erwartungsgemäß eine positive Korrelation zwischen der Attribution auf die Person des Täters und der Strafhöhe. Nach dem von Jones und Davis (1965) vorgeschlagenen Prinzip der nicht gemeinsamen Effekte wird bei Körperverletzungsdelikten eher auf eine stabile Disposition des Täters geschlossen, weil die Gewaltanwendung gegen eine andere Person und deren Verletzungen als Konsequenz der Tat ungewöhnlicher sind als ein Diebstahlsdelikt. Eine Attribution auf stabile Faktoren

führt laut Carroll und Payne (1976) zu einer Präferenz für eine Freiheitsstrafe, welche in der vorliegenden Untersuchung härter eingestuft wurde als eine Geldstrafe. Die durch die Ungewöhnlichkeit der Handlung hervorgerufene Fokussierung auf den Täter als handelnde Person könnte auch der Grund dafür sein, daß das Ausmaß der Attribution auf das soziale System keinen Vorhersagewert für das Strafzumessungsverhalten bei Körperverletzungsdelikten hat.

Der bei Diebstahlsdelikten gefundene positive Zusammenhang zwischen der Attribution auf das soziale System und der Strafhärte widerspricht nur scheinbar den theoretischen Annahmen von Carroll und Payne (1976) und Oswald und Bilsky (1991). Tatsächlich ist das erwartungskonträre Ergebnis in der Auswahl der Items begründet. Von den 12 Items der Skala „soziales System“, die nach Durchführung der MDS in die weitere Datenanalyse eingingen, beinhalten sechs Items (4, 12, 15, 18, 33 und 39) Forderungen nach einem härteren Vorgehen gegenüber Straftätern. Wenn also eine Versuchsperson harte Strafen generell befürwortet, so wird sie hohe Werte bei diesen Items und somit auch auf der Skala „soziales System“ aufweisen. Gleichzeitig wird sie eine entsprechend hohe Strafe als Reaktion auf das in Teil I des Fragebogens beschriebene Diebstahlsdelikt auswählen.

Bezüglich des Einflusses von Tätervariablen auf die Strafzumessung zeigte sich, daß die Vorstrafe eines Täters von erheblicher Bedeutung ist. Auch Hupfeld (1996) und Oswald (1994) fanden bei der Analyse realer Gerichtsurteile einen großen Einfluß der Vorstrafenbelastung von (jugendlichen) Tätern auf die Strafhärte. Hupfeld (1996) spricht in diesem Zusammenhang kritisch von einer „Sanktionseskalation“ (u.a. S. 81). Auch Oswald (1997) macht die Bedenklichkeit dieser Entwicklung deutlich: „..., da es anerkanntermaßen *nicht* darum gehen sollte, so etwas wie ‘Lebensschuld’ zu begleichen“ (S. 254). Im Zusammenhang mit der vorliegenden Untersuchung könnte der Einfluß der Vorstrafen zusätzlich erhöht sein, da durch die Kürze der Deliktbeschreibung die Information über die Vorstrafenbelastung des Täters möglicherweise stärker gewichtet wird.

Die Information darüber, ob der Täter in einem festen Arbeitsverhältnis steht oder arbeitslos ist, hat für sich genommen nur einen geringen Einfluß auf das Strafverhalten. Beim Diebstahlsdelikt werden arbeitslose Täter erwartungsgemäß weniger hart bestraft als Täter, die in einem festen Arbeitsverhältnis stehen. Gemäß dem von Oswald und Bilsky (1991) vertretenen Modell der Verantwortungszuschreibung könnte die finanzielle Notlage durch die Arbeitslosigkeit als entschuldigender Faktor gewertet werden, der die Verantwortungszuschreibung und somit die Strafzumessung mindert. Beim Körperverletzungsdelikt gibt es einen Interaktionseffekt zwischen Vorstrafe und sozio-ökonomischem Status. Während ein nicht vorbestrafter Täter, der in einem festen Arbeitsverhältnis steht, wesentlich härter bestraft wird als ein Täter, der arbeitslos ist, kehrt sich dieses Verhältnis bei vorbestraften Tätern um. Offensichtlich führt eine Kombination der Informationen über Vorstrafe und Arbeitslosigkeit des Täters in Verbindung mit einem Körperverletzungsdelikt tatsächlich dazu, daß eine starke Verantwortungszuschreibung auf den Täter und somit eine härtere Bestrafung erfolgt.

Zusammenfassend läßt sich feststellen, daß die Befunde zur Dimensionalität von Ursachenzuschreibungen für Kriminalität eine eindeutige Partitionierung in die Bereiche „Person“, „Interaktion“ und „soziales System“ zeigen. Der verhältnismäßig große Einfluß von Vorstrafen auf die Strafzumessung konnte ein weiteres Mal eindeutig nachgewiesen werden. Für den sozio-ökonomischen Status des Täters wurde ein mäßiger Einfluß auf die Strafzumessung ermittelt. Zur Deliktspezifität von Ursachenzuschreibungen und zum

Zusammenhang zwischen Ursachenzuschreibungen für Kriminalität und der Strafzumessung sind die Ergebnisse weniger eindeutig. Da sich die Frage stellt, inwieweit die Abweichungen von den erwarteten Ergebnissen auf die Operationalisierung im Rahmen dieser Arbeit zurückgehen, sollen im folgenden einige Aspekte der Untersuchungsmethode kritisch betrachtet werden.

Die Ergebnisse der Strafhärteskalierung zeigen, daß die häufig vertretene Annahme, Freiheitsstrafen seien auf jeden Fall härter zu beurteilen als Geldstrafen, nicht zutrifft. Es scheint also erforderlich zu sein, die von den Versuchspersonen subjektiv wahrgenommene Strafhärte der einzelnen Sanktionen in der jeweils interessierenden Stichprobe empirisch zu erfassen. Allerdings muß berücksichtigt werden, daß in der vorliegenden Untersuchung die Auswahl der zur Verfügung stehenden Sanktionsmöglichkeiten möglicherweise die Strafhärteskalierung beeinflußt hat. Einige Versuchspersonen haben eine Rangreihe nach dem Muster: niedrigste Geldstrafe - niedrigste Bewährungsstrafe - zweitniedrigste Geldstrafe - zweitniedrigste Bewährungsstrafe gebildet. Somit wäre der Rangplatz einer Geldstrafe von 30 Tagessätzen ein anderer, wenn die Geldstrafe von 15 Tagessätzen nicht zur Auswahl gestanden hätte. Die Strafhärtewerte sind also nur in Relation zueinander zu deuten.

Weiterhin muß zum Vorgehen bei der Strafhärteskalierung angemerkt werden, daß es den Versuchspersonen zum Teil schwerfiel, Geldstrafen und Freiheitsstrafen in ihrer Strafhärte zu vergleichen, ohne einen konkreten Täter vor Augen zu haben. Da es nachvollziehbar scheint, daß vor allem Geldstrafen verschiedene Personen je nach ihrer finanziellen Situation unterschiedlich hart treffen, wäre es unter Umständen ratsam, bei der Erfassung der wahrgenommenen Strafhärte von Sanktionen ein spezifisches Täterprofil vorzugeben.

Bezüglich des Zusammenhangs zwischen Ursachenzuschreibungen für Kriminalität und Strafhärte sind nach wie vor viele Fragen offen. Die praktische Relevanz des Problems der Strafzumessungsunterschiede in der Rechtsprechung sowie die methodischen Herausforderungen auf diesem interdisziplinären Forschungsfeld lassen weitere Untersuchungen lohnend erscheinen.

## Literatur

- Bilsky, W., Oswald, M.E. & Marciszewski, B. (1989). *Judges' subjective theories of criminal behavior*. Paper presented at the Second International Facet Theory Conference, July 4-6, 1989, University of Surrey, UK.
- Carroll, J.S. & Payne, J.W. (1976). The psychology of the parole decision process: A joint application of attribution theory and information-processing psychology. In J.S. Carroll & J.W. Payne (Eds.), *Cognition and social behavior* (S. 13-32). Hillsdale, New Jersey: Lawrence Erlbaum.
- Erskine, H. (1974). The polls: Causes of crime. *Public Opinion Quarterly*, 38, 288-298.
- Furnham, A. (1988). *Lay Theories: Everyday understanding of problems in social science*. Oxford: Pergamon Press.
- Furnham, A. & Henderson, M. (1983). Lay theories of delinquency. *European Journal of Social Psychology*, 13, 107-120.
- Gabriel, U. & Greve, W. (1996). „Strafe muß sein!“. Sanktionsbedürfnisse und strafbezogene Einstellungen: Versuch einer systematischen Annäherung. In C. Pfeiffer & W. Greve (Hrsg.), *Forschungsthema „Kriminalität“* (S. 185-214). Baden-Baden: Nomos.

- Görge, T. (1996). *Kriminalitätstheorien angehender Juristinnen und Juristen: eine attributionstheoretische Analyse* (Giessener Kriminalwissenschaftliche Schriften; Bd. 7). Mönchengladbach: Forum Verlag Godesberg.
- Hagan, J. (1989). Strafzumessungsforschung in Nord-Amerika. In C. Pfeiffer & M.E. Oswald (Hrsg.), *Strafzumessung: Empirische Forschung und Strafrechtsdogmatik im Dialog*. Stuttgart: Enke.
- Hollin, C.R. & Howells, K. (1987). Lay explanations of delinquency: Global or offense specific? *British Journal of Social Psychology*, 26, 203-210.
- Hupfeld, J. (1996). *Jugendrichterliches Handeln: Eine Analyse der Reaktionen auf Rückfalldelinquenz aus psychologischer Perspektive*. Baden-Baden: Nomos.
- Jones, E.E. & Davis, K.E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, Vol 2. New York: Academic Press.
- Kahe, K. & Walther, K. (1994). *Delikt- und Geschlechtsspezifität subjektiver Theorien zur Entstehung von Kriminalität sowie von Prozessen der Schuldzuschreibung und Strafzumessung*. Unveröffentlichte Diplomarbeit, Institut für Psychologie, TU Braunschweig.
- Kidder, L. & Cohn, E. (1979). Public views of crime and crime prevention. In I. Frieze, D. Baral & J. Carroll (Eds.), *New approaches to social problems*. San Francisco: Jossey-Bass.
- Langer, W. (1994). *Staatsanwälte und Richter: Justizielles Entscheidungsverhalten zwischen Sachzwang und lokaler Justizkultur*. Stuttgart: Enke.
- Oswald, M.E. (1994). *Psychologie des richterlichen Strafens*. Stuttgart: Enke.
- Oswald, M.E. (1997). Richterliche Urteilsbildung. In M. Steller & R. Volbert (Hrsg.), *Psychologie im Strafverfahren*. Bern: Huber.
- Oswald, M.E. & Bilsky, W. (1991). Subjektive Theorien über Kriminalitätsursachen und richterliche Schuldzuschreibung, *Monatsschrift für Kriminologie und Strafrechtsreform*, 74, 129-145.
- Reichert, A. (1999). *Zusammenhang zwischen Ursachenzuschreibung für Kriminalität und Strafhärte*. Unveröffentlichte Diplomarbeit, Westfälische Wilhelms-Universität Münster.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York, Berlin: Springer.
- Wilken, M. (1997). *Deliktsspezifität subjektiver Theorien zur Entstehung von Kriminalität unter Berücksichtigung von Strafhärte und Strafzweck*. Unveröffentlichte Diplomarbeit, Westfälische Wilhelms-Universität Münster.

## Anhang

Tabelle 1: Items zur Erfassung von Ursachenzuschreibungen für Kriminalität

Skala	Item	Wortlaut	D	K
		Zu Straftaten kommt es, weil ...		
Person	6	... die Täter an seelischen Störungen leiden		x
	7	... die Täter sich selbst nicht mögen		x
	9	... die Täter von Zorn auf ihre Opfer erfüllt sind		x
	13	... die Täter an Störungen der Hirnfunktion leiden		x
	14	... die Täter mit Situationen, in denen sie sich provoziert fühlen, nicht umgehen können		x
	20	... die Täter sich in einer materiellen Notlage befinden	x	
	21	... die Täter nur über geringe geistige Fähigkeiten verfügen		x
	23	... kriminelles Handeln einfach in der Natur des Menschen liegt	x	x
	24	... die Täter die Folgen ihres Tuns für sich selbst nicht bedenken	x	x
	25	... die Täter drogenabhängig sind	x	x
	27	... die Täter Menschen sind, die Nervenkitzel und Abenteuer suchen	x	
	28	... die Täter habgierige Menschen sind	x	
	30	... die Täter glauben, sie seien im Recht	x	x
34	... im Erbgut der Täter eine kriminelle Veranlagung enthalten ist	x	x	
Interaktion	1	... die Eltern der Täter es nicht geschafft haben, ihren Kindern den Unterschied zwischen gut und böse beizubringen	x	x
	3	... die Täter sich an Gruppen orientieren, die die gesellschaftlichen Normen nicht akzeptieren	x	x
	5	... die Täter bei ihren Eltern häufig aggressives Verhalten beobachten oder beobachtet haben		x
	11	... die Eltern der Täter es versäumt haben, ihre Erziehung an festen Prinzipien auszurichten	x	x
	16	... die Täter ohne Vater aufwachsen mußten	x	x
	19	... die Eltern der Täter ihren Kindern nicht genügend moralische Werte vermittelt haben	x	x
	26	... die Eltern der Täter ihren Kindern ein schlechtes Beispiel geben oder gegeben haben	x	x
	29	... die Eltern ihren Kindern nicht genügend religiöse Werte vermittelt haben	x	x
	32	... die Täter in ihrer Kindheit einen oder beide Eltern verloren haben	x	x
35	... die Täter in ihrer Kindheit nicht genügend väterliche Zuwendung erfahren haben	x	x	
38	... die Täter in ihrer Kindheit von den Eltern vernachlässigt worden sind	x	x	
soziales System	2	... die Massenmedien den Eindruck erwecken, strafbare Handlungen seien etwas ganz Normales	x	x
	4	... die Polizei personell nicht hinreichend ausgestattet ist	x	x
	8	... viele Menschen in Armut leben	x	
	10	... die Kirche viel von ihrem Einfluß verloren hat	x	x
	12	... das Risiko, bestraft zu werden, zu gering ist	x	x
	15	... Recht und Gesetz in unserer Gesellschaft zu wenig gelten	x	x
	17	... in der Gesellschaft das Geld so einen hohen Stellenwert hat	x	
	18	... der Strafvollzug heute seine abschreckende Wirkung verloren hat	x	x
	22	... es in der Gesellschaft zu wenig gemeinsame Werte und Normen gibt	x	x
	31	... der gesellschaftliche Wohlstand ungleich verteilt ist	x	
	33	... der Staat nicht hart genug gegen Straftäter vorgeht	x	x
	36	... die Täter häufig Gewaltdarstellungen in den Medien konsumieren		x
	37	... wir in einer Gesellschaft leben, in der fast alles erlaubt ist	x	x
39	... straffällig gewordene Ausländer nicht in ausreichendem Maße abgeschoben werden	x	x	

Anm.: Item = in der vorliegenden Untersuchung verwendete Itembezeichnung; D = zur Erklärung von Diebstahlsdelikten geeignet; K = zur Erklärung von Körperverletzungsdelikten geeignet

Tabelle 2: Strafhärteskalisierung der Sanktionsmöglichkeiten

Sanktion	M	SD
Ermahnung mit Einstellung	1.00	0.00
Geldstrafe, 15 Tagessätze	3.34	1.95
Geldstrafe, 30 Tagessätze	4.50	2.02
Geldstrafe, 90 Tagessätze	6.16	2.22
Geldstrafe, 180 Tagessätze	7.67	2.48
3 Monate Freiheitsstrafe <u>mit</u> Bewährung	4.07	1.85
6 Monate Freiheitsstrafe <u>mit</u> Bewährung	5.49	1.96
1 Jahr Freiheitsstrafe <u>mit</u> Bewährung	7.09	2.21
2 Jahre Freiheitsstrafe <u>mit</u> Bewährung	8.50	2.47
1 Monat Freiheitsstrafe <u>ohne</u> Bewährung	8.67	2.04
3 Monate Freiheitsstrafe <u>ohne</u> Bewährung	10.15	1.48
6 Monate Freiheitsstrafe <u>ohne</u> Bewährung	11.58	0.98
1 Jahr Freiheitsstrafe <u>ohne</u> Bewährung	12.80	0.67

Anm.: Skala von 1 (geringste Strafhärt) bis 13 (höchste Strafhärt)

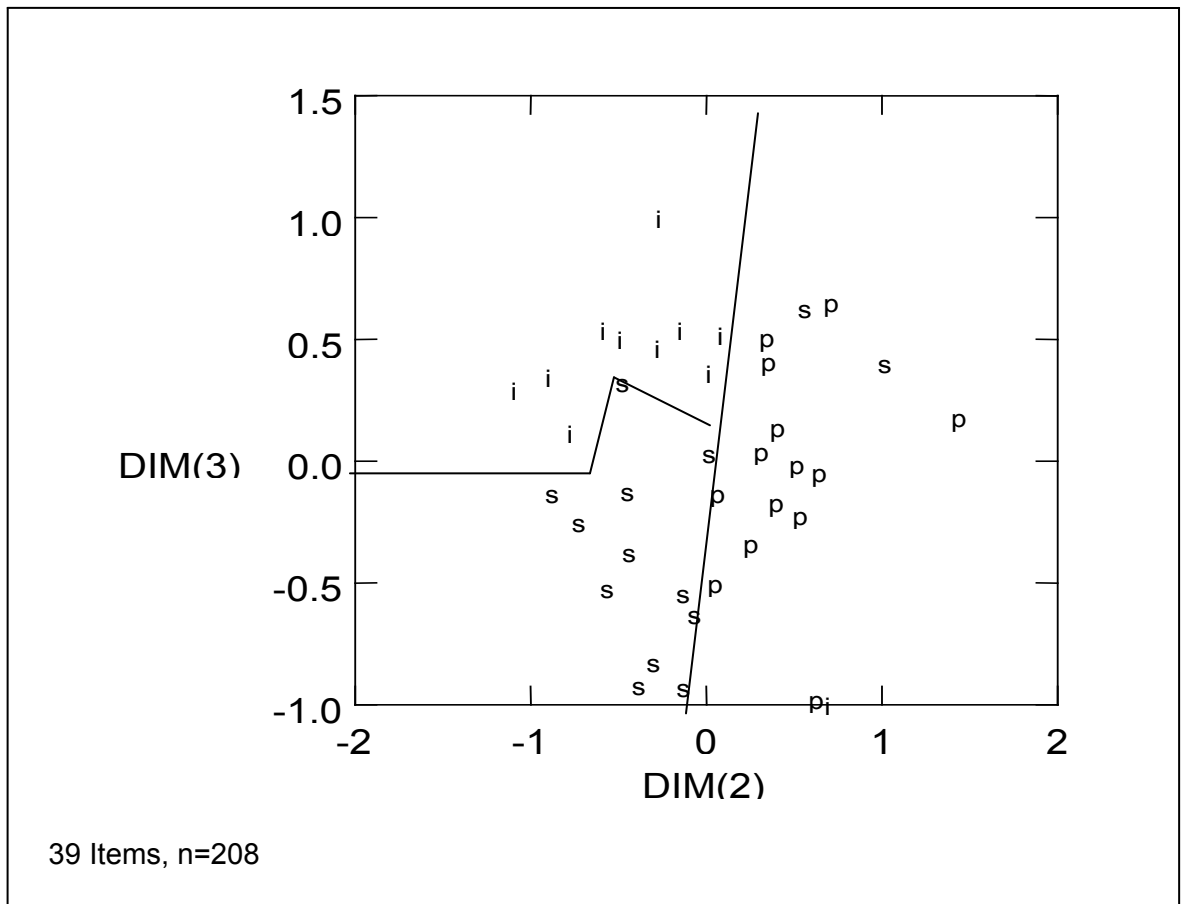


Abbildung 1: Multidimensionale Skalierung der 39 Items zur Ursachenzuschreibung  
 N=208, p = Person, i = Interaktion, s = soziales System

<sup>1</sup> Die der MDS zugrundeliegende Datenmatrix wurde den Autoren von Herrn Görden freundlicherweise zur Verfügung gestellt.

# Strukturierte Diagnostik psychischer Störungen bei Sexualstraftätern im Maßregelvollzug

Anja Schmidt<sup>1</sup>, Jürgen Hoyer<sup>1</sup> & Heike Kunst<sup>2</sup>

<sup>1</sup> Technische Universität Dresden, <sup>2</sup> Justizvollzugsanstalt Kassel II

## Einleitung

Bei der Diskussion um das Problem der Sexualdelinquenz stehen zumeist Fragen nach der Art der Delikte, der geeigneten Unterbringung der Sexualstraftäter, dem Schutz der Bevölkerung vor dem Täter, dem aus Sicherheitsgründen angemessenen Vorgehen bei der Lockerung des Strafvollzuges und Fragen einer geeigneten Therapie im Vordergrund der Betrachtung (vgl. z.B. Schüler-Springorum et al., 1996). Dies sind zweifellos wichtige Gesichtspunkte. Doch fallen bei Sichtung der gegenwärtigen Literatur zur Sexualdelinquenz zweierlei Aspekte auf: (I) Die Forschung zur Sexualdelinquenz zeichnet sich auf der einen Seite durch eine deliktorientierte Betrachtung der Sexualstraftäter aus. Sexualstraftäter werden in der Regel als Vergewaltiger und Missbrauchstäter bezeichnet; psychologische oder psychopathologische Unterschiede werden seltener berücksichtigt. (II) Die Untersuchung psychischer Störungen beschränkt sich weitestgehend auf schwere psychiatrische Störungen, die die Steuerungs- und Einsichtsfähigkeit einer Person beeinträchtigen können, z.B. Schizophrenie, Minderbegabung oder bestimmte Persönlichkeitsstörungen. Die Rolle anderer psychopathologischer Störungen wird hingegen vernachlässigt. Daraus ergeben sich folgende Fragen:

- Ist das Spektrum psychischer Störungen bei Sexualstraftätern tatsächlich auf schwere psychiatrische Störungen begrenzt? Aktuelle wissenschaftliche Untersuchungen an Sexualstraftätern zeigen, dass Sexualstraftäter neben schweren psychiatrischen Störungen auch klinische Störungen aufweisen, die nicht primär die Steuerungs- und Einsichtsfähigkeit einer Person beeinträchtigen (Berner, Berger, Guitierrez, Jordan & Berger, 1992; McElroy et al., 1999; Raymond, Coleman, Ohlerking, Christenson & Miner, 1999), aber zumindest therapierelevant sind. Zwar muss nicht jede psychische Störung therapierbar sein, aber das Wissen um das Vorliegen psychischer Störungen sollte einen wesentlichen Einfluss im Hinblick auf Lockerungsentscheidungen und Rückfallprävention haben.
- Haben Subgruppen von Sexualstraftätern gleiche oder unterschiedliche psychische Störungen? Die Studien von Hoyer, Kunst und Schmidt (in press), McElroy et al. (1999) und Raymond et al. (1999) weisen darauf hin, dass sich Subgruppen von Sexualstraftätern, wie zum Beispiel Paraphile im Hinblick auf vorliegende komorbide psychische Störungen durchaus von nicht-paraphilen Sexualstraftätern unterscheiden.

Ziel des vorliegenden Beitrages ist es: (1) einen Ansatz zur störungs- anstelle einer deliktbezogenen Klassifikation von Sexualstraftätern im Maßregelvollzug aufzuzeigen und (2) Aktual- und Lebenszeitprävalenzen ausgewählter Achse-I-Störungen und Persönlichkeitsstörungen bei Sexualstraftätern im Maßregelvollzug darzustellen. Dabei wurden in der vorliegenden Studie ausschließlich im Maßregelvollzug untergebrachte Sexualstraftäter mit Hands-on-Sexualdelikten (Körperkontakt umfassend) untersucht. Die Ergebnisse dieser Untersuchung sind damit für die genannte Stichprobe spezifisch und nicht auf die Gesamtpopulation der Sexualstraftäter verallgemeinerbar.

## Sexualdelinquenz: Deliktbezogene vs. störungsbezogene Perspektive

Über die Klassifikation von Sexualstraftätern besteht bisher in der Wissenschaft noch wenig Einigkeit. In der Regel werden Sexualstraftäter anhand ihrer begangenen Delikte in Vergewaltiger (Rapists) und Missbrauchstäter (Child Molesters) unterteilt (z.B. Marshall, Laws & Barbaree, 1990; Rehder, 1996 a, b). Hoyer, Borchard und Kunst (2000, S. 8) sprechen in diesem Zusammenhang von einer „kriminologischen Forschungsperspektive“, die relevant ist, wenn man Aussagen über die Gesamtheit der Täter machen will. Es werden Delinquente untersucht und nicht Deviante; es wird eine soziologische Abweichung – Kriminalität – untersucht, keine psychopathologische. Beschränkt man sich also auf die Untersuchung von Sexualstraftätern mit psychischen Störungen, so ist ein Perspektivenwechsel angezeigt, da es sich bei den Kategorien – Vergewaltiger und Missbrauchstäter – nicht um psychologische Variablen oder theoretisch begründete, diagnostisch valide Konstrukte handelt (Hoyer, 2001).

In der Literatur existieren verschiedene Lösungsversuche, Sexualstraftäter zu klassifizieren. Einige seien an dieser Stelle exemplarisch skizziert: Howells (1981) sowie Groth, Hobson und Gary (1982) unterschieden Sexualstraftäter mit stabiler erotischer Präferenz für Kinder und sexuellen Missbrauch von Kindern als Ersatz für nicht vorhandene erwachsene Partner. Knight und Prentky (1990) identifizierten eine große Anzahl von Subtypen des „Vergewaltigers“ oder des „Missbrauchstäters“. Dabei unterteilten die Autoren Missbrauchstäter nach dem Ausmaß der sexuellen Präferenz für Kinder (Degree of Fixation), der Anzahl und der Art der Kontakte zu Kindern (Amount of Contact), Motiven der Kontaktierung und der sozialen Kompetenz; Vergewaltiger wurden bezüglich vorliegender sadistischer Phantasien (sexual type), dem Gewaltanteil gegenüber Frauen (vindictive type) und dem Vorliegen impulskontrollgestörten Verhaltens während der Taten (opportunistic type und pervasively type) unterschieden. Rehder (1996 a, b) ermittelte sechs Klassen von Vergewaltigern und vier Klassen von Missbrauchstätern. Gemeinsam ist diesen Klassifikationsansätzen die primäre Betrachtung der Delikte. In weiteren Ansätzen bestimmen Merkmale des Opfers die Zugehörigkeit des Täters zu einer Gruppe. So wurden Missbrauchstäter unter anderem danach unterschieden, ob es sich um ein hetero- oder ein homosexuelles, ein inzestuöses oder ein nicht-inzestuöses Delikt handelt.

Die Vielzahl der Klassifikationsversuche hat der Heterogenität von Untersuchungsgruppen, den zum Teil schwer objektivierbaren Kategorien und den daraus resultierenden widersprüchlichen Ergebnissen bisher kaum entgegenwirken können, so dass Okami und Goldberg (1992, S.302) zusammenfassend von einem „definitorischen und diagnostischen Chaos“ sprechen. Möglicherweise können die Probleme mangelnder Homogenität von Tätergruppen und eingeschränkter Vergleichbarkeit zwischen Studien durch die Verwendung diagnostischer Kategorien verringert werden. In der vorliegenden Studie wurden in Anlehnung an Hoyer, Kunst, Borchard und Stangier (1999) störungsbezogene Kategorien zur Klassifikation der untersuchten Sexualstraftäter herangezogen. Dabei wurde zum einen die im DSM IV (1996) beschriebene Diagnosengruppe der Paraphilien (DSM IV 302.xx) verwendet. Die Hauptmerkmale der Paraphilien sind dabei:

- wiederkehrende intensive sexuell erregende Phantasien über einen Zeitraum von mindestens sechs Monaten
- sexuell dranghafte Bedürfnisse oder Verhaltensweisen, die sich im allgemeinen auf nichtmenschliche Objekte, das Leiden oder die Demütigung der eigenen Person oder des Partners, Kinder (...) beziehen
- die Phantasien oder Verhaltensweisen führen in klinisch bedeutsamer Weise zu Leiden oder Beeinträchtigung.



Zum anderen wurden auf der Grundlage der Überlegung, dass nicht alle Täter eine Paraphilie aufweisen, die Störungen der Impulskontrolle als weitere psychische Störungen erachtet, die das Risiko gefährlicher Straftaten erhöhen. Berner und Karlick-Bolten haben bereits 1985 vorgeschlagen, der Gruppe der Paraphilien eine Gruppe der sogenannten impulskontrollgestörten Täter gegenüber zu stellen. Diese Sexualstraftäter lassen sich dadurch kennzeichnen, dass sie sexuelle und aggressive Impulse nicht angemessen kontrollieren können (vgl. Barbaree, 1990). Gewalt, Nichteinwilligung und Angst des Opfers führen nicht zu einer Hemmung der sexuellen Erregung (Barbaree, Marshall & Lanthier, 1979; Quinsey, Chaplin & Upfold, 1984). In der Tatsituation stehen keine hinreichenden kognitiven Fähigkeiten zur Verfügung, um sich von intensiven Affekten, wie Wut und Hilflosigkeit, zu distanzieren (Hoyndorf, Reinhold & Christmann, 1995). Die Sexualstraftäter weisen in der Regel eine heterogene Deliktbiographie auf (Wulfert, Greenway & Dougher, 1996). Die Sexualdelikte werden nicht, wie bei Paraphilien, aufgrund sexuell abweichender Phantasien begangen, sondern situationsabhängig als Folge eines gravierenden Verstärkerverlustes. Erst mit der fremdgefährdenden Handlung findet der Spannungszustand ein abruptes Ende. Die diagnostischen Merkmale der Störungen der Impulskontrolle lassen sich wie folgt zusammenfassen:

- ansteigendes Gefühl von Spannung vor der Handlung
- Unfähigkeit, dem Impuls zur Handlung zu widerstehen (bzw. mangelnde Erregungshemmung)
- Erleben von Befriedigung, Vergnügen oder Erleichterung während der Durchführung der Tat.

Für die vorliegende Arbeit ist die Restkategorie „Störung der Impulskontrolle nicht andernorts klassifiziert“ (DSM IV 312.30) relevant, da im DSM IV keine eigene Kategorie für Impulskontrollstörungen mit sexueller Aggression vorgesehen ist. In einer ersten Studie an 72 Sexualstraftätern im Maßregelvollzug erwiesen sich die von Hoyer et al. (1999) vorgeschlagenen psychodiagnostischen Kategorien zur Klassifikation der Sexualstraftäter im Maßregelvollzug als reliabel und valide.

### **Achse-I-Störungen bei Sexualstraftätern**

Anhaltspunkte für Achse-I-Störungen bei Sexualstraftätern lieferten unter anderem Studien von Rader (1977), Segal und Marshall (1985), Overholser und Beck (1986), Chantry und Craig (1994), Hayashino, Wurtele und Klebe (1995) und Hoyer et al. (1999). Sexualstraftäter beschrieben sich darin zum Beispiel als sozial ängstlich oder depressiv und gaben den Missbrauch von Alkohol an. Die genannten Studien verwendeten Selbstbeschreibungsverfahren zur Untersuchung einzelner Störungen. Dazu ist einschränkend festzuhalten, dass der Einsatz von Fragebögen lediglich die Erfassung subklinischer Ausprägungen psychischer Störungen ermöglicht. Fragebögen sind hinsichtlich der diagnostischen Kriterien einer psychischen Störung begrenzt spezifisch. Das bedeutet: Sie erlauben ein Screening, aber keine Diagnostik psychischer Störungen (Stangier & Heidenreich, 1997). In den vorgenannten Studien wurden Sexualstraftäter, mit Ausnahme der Arbeit von Hoyer et al. (1999), deliktbezogen kategorisiert, so dass mit großer Wahrscheinlichkeit ein nicht bestimmbarer Anteil paraphiler und nicht-paraphiler Sexualstraftäter in den Untersuchungsgruppen enthalten war. Studien, die psychische Störungen unter Verwendung der diagnostischen Kriterien des DSM III bzw. des DSM IV vornahmen und die Sexualstraftäter störungsbezogenen Kategorien zuordneten, geben einen Einblick in das Spektrum psychischer Störungen bei

Sexualstraftätern (Berner & Karlick-Bolten, 1985; McElroy et al., 1999; Raymond et al., 1999). In der Studie von Raymond et al. (1999) wiesen 93% der 45 pädophilen Straftäter eine komorbide Achse-I-Störung auf. Siebenundneunzig Prozent der 36 paraphilen und nicht-paraphilen Sexualstraftäter erfüllten in der Studie von McElroy et al. (1999) mindestens eine weitere Achse-I-Störung; 78% hatten sogar drei oder mehr Achse-I-Störungen. Als hoch prävalente Achse-I-Störungen kristallisierten sich bei den untersuchten Sexualstraftätern die Affektiven Störungen, die Angststörungen und die Substanzbezogenen Störungen heraus. Unter den Angststörungen kam die Sozialphobie am häufigsten vor. Prävalenzraten variierten zwischen 19-31% bei paraphilen Straftätern; bei nicht-paraphilen Straftätern ergab sich eine Prävalenz von 13% (vgl. McElroy et al., 1999; Raymond et al., 1999). Erhöhte soziale Angst bei Sexualstraftätern wird als eine Entstehungsbedingung für Sexualdelinquenz diskutiert (z.B. Stermac, Segal & Gillis, 1990). In verschiedenen Arbeiten wird davon ausgegangen, dass Defizite der sozialen Kompetenz (Segal & Marshall, 1985; Overholser & Beck, 1986; Knight & Prentky, 1990) und Minderbegabung (Heim & Morgner, 1985; Möller & Bier-Weiß, 1995) die Entstehung der Sozialphobie bei Sexualdelinquenten mitbedingen können. Viele Sexualstraftäter waren selbst Opfer sexuellen oder physischen Missbrauchs. Schätzungen variieren hier zwischen 25-58% (vgl. Ford & Linney, 1995; McElroy et al., 1999). Diese Erfahrungen können das Risiko einer Angststörung erhöhen (Stein et al., 1996).

Wissenschaftliche Arbeiten ermittelten bei Sexualstraftätern auch einen häufigen Missbrauch bzw. eine Abhängigkeit von Alkohol und/oder Drogen. Leygraf (1988) wies darauf hin, dass Sexualdelikte mit Gewaltanwendung in 45% der Fälle unter dem Einfluss von Alkohol begangen werden. Die Variable „Alkoholeinfluss während der Tat“ hatte in der Arbeit von Kunst, Hoyer und Borchard (2000) eine diskriminative Funktion: Dreimal mehr impuls Kontrollgestörte Sexualstraftäter als paraphile Straftäter verübten ihre Delikte unter dem Einfluss von Alkohol. Allnutt, Bradford, Greenberg und Curry (1996) zeigten demgegenüber eine hohe Komorbidität mit Alkoholismus bei paraphilen Sexualstraftätern. Auch McElroy et al. (1999) und Raymond et al. (1999) konnten Prävalenzraten zwischen 51-67% der Alkoholabhängigkeit bei paraphilen Sexualstraftätern ermitteln. Unwesentlich geringer war die Prävalenzrate mit 53% bei 15 untersuchten nicht-paraphilen Straftätern. Im Vergleich zur Allgemeinbevölkerung lag die Lebenszeitprävalenz der Alkoholabhängigkeit bei den untersuchten Sexualstraftätern damit um das Zehnfache höher als in der Allgemeinbevölkerung mit 5-8% (DSM IV, 1996). Erklärungen zur Komorbidität zwischen Alkoholeinfluss und Sexualstraftaten gehen von einem multifaktoriellen Gefüge aus: Alkohol mindere Angst und verstärke das Auftreten aggressiver Impulse (Allnutt et al., 1996), Alkohol verhindere adäquate kognitive Prozesse, schränke die Wahrnehmungsfähigkeit ein und diene in belastenden Situationen der positiven Stimmungsinduktion (Comer, 1995). Studien an klinischen – nicht forensischen – Stichproben zeigten eine hohe Komorbidität zwischen Substanzbezogenen Störungen, Angststörungen, z.B. Sozialphobie, und Affektiven Störungen (Heckelman & Scheier, 1995; Stangier & Heidenreich, 1997). Befunden von McElroy et al. (1999) und Raymond et al. (1999) zufolge waren dies auch die häufigsten bei Sexualstraftätern diagnostizierten Achse-I-Störungen.

### **Persönlichkeitsstörungen bei Sexualstraftätern**

Die Prävalenz von Persönlichkeitsstörungen wird bei Delinquenten allgemein hoch eingeschätzt. Sexualstraftäter bilden diesbezüglich keine Ausnahme. Empirisch begründete Schätzungen des Anteils von Persönlichkeitsstörungen in Straftäter-Populationen gehen von

etwa einem Drittel (Leygraf, 1988; Schüler-Springorum et al., 1996) bis zu zwei Drittel aus (Berner et al., 1992; Berger, Berner, Bolterauer, Guitierrez & Berger, 1999). In der Forschung zur Sexualdelinquenz wird bei der Diagnostik der Persönlichkeitsstörungen – wie bei der Diagnostik der Achse-I-Störungen auch – eine Differenzierung hinsichtlich des Schweregrades der Störungen vorgenommen. Dabei werden „schwere“, d.h. die Steuerungsfähigkeit einer Person beeinträchtigende Persönlichkeitsstörungen und die hinsichtlich ihres Einflusses auf die Steuerungsfähigkeit und das Hemmvermögen einer Person „zunächst recht unverdächtigen“ Persönlichkeitsstörungen (Kröber, 1997, S.169) unterschieden. Persönlichkeitsstörungen, denen im Rahmen der psychiatrischen Begutachtung der Sexualstraftäter eine wichtige Bedeutung zukommt, da sie eine Ursache verminderter Schuldfähigkeit darstellen können, sind nach Schüler-Springorum et al. (1996) und Kröber (1997) die Antisoziale Persönlichkeitsstörung, die Borderline Persönlichkeitsstörung, die Narzisstische Persönlichkeitsstörung und die Schizotypische Persönlichkeitsstörung. Diese Aussagen werden auch durch Befunde von Berner und Karlick-Bolten (1985) und McElroy et al. (1999) an paraphilen und nicht-paraphilen Sexualstraftätern gestützt. Nicht minder beachtenswert sind die Befunde von Chantry und Craig (1994), Rehder (1996a) und Raymond et al. (1999), die auch weniger „schwere“ Persönlichkeitsstörungen bei Sexualstraftätern fanden, z.B. die Zwanghafte Persönlichkeitsstörung, die Selbstunsichere Persönlichkeitsstörung und die Passiv-aggressive Persönlichkeitsstörung.

Die Forschungspraxis zur Diagnostik psychischer Störungen bei Sexualstraftätern im Maßregelvollzug lässt sich folgendermaßen zusammenfassen:

- (1) Wenn in den vergangenen Jahren Sexualstraftäter untersucht wurden, dann zumeist hinsichtlich psychischer Störungen, die einen Einfluss auf die Steuerungsfähigkeit einer Person haben können.
- (2) Bisher existieren nur wenige Studien, die umfassend und unter Verwendung der diagnostischen Kriterien des DSM IV psychische Störungen bei Sexualstraftätern mit schweren Hands-on-Delikten untersuchten. Diese Studien zeigen, dass das Spektrum psychischer Störungen *nicht* auf schwere psychiatrische Störungen begrenzt ist.

## **Methode**

In der vorliegenden Studie wurden ausschließlich männliche Sexualstraftäter, die in den Maßregelvollzugseinrichtungen des Sächsischen Krankenhauses für Psychiatrie und Neurologie Arnsdorf und im Landeskrankenhaus Moringen untergebracht waren, untersucht. Die Probanden wurden anhand der diagnostischen Kriterien der Paraphilie (DSM IV 302.xx) und der Impulskontrollstörung nicht näher bezeichnet (n.n.b., DSM IV 312.30) den Untersuchungsgruppen zugeordnet. Dabei wurden in die Stichprobe ausschließlich Sexualstraftäter mit Hands-on-Sexualdelikten aufgenommen. N = 30 Sexualstraftäter wurden als paraphil und N = 25 als impulskontrollgestört klassifiziert (Alter:  $\underline{M}$  = 33.85,  $\underline{SD}$  = 9.75; Unterbringungsdauer in Monaten:  $\underline{M}$  = 52.67,  $\underline{SD}$  = 37.13). Achtzehn der paraphilen Sexualstraftäter waren pädophil, und bei zehn Paraphilen wurde ein Sexueller Sadismus diagnostiziert. Die Interrater-Reliabilität für die Zuordnung der untersuchten Sexualstraftäter zu den diagnostischen Kategorien betrug  $\kappa = .85$ . Die impulskontrollgestörten Sexualstraftäter waren jünger als die paraphilen ( $t_{(53)} = 2.29$ ;  $p = 0.03$ ). Korrelationsanalysen erbrachten keine deutlichen Hinweise auf eine Altersabhängigkeit hinsichtlich der diagnostizierten klinischen Störungen ( $p > 0.05$ , in allen Analysen). Bezüglich der Unterbringungsdauer zeigten sich keine statistisch bedeutsamen Unterschiede zwischen beiden Gruppen ( $t_{(53)} = 0.88$ ;  $p = 0.38$ ). Je ein Drittel der paraphilen und der impulskontrollgestörten Sexualstraftäter hatte keinen

Schulabschluss und jeweils die Hälfte der Sexualstraftäter beider Gruppen hatte keinen Berufsabschluss. Etwa zwei Drittel der Paraphilen und der Impulskontrollgestörten waren ledig. Bezüglich Ausbildungsstand (Schul- / Berufsabschluss) und Familienstand zeigten sich keine statistisch bedeutsamen Unterschiede zwischen beiden Gruppen ( $p > 0.25$ , in allen Analysen).

*Diagnostische Verfahren:* Die Diagnostik ausgewählter Achse-I-Störungen erfolgte unter Verwendung des Diagnostischen Kurz-Interview bei psychischen Störungen (Mini-DIPS, Margraf, 1994). Das Mini-DIPS ist ein strukturiertes Interview, mit dem Achse-I-Störungen entsprechend den Kriterien des DSM IV erfasst werden können. Das Mini-DIPS ermöglicht die Diagnostik aller Angststörungen, der Affektiven Störungen (außer der Zylothymie), der Somatoformen Störungen: Hypochondrie, Somatisierungssyndrom, Konversionssyndrom, Somatoformes Schmerzsyndrom, der Essstörungen: Anorexia Nervosa, Bulimia Nervosa und des Substanzmissbrauch sowie der -abhängigkeit. Desweiteren erlaubt das Mini-DIPS die Erfassung soziodemographischer Daten, ein Psychose-Screening, ein Screening bezüglich der Einnahme von Medikamenten und erfahrener psychotherapeutischer Behandlungen. Einschätzungen zu psychosozialen Faktoren und zum Funktionsniveau können unter Verwendung der Achsen IV und V vorgenommen werden. Das Mini-DIPS ist ein vergleichbar genaues Verfahren wie das DIPS (Margraf, Schneider, Ehlers, DiNardo & Barlow, 1991). Diagnosen können mit beiden Verfahren zuverlässig erfasst werden. Die Interrater-Reliabilität liegt um  $\kappa = .80$  (Margraf, 1994).

Die Diagnostik der Persönlichkeitsstörungen wurde mit dem Strukturierten Klinischen Interview für DSM IV Achse II (SKID-II, Wittchen, Zaudig & Fydrich, 1997) vorgenommen. Das SKID-II ist ein zweistufiges Verfahren bestehend aus einem Fragebogen- und einem Interviewteil, das die Erfassung der im DSM IV operationalisierten Persönlichkeitsstörungen in einer ökonomischen Weise ermöglicht. Das zweistufige Vorgehen berücksichtigt die subjektive Einschätzung des Patienten und die klinische Beurteilung des Diagnostikers. Die Interrater-Reliabilität ist für das Vorliegen überhaupt einer Persönlichkeitsstörung mäßig ( $\kappa = .68$ ). Für die Diagnose spezifischer Persönlichkeitsstörungen schwankt die Übereinstimmung zwischen den Urteilern zwischen  $.41 - .73$ . Die Diagnostik von Persönlichkeitsstörungen kann sich trotz der Verwendung der DSM IV-Kriterien aus verschiedenen Gründen schwierig gestalten: Die Grenzen zwischen Diagnose vs. Nicht-Diagnose sind häufig fließend, die Art der Funktionsstörung ist nicht immer eindeutig, bei den Betroffenen kann nicht in jedem Fall ein Problembewusstsein vorausgesetzt werden, und das statusdiagnostische Vorgehen kann u.U. die vergleichsweise schlechte Interrater-Reliabilität mitbedingen. Die klinischen Diagnosen der vorliegenden Studie wurden durch zwei erfahrene klinische Psychologen supervidiert. Unklare Diagnosen wurden besprochen und schließlich auf der Basis eines Konsensurteils vergeben.

## Ergebnisse

Sechsendneunzig Prozent ( $N = 53$ ) der untersuchten Sexualstraftäter erfüllten zusätzlich zur Paraphilie oder zur nicht näher bezeichneten Störung der Impulskontrolle die Diagnosekriterien für mindestens eine Achse-I-Störung oder eine Persönlichkeitsstörung. Dabei hatten alle Impulskontrollgestörten ( $N = 25$ ) und 93% ( $N = 28$ ) der paraphilen Sexualstraftäter mindestens eine weitere psychische Störung. Eine Achse-I-Störung *und* eine Persönlichkeitsstörung hatten 70% ( $N = 21$ ) der paraphilen und 48% ( $N = 12$ ) der impulskontrollgestörten Straftäter (Chi-Quadrat-Tests:  $p > 0.05$ ).

Die Angststörungen waren in der untersuchten Stichprobe hoch prävalent. Dreiundsiebzig Prozent der Paraphilen ( $N = 22$ ) und 64% der Impulstäter ( $N = 16$ ) wiesen mindestens eine

Angststörung auf. Paraphile hatten sogar doppelt so häufig mindestens zwei Angststörungen (16 Paraphile, 7 Impulstäter). Im Vorliegen mindestens einer Angststörung unterschieden sich beide Gruppen nicht statistisch signifikant ( $p > 0.05$ ); für zwei Angststörungen zeichnete sich mit einem  $p$  von 0.06 ein Unterschied im Bereich einer statistischen Tendenz ab. Unter den Angststörungen wurden die Lebenszeitdiagnosen<sup>1</sup> der Sozialphobie (53% der Paraphilen, 20% der Impulstäter) und der Spezifischen Phobie (43% der Paraphilen, 36% der Impulstäter) am häufigsten ermittelt; gefolgt von der Posttraumatischen Belastungsstörung, die bei 23% der paraphilen und bei 32% der impuls kontrollgestörten Straftäter als Lebenszeitdiagnose vorlag.

Die Sozialphobie kam sowohl als aktuelle Diagnose als auch als Lebenszeitdiagnose bei paraphilen Straftätern etwa dreimal häufiger vor als bei Impulstägern. Im Vergleich zur männlichen Allgemeinbevölkerung, in der die Sozialphobie bei 11% vorkommt (Kessler et al., 1994), wurde die Lebenszeitdiagnose der Sozialphobie bei Paraphilen dieser Stichprobe etwa fünfmal häufiger diagnostiziert. Der Unterschied zwischen Paraphilen und Impulstägern war für die Lebenszeitdiagnose der Sozialphobie auf dem Alpha-Niveau von 0.01 statistisch signifikant ( $\chi^2_{(1)} = 6.42$ ;  $p = 0.01$ , einseitige Prüfung). Paraphile und Impulskontrollgestörte unterschieden sich nicht statistisch bedeutsam bezüglich der Posttraumatischen Belastungsstörung ( $\chi^2_{(1)} = 0.52$ ;  $p > 0.05$ ). Dies galt auch für die aktuelle und die Lebenszeitdiagnose der Spezifischen Phobie (aktuelle Diagnose:  $\chi^2_{(1)} = 0.72$ ,  $p > 0.05$ ; Lebenszeitdiagnose:  $\chi^2_{(1)} = 0.31$ ;  $p > 0.05$ ). Die bei 22 Sexualstraftägern ermittelte Lebenszeitdiagnose der Spezifischen Phobie umfasste bei 59% (N = 13) der Paraphilen und 27% (N = 6) der Impulskontrollgestörten eine Dentalphobie.

Tabelle1: Punkt- und Lebenszeitprävalenzen der Angststörungen Paraphiler (N = 30) und Impulskontrollgestörter (N =25)

Störungen	Aktualdiagnose				Lebenszeitdiagnose			
	Paraphilie		Impulskontrollstörung		Paraphilie		Impulskontrollstörung	
	N	%	N	%	N	%	N	%
Paniksyndrom ohne AG	0	0	0	0	1	3	0	0
Paniksyndrom mit AG	1	3	0	0	1	3	1	4
Agoraphobie	0	0	0	0	1	3	0	0
Sozialphobie	7	23	2	8	16	53	5	20
Spezifische Phobie	9	30	5	20	13	43	9	36
Posttraum. Belastgsstörg.	1	3	3	12	7	23	8	32
Generalisierte Angststörung	0	0	0	0	3	10	3	12

Anmerkung: AG = Agoraphobie

Substanzbezogene Störungen, wie Alkohol-, Drogen- und Medikamentenabhängigkeit, bildeten nach den Angststörungen die zweithäufigste Störungsgruppe unter den untersuchten Achse-I-Störungen. Tabelle 2 verdeutlicht die Lebenszeitprävalenzen der diagnostizierten Substanzbezogenen Störungen bei den untersuchten Sexualstraftägern. Sechsfundfünfzig Prozent (17 Paraphile, 14 Impulskontrollgestörte) der Sexualstraftäter zeigten eine Substanzabhängigkeit von Alkohol, Drogen und/oder Medikamenten als Lebenszeitdiagnose. Beide Gruppen unterschieden sich nicht statistisch bedeutsam im Vorliegen Substanzbezogener Störungen ( $\chi^2_{(1)} = 0.002$ ;  $p = 0.96$ ). Alkoholabhängigkeit kam in beiden Gruppen gleich häufig vor und lag mit einer Prävalenz zwischen 43-52% etwa um das Achtfache höher als in der Allgemeinbevölkerung mit 5-8% (DSM IV, 1996). Die Kriterien

für einen Substanzmissbrauch von Alkohol und/oder Drogen erfüllten 9% (vier Paraphile, ein Impulstäter) der Sexualstraftäter.

Tabelle 2: Lebenszeitprävalenzen der Substanzbezogenen Störungen

Lebenszeitdiagnosen	Paraphilie (N = 30)		Impulskontrollstörung (N = 25)	
	N	%	N	%
Alkoholmissbrauch	3	10	1	4
Alkoholabhängigkeit	13	43	13	52
Drogenmissbrauch	1	3	0	0
Drogenabhängigkeit	6	20	2	8
Medikamentenabhängigkeit <sup>1</sup>	1	3	0	0

Statistische Berechnungen zum Alkoholmissbrauch, zur Drogenproblematik und zur Medikamentenabhängigkeit konnten aufgrund zu geringer Fallzahlen nicht vorgenommen werden.

Affektive Störungen kamen bei Impulstägern (N = 21) etwa doppelt so häufig vor wie bei Paraphilen (N = 11); der Unterschied war diesbezüglich zwischen beiden Gruppen statistisch signifikant ( $\chi^2_{(1)} = 10.41$ ;  $p = 0.001$ ). In der untersuchten Stichprobe wurden das Schwere Depressive Syndrom (SDS) und die Dysthymie diagnostiziert. Keiner der Probanden wies eine Bipolare Störung auf. Die Affektiven Störungen wurden hauptsächlich als Lebenszeitdiagnose ermittelt. Aktuelle Beeinträchtigungen durch Symptome des Schweren Depressiven Syndrom bzw. der Dysthymie waren, wie Tabelle 3 zeigt, selten.

Tabelle 3: Punkt- und Lebenszeitprävalenzen der Affektiven Störungen

Störungen	Aktualdiagnose				Lebenszeitdiagnose			
	Paraphilie <sup>1</sup>		Impulskontrollstörung <sup>2</sup>		Paraphilie <sup>1</sup>		Impulskontrollstörung <sup>2</sup>	
	N	%	N	%	N	%	N	%
SDS	0	0	1	4	9	30	20	80
Dysthymie	0	0	0	0	2	7	1	4

Anmerkung: <sup>1</sup> Paraphile N = 30; <sup>2</sup> Impulskontrollgestörte N = 25; SDS = Schweres Depressives Syndrom

Das Schwere Depressive Syndrom wurde bei 53% (N = 29) der Sexualstraftäter diagnostiziert und erwies sich damit als häufigste Achse-I-Lebenszeitdiagnose in der untersuchten Stichprobe. Impulstäter erfüllten signifikant häufiger als Paraphile die Kriterien für die Lebenszeitdiagnose des Schweren Depressiven Syndrom ( $\chi^2_{(1)} = 13.68$ ;  $p = 0.00$ ). Bei acht Paraphilen und 12 Impulstägern entwickelte sich das Schwere Depressive Syndrom während der Unterbringung.

In der untersuchten Stichprobe konnte bei 70% (N = 21) der Paraphilen und bei 48% (N = 12) der Impulskontrollgestörten mindestens eine Persönlichkeitsstörung diagnostiziert werden. Der Unterschied zwischen beiden Gruppen lag diesbezüglich im Bereich einer statistischen Tendenz ( $\chi^2_{(1)} = 2.80$ ,  $p = 0.097$ ).

In der untersuchten Stichprobe wurden Cluster B–Persönlichkeitsstörungen (Antisoziale, Borderline, Narzisstische Persönlichkeitsstörung) und Cluster C–Persönlichkeitsstörungen (Selbstunsichere, Dependente, Zwanghafte Persönlichkeitsstörung) diagnostiziert. Cluster A–Persönlichkeitsstörungen (Schizotypische, Schizoide, Paranoide Persönlichkeitsstörung)

wurden nicht diagnostiziert. Diese Persönlichkeitsstörungen beinhalten u.a. Symptome der Schizophrenie. Probanden mit einer Schizophrenie wurden nicht in die Stichprobe aufgenommen. Vereinzelt wurden Symptome der Paranoiden oder der Schizoiden Persönlichkeitsstörung bei den untersuchten Probanden ermittelt. Das Stellen einer Diagnose war nicht gerechtfertigt. Die im Anhang B des DSM IV beschriebenen und mit dem SKID-II erfassbaren Diagnosen der Passiv-aggressiven und der Depressiven Persönlichkeitsstörung wurden ebenfalls nicht gestellt. Die Symptome der Depressiven Persönlichkeitsstörung konnten besser durch die zu den Achse-I-Störungen gehörenden Affektiven Störungen erklärt werden.

In beiden Sexualstraftäter-Gruppen waren Cluster B-Persönlichkeitsstörungen, mit Ausnahme der nicht diagnostizierten Histrionischen Persönlichkeitsstörung, und Cluster C-Persönlichkeitsstörungen hoch prävalent. Siebenundvierzig Prozent (N = 14) der Paraphilen und 40% (N = 10) der Impulstätter wiesen mindestens eine Cluster B-Persönlichkeitsstörung auf. Cluster C-Persönlichkeitsstörungen wurden bei 40% (N = 12) der Paraphilen und 20% (N = 5) der Impulskontrollgestörten diagnostiziert. Der Unterschied zwischen beiden Gruppen war bezüglich der Cluster B-Persönlichkeitsstörungen nicht statistisch signifikant ( $\chi^2_{(1)} = 0.29, p = 0.86$ ). Demgegenüber unterschieden sich beide Gruppen hinsichtlich der Cluster C-Persönlichkeitsstörungen auf dem Alpha-Niveau von 5% statistisch signifikant ( $\chi^2_{(1)} = 6.94, p = 0.03$ ). Im einzelnen wurden die Antisoziale Persönlichkeitsstörung, die Selbstunsichere Persönlichkeitsstörung und die Borderline Persönlichkeitsstörung in der vorliegenden Studie am häufigsten diagnostiziert. Die Antisoziale Persönlichkeitsstörung wurde bei 37% (N = 11) der Paraphilen und bei 32% (N = 8) der Impulskontrollgestörten diagnostiziert. Die Selbstunsichere Persönlichkeitsstörung kam mit 30% (N = 9) bei den paraphilen Sexualstraftätern etwa doppelt so häufig vor wie bei den impuls kontrollgestörten Sexualstraftätern mit 16% (N = 4). Beide Gruppen unterschieden sich nicht statistisch signifikant in der Prävalenz der Antisozialen und der Selbstunsicheren Persönlichkeitsstörung ( $p > 0.20$ ). Als dritthäufigste Persönlichkeitsstörung kam die Borderline Persönlichkeitsstörung vor (5 Paraphile / 3 Impulskontrollgestörte). Eine statistische Berechnung konnte aufgrund der niedrigen Fallzahlen nicht vorgenommen werden. Abbildung 1 fasst die Lebenszeitprävalenzen der diagnostizierten Persönlichkeitsstörungen bei Paraphilen und Impulskontrollgestörten zusammen.

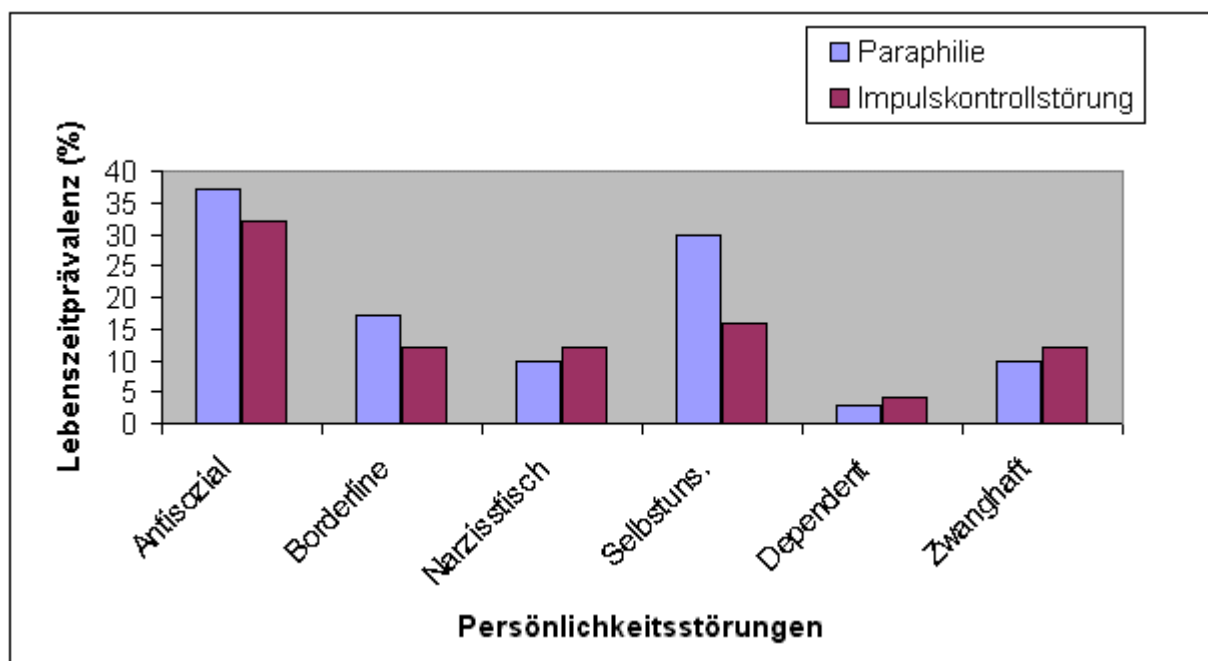


Abb. 1: Lebenszeitprävalenzen der Persönlichkeitsstörungen bei Sexualstraftätern im Maßregelvollzug

## Diskussion

Die Befunde der vorliegenden Arbeit zeigen, dass Sexualstraftäter mit Hands-on-Delikten zahlreiche Achse-I-Störungen und Persönlichkeitsstörungen aufweisen, die nicht unmittelbar die Einsichts- und/oder Steuerungsfähigkeit einer Person beeinträchtigen müssen – nämlich Angststörungen, Affektive Störungen und Alkohol- oder Drogenabhängigkeit. Die bisherige Forschungs- und Gutachtenpraxis vermittelte den Eindruck, bei Sexualstraftätern lägen nur schwere psychische Störungen – z.B. Schizophrenien, Minderbegabung – vor.

Die Angststörungen bildeten in der untersuchten Stichprobe die häufigste Störungsgruppe, gefolgt von den Substanzbezogenen Störungen. Die Sozialphobie wurde in Übereinstimmung mit Befunden von McElroy et al. (1999) und Raymond et al. (1999) als häufigste Angststörung bei Paraphilen diagnostiziert. Paraphile leiden bzw. litten auch häufiger als Impulstäter unter einer Dentalphobie als Form der Spezifischen Phobie. Die häufigste Achse-I-Störung bei impulskontrollgestörten Sexualstraftätern war die Lebenszeitdiagnose des Schweren Depressiven Syndrom. Dieses entwickelte sich bei der Mehrzahl der Paraphilen und Impulskontrollgestörten im Verlauf der Unterbringung. Möglicherweise ist die Entwicklung einer Depression während der Unterbringung das Resultat der erlebten Konsequenzen oder auch in einigen Fällen der Verantwortungsübernahme in Bezug auf das begangene Delikt.

Alkoholabhängigkeit wurde, entgegen der Annahme von Kunst, Hoyer und Borchard (2000), in beiden Subgruppen gleich häufig diagnostiziert. Die Autoren hatten vermutet, dass Alkoholmissbrauch bei Impulstätern wahrscheinlicher sei, da Paraphile Sexualdelikte infolge ihrer sexuellen Präferenz begingen – *nicht* aufgrund von Enthemmung durch Alkohol. Auch im Gegensatz zu den Ausführungen von Allnutt et al. (1996), die ein Alkoholproblem im besonderen bei Paraphilen fanden, stellen sich Alkoholmissbrauch bzw. -abhängigkeit in dieser Studie als klinische Störungen bei den untersuchten Sexualstraftätern *insgesamt* dar. Die untersuchten Sexualstraftäter litten aktuell, d.h. zum Untersuchungszeitpunkt, und zu einem früheren Zeitpunkt – vor oder während der Unterbringung – unter Angststörungen und Affektiven Störungen (vgl. Tab. 1 und 3). Dies unterstreicht die Bedeutung einer umfassenden Diagnostik, die nicht nur die aktuelle Beeinträchtigung durch psychische Störungen berücksichtigt, sondern im Hinblick auf eine erfolgreiche Therapie und Rückfallprävention auch prüfen sollte, ob zu einem früheren Zeitpunkt Symptome einer psychischen Störung vorhanden waren. Auch wenn während der Unterbringung die Symptome einer Störung für den Patienten nicht beeinträchtigend sind, heißt das nicht, dass die jeweilige Störung beseitigt ist. Vielmehr ist anzunehmen, dass bestimmte angstausslösende oder belastende Ereignisse während der Unterbringung weniger stark auftreten, so dass die Beeinträchtigung in Belastungssituationen ohne therapeutische Einflussnahme durchaus wiederkehren kann (vgl. hierzu auch Hoyer, Borchard & Kunst, 2000).

In Übereinstimmung mit Befunden von McElroy et al. (1999) wurde die Antisoziale Persönlichkeitsstörung auch in dieser Stichprobe als häufigste Persönlichkeitsstörung ermittelt, gefolgt von der Selbstunsicheren Persönlichkeitsstörung (vgl. auch Berner et al., 1992). Paraphile wiesen in dieser Studie die genannten Diagnosen häufiger auf als Impulstäter. Die Antisoziale Persönlichkeitsstörung und die Borderline Persönlichkeitsstörung beinhalten Merkmale der Straftäter, die sich auch in den Delikten selbst widerspiegeln.

Die ermittelten Befunde zu den Achse-I-Störungen und den Persönlichkeitsstörungen bei Sexualstraftätern im Maßregelvollzug weisen darauf hin, dass:

- 1) die Prävalenz der meisten diagnostizierten psychischen Störungen wesentlich höher ist als



in der Allgemeinbevölkerung

- 2) bei Paraphilen und Impulskontrollgestörten eine vergleichbare Anzahl von Angststörungen, Substanzbezogenen Störungen und Cluster B–Persönlichkeitsstörungen diagnostiziert werden konnte; Paraphile häufiger als Impulstäter eine Cluster C–Persönlichkeitsstörung aufwiesen, und Impulstäter stärker als Paraphile unter Affektiven Störungen litten
- 3) sich die untersuchten Sexualstraftäter u.a. in der Art der im einzelnen diagnostizierten Störungen unterschieden.

Dies legt eine therapierelevante Differenzierung beider Subgruppen nahe: Während bei Paraphilen die sexuell deviante Präferenz im Vordergrund der Behandlung stehen sollte, ist bei Impulskontrollgestörten, deren sexuelle Orientierung nicht grundsätzlich gestört sein muss, vielmehr auf die alternative Bewältigung intrapsychischer Spannungssituationen oder persönlicher Konflikte – nicht mehr durch Gewaltausübung – einzugehen (vgl. dazu auch Füllgrabe, 1994).

Sowohl bei Paraphilen als auch bei Impulskontrollgestörten zeichnen sich die Behandlung der Alkoholproblematik, der Cluster B–Persönlichkeitsstörungen und das Training sozialer Fertigkeiten als wesentliche Bestandteile der Therapie ab. Dabei sollte sich das Erlernen sozialer Fertigkeiten bei Paraphilen im Besonderen auf die Beseitigung von Defiziten in der Selbstbehauptung beziehen und bei Impulskontrollgestörten auf die Kontrolle von Ärger und aggressivem Verhalten (vgl. Hoyer, Borchard & Kunst, 2000).

## Literatur

- Allnutt, S.H., Bradford, J.M.W., Greenberg, D.M. & Curry, S. (1996). Co-morbidity of alcoholism and the paraphilias. *Journal of Forensic Sciences*, 41, 234-239.
- Barbaree, H.E. (1990). Stimulus control of sexual arousal. In W.L. Marshall, D.R. Laws & H.E. Barbaree (Eds.), *Handbook of Sexual Assault*. (S. 115-141). New York: Plenum Press.
- Barbaree, H.E., Marshall, W.L. & Lanthier, R.D. (1979). Deviant sexual arousal in rapists. *Behaviour Research and Therapy*, 17, 215-222.
- Berger, P., Berner, W., Bolterauer, J., Guitierrez, K. & Berger, K. (1999). Sadistic personality disorder in sex offenders: Relationship to antisocial personality disorder and sexual sadism. *Journal of Personality Disorders*, 13, 175-186.
- Berner, W., Berger, P., Guitierrez, K., Jordan, B. & Berger, K. (1992). The role of personality disorders in the treatment of sex offenders. *Journal of Offender Rehabilitation*, 18, 25-37.
- Berner, W. & Karlick-Bolten, E. (1985). Vergleich zwischen „Paraphilie“ und „sexuellen Impulshandlungen“ bei Sexualdelinquenten. *Forensia*, 5, 157-173.
- Chantry, K. & Craig, R.J. (1994). Psychological screening of sexually violent offenders with the MCMI. *Journal of Clinical Psychology*, 50, 430-435.
- Comer, R.J. (1995). *Klinische Psychologie*. Heidelberg: Spektrum Verlag.
- Ford, M.E. & Linney, J.A. (1995). Comparative analysis of juvenile sexual offenders, violent nonsexual offenders, and status offenders. *Journal of Interpersonal Violence*, 10, 56-70.
- Füllgrabe, U. (1994). „Sexual“verbrechen aus Gewaltmotivation. Basic ID – Die realitätsgerechte Analyse menschlichen Verhaltens. *Kriminalistik*, 48, 241-247.
- Groth, A.N., Hobson, W.F. & Gary, T.S. (1982). The child molester: Clinical observations. *Journal of Social Work and Human Sexuality*, 1, 129-144.
- Hayashino, D.S., Wurtele, S.K. & Klebe, K.J. (1995). Child molester – An examination of cognitive factors. *Journal of Interpersonal Violence*, 10, 106-116.

- Heckelman, L.R. & Schneier, F.R. (1995). Diagnostic Issues. In R.G. Heimberg, M.R. Liebowitz, D.A. Hope & F.R. Schneier (Eds.), *Social phobia – diagnosis, assessment, and treatment*. New York: The Guilford Press.
- Heim, M. & Morgner, J. (1985). Der pädophile Straftäter. *Psychiatrie, Neurologie und Medizinische Psychologie*, 37, 107-112.
- Howells, K. (1981). Adult sexual interest in children: Considerations relevant to theories of aetiology. In M. Cook & K. Howells (Eds.), *Adult sexual interest in children*. London: Academic Press.
- Hoyer, J. (2001). Psychodiagnostische Kategorisierung von gefährlichen Sexualdelinquenten. In J. Hoyer & H. Kunst (Hrsg.), *Psychische Störungen bei Sexualdelinquenten*. Lengerich: Papst.
- Hoyer, J., Borchard, B. & Kunst, H. (2000). Diagnostik und störungsspezifische Therapie bei Sexualdelinquenten mit psychischen Störungen. *Verhaltenstherapie*, 10, 7-15.
- Hoyer, J., Kunst, H. & Schmidt, A. (in press). Social phobia in sex offenders. *The Journal of Nervous and Mental Disease*.
- Hoyer, J., Kunst, H., Borchard, B. & Stangier, U. (1999). Paraphile versus impulskontrollgestörte Sexualstraftäter: Eine psychologisch valide Differenzierung? *Zeitschrift für Klinische Psychologie*, 28, 37-44.
- Hoynorf, S., Reinhold, M. & Christmann, F. (1995). *Behandlung sexueller Störungen – Ätiologie, Diagnostik, Therapie: Sexuelle Dysfunktionen, Missbrauch, Delinquenz*. Weinheim: Beltz/PVU.
- Kessler, R.C., McGonagle, K.A., Zhao, S., Nelson, C.B., Hughes, M., Eshleman, S., Wittchen, H.-U. & Kendler, K.S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. *Archives of General Psychiatry*, 51, 8-19.
- Kleinbaum, D.G., Kupper, L. & Morgenstern, H. (1982). *Epidemiologic research – principles and quantitative methods*. New York: VNR.
- Knight, R.A. & Prentky, R.A. (1990). Classifying sexual offenders – The development and corroboration of taxonomic models. In W.L. Marshall, D.R. Laws & H.E. Barbaree (Eds.), *Handbook of Sexual Assault* (S. 23-52). New York: Plenum Press.
- Kröber, H.-L. (1997). Strafrechtliche Begutachtung von Persönlichkeitsstörungen. *Persönlichkeitsstörungen*, 4, 161-171.
- Kunst, H., Hoyer, J. & Borchard, B. (2000). Alkoholeinfluss bei Sexualdelikten unter differential-diagnostischer Perspektive. *Sucht*, 46, 137-141.
- Leygraf, N. (1988). *Psychisch kranke Straftäter – Epidemiologie und aktuelle Praxis des psychiatrischen Maßregelvollzugs*. Berlin: Springer Verlag.
- Margraf, J. (1994). *Diagnostisches Kurz-Interview bei psychischen Störungen – Mini-DIPS, Handbuch und Interviewleitfaden*. Berlin: Springer Verlag.
- Margraf, J., Schneider, S., Ehlers, A., DiNardo, P. & Barlow, D. (1991). *Diagnostisches Interview bei psychischen Störungen. Interviewleitfaden*. Berlin: Springer Verlag.
- Marshall, W.L., Laws, D.R. & Barbaree, H.E. (Eds.) (1990). *Handbook of sexual assault: Issues, theories and treatment of the offender*. New York: Plenum Press.
- McElroy, S.L., Soutullo, C.A., Taylor, P., Nelson, E.B., Beckman, D.A., Brusman, L.A., Ombaba, J.M., Strakowski, S.M. & Keck, P.E. (1999). Psychiatric features of 36 men convicted of sexual offenses. *Journal of Clinical Psychiatry*, 60, 414-420.
- Möller, A. & Bier-Weiss, I. (1995). Pädophile Straftäter – Versuch einer Tätertypologie. *Psychiatrische Praxis*, 22, 24-26.
- Okami, P. & Goldberg, A. (1992). Personality correlates of pedophilia: Are they reliable indicators? *The Journal of Sex Research*, 29, 297-328.

- Overholser, J.C. & Beck, S. (1986). Multimethod assessment of rapists, child molesters, and three control groups on behavioral and psychological measures. *Journal of Consulting and Clinical Psychology, 54*, 682-687.
- Quinsey, V.L., Chaplin, T.C. & Upfold, D. (1984). Sexual arousal to nonsexual violence and sadomasochistic themes among rapists and non-sex-offenders. *Journal of Consulting and Clinical Psychology, 52*, 651-657.
- Rader, C.M. (1977). MMPI profile types of exposer, rapists, and assaulters in a court services population. *Journal of Consulting and Clinical Psychology, 45*, 61-69.
- Raine, A. (1993). *The psychopathology of crime*. San Diego: Academic Press.
- Raymond, N.C., Coleman, E., Ohlerking, F., Christenson, G.A. & Miner, M. (1999). Psychiatric comorbidity in pedophilic sex offenders. *American Journal of Psychiatry, 156*, 786-788.
- Rehder, U. (1996a). Klassifizierung inhaftierter Sexualdelinquenten – 1. Teil: Wegen Vergewaltigung und sexueller Nötigung Erwachsener Verurteilte. *Monatsschrift für Kriminologie und Strafrechtsreform, 79*, 291-304.
- Rehder, U. (1996b). Klassifizierung inhaftierter Sexualdelinquenten – 2. Teil: Wegen sexuellen Missbrauchs von Kindern Verurteilte. *Monatsschrift für Kriminologie und Strafrechtsreform, 79*, 373-385.
- Saß, H., Wittchen, H.-U. & Zaudig, M. (1996). *Diagnostisches und Statistisches Manual Psychischer Störungen DSM IV*. Göttingen: Hogrefe.
- Schüler-Springorum, H., Berner, W., Cirullies, B., Leygraf, N., Nowara, S., Pfäfflin, F., Schott, M. & Volbert, R. (1996). Sexualstraftäter im Maßregelvollzug. *Monatsschrift für Kriminologie und Strafrechtsreform, 79*, 147-209.
- Segal, Z.V. & Marshall, W.L. (1985). Heterosexual skills in a population of rapists and child molesters. *Journal of Consulting and Clinical Psychology, 53*, 55-63.
- Stangier, U. & Heidenreich, T. (1997). Diagnostik der Sozialen Phobie. *Verhaltenstherapie, 7*, 107-118.
- Stein, M.B., Walker, J.R., Anderson, G., Hazen, A.L., Ross, C.A., Eldridge, G. & Forde, D.R. (1996). Childhood physical and sexual abuse in patients with anxiety disorders and in a community sample. *American Journal of Psychiatry, 153*, 275-277.
- Stermac, L.E., Segal, Z.V. & Gillis, R. (1990). Social and cultural factors in sexual assault. In W.L. Marshall, D.R. Laws & H.E. Barbaree (Eds.), *Handbook of Sexual Assault* (S.143-159). New York: Plenum Press.
- Wittchen, H.-U., Zaudig, M. & Fydrich, T. (1997). *Strukturiertes klinisches Interview für DSM IV Achse-I und Achse-II, Handanweisung*. Göttingen: Hogrefe.
- Wulfert, E., Greenway, D.E. & Dougher, M.J. (1996). A logical functional analysis of reinforcement-based disorders: Alcoholism and pedophilia. *Journal of Consulting and Clinical Psychology, 64*, 1140-1151.

<sup>1</sup> Zum Zeitpunkt der Untersuchung liegen die Symptome einer Störung vor, oder die Symptome sind teilweise oder vollständig remittiert und waren zu einem früheren Zeitpunkt vollständig erfüllt (Kleinbaum, Kupper & Morgenstern, 1982).



# Schuldig bei Verdacht: Wie konfirmatorisches Hypothesentesten zu fälschlicher Beschuldigung wegen sexuellen Kindesmißbrauchs führt

*Stefan Schulz-Hardt<sup>1</sup>, Eberhard Höfer<sup>2</sup> & Günter Köhnken<sup>2</sup>*

*<sup>1</sup>Institut für Psychologie, Universität München, <sup>2</sup>Institut für Psychologie, Universität Kiel*

## **Einleitung**

Im Verlauf der letzten 15 Jahre haben viele spektakuläre Anklagen wegen vermeintlichen sexuellen Kindesmißbrauchs starke öffentliche Beachtung gefunden. Als international bekannte Beispiele zu nennen sind der McMartin und der Kelley Michaels Fall in den USA sowie die Cleveland-Affäre in Großbritannien. In Deutschland haben vor allem der Montessori-Prozeß in Münster sowie die sog. Wormser (bzw. Mainzer) „Kinderschänderprozesse“ zu kontroversen Diskussionen über Ermittlungs- und Strafverfahren zu sexuellem Mißbrauch von Kindern geführt (vgl. hierzu das Sonderheft 1 aus Praxis der Rechtspsychologie, 2000).

Das gemeinsame Charakteristikum dieser Strafverfahren läßt sich wie folgt beschreiben: Es lagen Aussagen von zahlreichen Kindern vor, in denen teilweise sehr detailliert Mißbrauchshandlungen beschrieben wurden; diese Handlungen sollten an ihnen selbst bzw. an anderen Kindern vorgenommen worden sein. Daraufhin erhoben die Justizbehörden, oftmals sogar gestützt auf Glaubwürdigkeitsgutachten psychologischer Sachverständiger, Anklage und nahmen die Beschuldigten zum Teil für viele Monate in Untersuchungshaft. In allen diesen Prozessen wurden die Angeklagten am Ende freigesprochen – zum Teil sogar ausdrücklich wegen erwiesener Unschuld. Zwar kann nicht in jedem Fall ausgeschlossen werden, daß die Vorwürfe einen wahren Kern besitzen (d.h. daß tatsächlich vereinzelt sexueller Mißbrauch stattgefunden hat); gleichwohl ist festzustellen, daß auch in diesen Fällen angebliche Ereignisse berichtet und in den Befragungen verfolgt wurden, die objektiv unmöglich waren (z.B. wurde ein Kind angeblich an den Haaren aufgehängt, obwohl dieses Kind einen Bürstenhaarschnitt besaß) oder durch die polizeilichen Ermittlungen widerlegt worden sind (z.B. das Abhalten „Schwarzer Messen“ in geheimen Kellergewölben). Insgesamt stellt man in diesen Verfahren im Nachhinein eine extreme Diskrepanz fest zwischen den Vorwürfen, von deren Richtigkeit viele Eltern, psychologische Sachverständige und Mitarbeiter von Strafverfolgungsbehörden sowie Kinderschutzorganisationen restlos überzeugt waren, und der empirischen Sachlage – eine Diskrepanz, die einer Erklärung bedarf.

Forschungsarbeiten, die durch diese Mißbrauchsprozesse angeregt wurden, konnten in der Folgezeit nachweisen, daß suggestive Befragungen die Ursache für die konfabulierten Kinderaussagen darstellten (vgl. z.B. Ceci & Bruck, 1993, 1995; Köhnken, 1997; Volbert & Pieters, 1996). Zu diesen Befragungen kam es, weil die Befragenden (Eltern, Polizei, Mitarbeiter von Kinderschutzorganisationen) schon vorher fest davon überzeugt waren, es müsse tatsächlich Mißbrauch stattgefunden haben, den es nun mit allen Mitteln aufzudecken gelte (Köhnken, 1997). Offenkundige Ungereimtheiten wurden in einer Weise ignoriert, die nachträglich – wohl auch für die beteiligten Personen – völlig unverständlich wirkt. Ein Beispiel: Im Montessori-Verfahren berichtete eine Mutter: „Ich habe ihr (der Tochter) davon erzählt und ihr eine Belohnung – sie wünschte sich einen Fahrradhelm – versprochen, wenn sie etwas erzählen würde. Ich habe ihr die Belohnung versprochen, weil sie vorher

monatelang nichts gesagt hatte. Sie sagte dann ganz spontan ...“ Daß die Tochter unter diesen Umständen wahrscheinlich auch von einem nächtlichen Besuch durch Marsmenschen berichtet hätte, kam niemandem in den Sinn.

Die rechtspsychologisch interessante Fragestellung lautet nun: Wie konnten die Beteiligten zu einer Überzeugung gelangen, die durch empirische Evidenz kaum gestützt wurde bzw. vielfach den Fakten sogar nachweislich widersprach? Diese Frage ist theoretisch und praktisch bedeutsam: Wenn man weiß, durch welche Prozesse solche Vorverurteilungen im Kopf der Beteiligten entstehen, kann man gezielt intervenieren und dadurch die Wahrscheinlichkeit verringern, daß die Kinder suggestiv befragt werden. Das wiederum hilft nicht nur, die Rate von Falschbeschuldigungen, die für die Betroffenen verheerend sein können, zu senken. Zugleich wird dadurch auch die Chance erhöht, daß bei tatsächlichem Vorliegen von sexuellem Mißbrauch der Täter rechtskräftig verurteilt werden kann: Da nämlich Kinderaussagen durch suggestive Befragungen entwertet werden und dann nachträglich zumeist der wahre Kern der Erinnerungen nicht mehr festgestellt werden kann, würden suggestive Befragungen von tatsächlichen Mißbrauchsopfern de facto die Täter vor Verurteilung schützen.

Für das Phänomen solcher realitätsfremder „Vorverurteilungen“ haben wir auf der Fachgruppentagung Rechtspsychologie in Erlangen 1999 einen Erklärungsansatz (inkl. erster Pilotuntersuchungen) vorgestellt, der auf sozialpsychologischen Überlegungen zu individueller und kollektiver Informationsverarbeitung aufbaut. Dieser Erklärungsansatz wurde von uns inzwischen weiterentwickelt (siehe Schulz-Hardt & Köhnken, 2000) und durch weitere Untersuchungen überprüft. Der aktuelle Stand soll an dieser Stelle referiert werden.

### **Konfirmatorisches Hypothesentesten beim Verdacht des sexuellen Kindesmissbrauchs**

Wenn Personen zu der Überzeugung kommen, es liege sexueller Mißbrauch vor, obwohl keine klare Evidenz dafür gegeben ist oder die Evidenz sogar dagegen spricht, dann könnte man zunächst geneigt sein zu vermuten, hier liege (im weitesten Sinne) „irrationales“ Denken vor – resultierend daraus, daß Personen das Vorliegen von Mißbrauch eingeredet wurde, daß sie (z.B. aus ideologischen Gründen) unbedingt einen Mißbrauch annehmen wollen, oder daß die Angst um das Wohl der Kinder einen Zusammenbruch des rationalen Denkens bewirkt hat. Die angesprochenen Mißbrauchsverfahren bieten durchaus Anhaltspunkte dafür, daß jeder dieser drei Mechanismen zumindest bei einigen Beteiligten gewirkt haben könnte. Wir wollen auch gar nicht bestreiten, daß solche Mechanismen in diesem Kontext auch eine (verstärkende) Rolle spielen können. Unsere Grundthese ist allerdings weitgehender: Durch allgemeine Gesetzmäßigkeiten beim Testen von sozialen Hypothesen (als welche man einen Mißbrauchsverdacht klassifizieren kann) neigen auch kühl und nüchtern analysierende, unvoreingenommene, normal begabte und um nichts als die Wahrheit bemühte Personen dazu, den Mißbrauchsverdacht fälschlicherweise zu bestätigen. Auch ohne solche Einschränkungen der Rationalität kommt es also zu einem Phänomen, das wir im Titel unseres Beitrags als „Schuldig bei Verdacht“ bezeichnet haben. Wie es dazu kommt, werden wir jetzt erläutern.

#### **Soziales Hypothesentesten**

Unser Ansatz beansprucht nicht, erklären zu können, wie überhaupt ein Verdacht des sexuellen Kindesmissbrauchs aufkommen kann. Die Erfahrung hat gezeigt, daß dies oft aus Zufälligkeiten heraus geschieht, etwa, wenn ein Kind eine merkwürdig klingende Äußerung

macht oder sich auffällig verhält. Daß in Zeiten, in denen sexueller Mißbrauch stark in den Medien und in der öffentlichen Diskussion präsent ist, dann auch Mißbrauch als denkbare Deutungsmuster für solche Auffälligkeiten herangezogen und weiterverfolgt wird, erscheint nicht überraschend. An dieser Stelle setzt unser Erklärungsansatz an, nämlich bei der Frage, warum dieser einmal aufgekommene Verdacht nicht angesichts fehlender Evidenz wieder verworfen wird, sondern sogar immer mehr an subjektiver Überzeugungskraft gewinnt. Wie bereits erwähnt, lokalisieren wir die Ursache hierfür in bestimmten Charakteristika des Testens sozialer Hypothesen.

Unter „sozialen Hypothesen“ versteht man ganz allgemein Annahmen über Zustände oder Beziehungen in der sozialen Umwelt. Beispiele hierfür sind „Hans ist einfühlsam“ oder „Wer als Kind geschlagen wird, schlägt als Erwachsener selbst“. Der Verdacht, Person X könne sexuellen Mißbrauch an den Kindern Y, Z, usw. begangen haben, läßt sich somit auch als eine soziale Hypothese klassifizieren. Die Überprüfung sozialer Hypothesen durch „Laien“ (als Überblick siehe Trope & Liberman, 1996) erfolgt vom Prinzip her nicht anders als beispielsweise die Überprüfung diagnostischer Hypothesen durch psychologische Sachverständige. Die dabei durchlaufenen Schritte sind (erläutert am Beispiel des Mißbrauchsverdachts):

1. *Formulierung von „Wenn-dann“-Regeln für mögliche Evidenz:* Es wird überlegt, welche Kennzeichen in der sozialen Umwelt beobachtbar sein müßten, wenn die Hypothese zutrifft oder nicht zutrifft. Beispiel: „Wenn Mißbrauch vorliegt, dann zeigen die Kinder Verhaltensauffälligkeiten.“

2. *Suche nach Evidenz:* Diese Suche nach den im ersten Schritt formulierten Realkennzeichen kann sowohl im Gedächtnis erfolgen (z.B. frühere Äußerungen und Verhaltensweisen des Kindes) als auch über Informationssuche in der aktuellen sozialen Umwelt (z.B. Beobachtung oder Befragung des Kindes) erfolgen.

3. *Identifizierung der Evidenz:* Oftmals spricht potentielle Evidenz nicht für sich, sondern muß „identifiziert“ werden, d.h. sie wird kategorisiert und interpretiert. Wenn das Kind zum Beispiel schweigt, so kann man das ganz unterschiedlich interpretieren (z.B. im Sinne der Mißbrauchshypothese als Traurigkeit oder vom Täter auferlegtes Schweigegebot, dagegen auch entgegen der Mißbrauchshypothese als Schüchternheit oder Nervosität).

4. *Rückschluß auf die getestete Hypothese:* Es wird überlegt, wie wahrscheinlich das Zutreffen der Hypothese vor dem Hintergrund der vorliegenden Evidenz ist (d.h. wie überzeugt man momentan von der Hypothese ist).

Natürlich können hierbei auch Schritte übersprungen werden, oder man kehrt von einem Schritt zu einem vorhergehenden zurück. Betont werden muß auch, daß Personen sich im Alltag – im Gegensatz etwa zum professionellen Diagnostiker – nicht immer bewußt sind, diese Schritte so zu durchlaufen, und selbst wenn sie es wären, würden sie sie evtl. nicht so benennen. Grundsätzlich kann man aber davon ausgehen, daß eine Person, die einem Mißbrauchsverdacht nachgeht, sich überlegt, woran man sehen kann, ob Mißbrauch gegeben ist oder nicht (Schritt 1), im Gedächtnis und in der Umwelt nach subjektiv relevanten Informationen sucht (Schritt 2), sich überlegt, was diese bedeuten (Schritt 3), und daß daraus dann resultiert, wie sehr sie an das Zutreffen oder Nicht-Zutreffen der Hypothese glaubt (Schritt 4).

#### Konfirmatorische Mechanismen beim sozialen Hypothesentesten

Zu Beginn der Forschung zum sozialen Hypothesentesten (vgl. z.B. Snyder & Swann, 1978) glaubte man, daß soziales Hypothesentesten stets konfirmatorisch ist, d.h. daß die getestete Hypothese immer eine größere Wahrscheinlichkeit besitzt, bestätigt zu werden, als auf

Grundlage der verfügbaren Evidenz gerechtfertigt wäre. Man ging also davon aus, der Prozeß der sozialen Hypothesentestung würde durchweg die getestete Hypothese begünstigen. Dies hat sich als falsch herausgestellt (z.B. Klayman & Ha, 1987; Trope & Liberman, 1996). Allerdings kann – da Menschen keine perfekten Informationsverarbeitungsmaschinen mit unbegrenzter Kapazität und vollständiger Information sind – soziales Hypothesentesten konfirmatorisch verlaufen, und es lassen sich auch sehr genau die Prozesse spezifizieren, durch die es dazu kommt, sowie die Bedingungen identifizieren, unter denen dies geschieht. Wie wir im folgenden ausführen werden, liegen bei der Testung eines Mißbrauchsverdachts eben jene Bedingungen vor, unter denen Hypothesen mit erhöhter Wahrscheinlichkeit fälschlicherweise bestätigt werden.

Die Prozesse, die zu einer Bestätigungsneigung für die getestete Hypothese führen, lassen sich in drei Kategorien unterteilen, nämlich kognitive Prozesse, konfirmatorische Teststrategie und motivationale Prozesse. In Abbildung 1 sind diese Mechanismen im Überblick dargestellt.

*Kognitive Prozesse.* Unterstellen wir einmal, daß die Beteiligten beim Testen des Mißbrauchsverdachts so sorgfältig wie professionelle Diagnostiker vorgehen. (Wenn man das nicht unterstellt, gelten die folgenden Ausführungen nicht weniger – dies soll aber zeigen, daß sorgfältige und intensive Testung nicht zwangsläufig vor fälschlicher Hypothesenbestätigung schützt.) Trope und Liberman (1996) nennen dieses Vorgehen auch „diagnostisches Hypothesentesten“. Beim diagnostischen Hypothesentesten wird die Evidenz sowohl hinsichtlich ihrer Vereinbarkeit mit der zu testenden Hypothese (Mißbrauch liegt vor) als auch hinsichtlich ihrer Vereinbarkeit mit der Alternativhypothese (Mißbrauch liegt nicht vor) beurteilt. Weiterhin wird berücksichtigt, daß Information fehlerhaft sein kann; und auch die unterschiedlichen Basisraten (Mißbrauch ist seltener als Nicht-Mißbrauch) werden einbezogen. Mechanismen, die trotzdem zu einer Bestätigungsneigung der getesteten Hypothese führen, sind (vgl. Trope & Liberman, 1996):

- *Überschätzung der a-priori-Wahrscheinlichkeit der Hypothese:* Durch die vermehrte gedankliche Beschäftigung mit der Hypothese kann sich subjektiv die Basisrate erhöhen (Fiedler, Armbruster, Nickel, Walter & Asbeck, 1996; Hodgins & Zuckerman, 1993). Da der Mißbrauchsverdacht (im Gegensatz zur Alternativhypothese, daß kein Mißbrauch vorliegt) in hohem Maße emotional involvierend ist, löst er besonders starke gedankliche Beschäftigung mit der Hypothese aus.



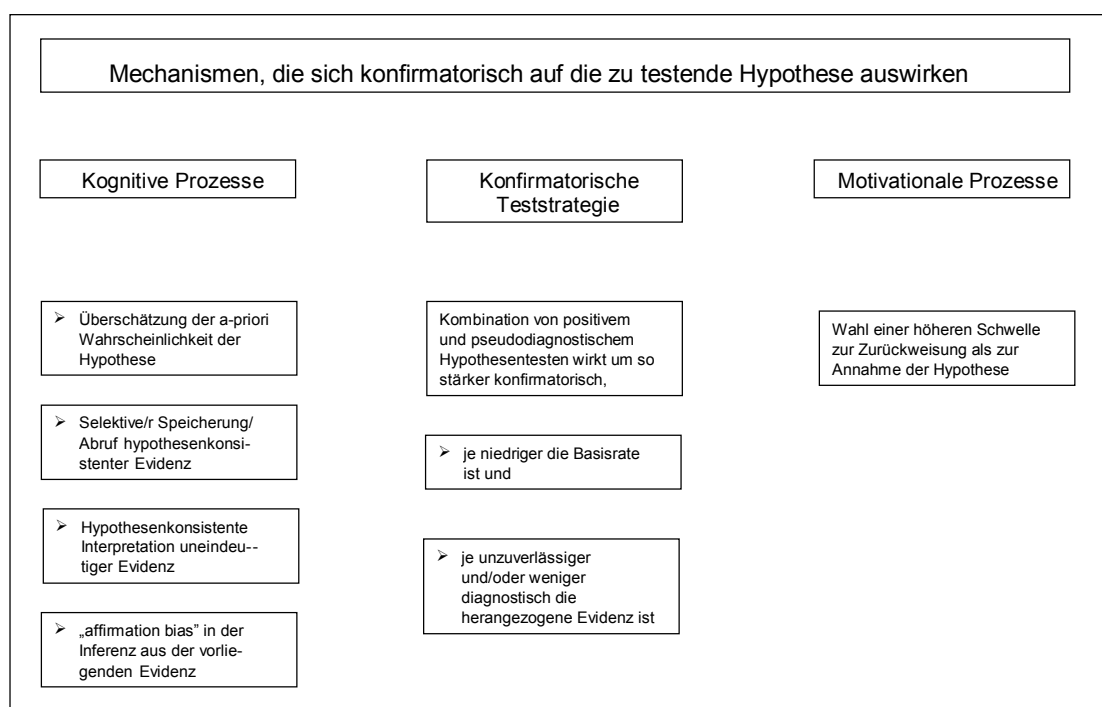


Abb. 1: Selbstbestätigungsmechanismen beim konfirmatorischen Hypothesentesten (Schulz-Hardt & Köhnken, 2000)

- *Selektive Speicherung und selektiver Abruf hypothesenkonsistenter Evidenz*: Die Hypothese kann als Filter wirken, der zur Hypothese passende Informationen bei Informationsaufnahme und -abruf begünstigt. Die soziale Kognitionsforschung hat gezeigt, daß sowohl hypothesenkonsistente als auch hypotheseninkonsistente Informationen besser erinnert werden als Informationen, die für die Hypothese neutral sind (Stangor & McMillan, 1992). Im Kontext des sexuellen Mißbrauchs gibt es aber praktisch keine wirklich hypotheseninkonsistente Evidenz, denn dies müßten ja Informationen sein, die nur oder fast nur im Fall von Nicht-Mißbrauch auftreten. Also führt der obige Filtermechanismus zu einer einseitigen Begünstigung hypothesenkonsistenter Informationen und somit dazu, daß man den Anteil vermeintlich „verdächtiger“ Details systematisch überschätzt.
- *Hypothesenkonsistente Interpretation uneindeutiger Evidenz*: Mehrdeutige Äußerungen oder Verhaltensweisen bedürfen, wie bereits beschrieben, der Interpretation. Die getestete Hypothese stellt nun selbst ein sehr salientes Deutungsmuster für solche uneindeutige Evidenz dar (z.B.: Das Kind schweigt, weil der Täter ihm verboten hat, über den Mißbrauch zu sprechen). Mehrdeutige Evidenz kann daher bevorzugt im Sinne der Hypothese gedeutet werden (Darley & Gross, 1983; Snyder & Uranowitz, 1978; Srull & Wyer, 1989). Je höher der Anteil uneindeutiger Evidenz ist, desto stärker wirkt dieser Mechanismus. Im Fall des sexuellen Mißbrauchsverdachts bei Kindern gibt es, mit Ausnahme medizinischer Nachweise wie z.B. Spermaspuren oder Verletzungen im Genitalbereich, fast überhaupt keine eindeutige Evidenz – so sind beispielsweise Äußerungen und Verhaltensweisen von Kindern häufig mehrdeutig, ebenso beobachtete Interaktionen zwischen dem potentiellen Täter und vermeintlichen Opfern.
- *„affirmation bias“ in der Inferenz aus der vorliegenden Evidenz*: Menschen reagieren beim Testen stärker auf das Auftreten als das Ausbleiben bestimmter Evidenz (Hodgins & Zuckerman, 1993; Newman, Wolff & Hearst, 1980). Das kann bewirken, daß trotz einer

objektiv geringen Anzahl verdächtiger Kennzeichen die Hypothese bestätigt wird. Je weniger klar ausgeprägt die Vorstellungen darüber sind, welche Evidenz bei Zutreffen der Hypothese zu finden sein müßte (Schritt 1 im vorhin dargestellten Ablauf), desto weniger registriert man ausbleibende Evidenz. Solche Vorstellungen sind nun im Bereich von sexuellem Kindesmißbrauch besonders diffus, da (mit Ausnahme der oben erwähnten medizinischen Kennzeichen) selbst Experten sich im Hinblick auf Realkennzeichen von sexuellem Mißbrauch nicht einig sind.

Menschen sind sich solcher Urteilsverzerrungen üblicherweise nicht bewußt (Wilson & Brekke, 1994). Daher werden diese Mechanismen auch durch hohe Motivation und Fähigkeit zu besonders genauer Testung nicht eliminiert; sie können dadurch sogar verstärkt werden. (Auch psychologische Sachverständige können ihnen unterliegen!)

*Konfirmatorische Teststrategie.* Personen wenden beim Testen von Hypothesen (vor dem Hintergrund des vorhin angesprochenen Vier-Schritte-Rasters) unterschiedliche Vorgehensweisen an. Zwei für uns hier wesentliche Dimensionen der Unterscheidung sind

1. Diagnostisches vs. pseudodiagnostisches Testen
2. Positives vs. negatives Testen

Zu 1) Die Merkmale diagnostischen Hypothesentestens wurden bereits oben vorgestellt. Ein solches diagnostisches Testen stellt hohe Anforderungen an Motivation und Fähigkeit der Person; Personen setzen daher oftmals vereinfachte Teststrategien ein (Friedrich, 1993). Dem diagnostischen Testen steht daher ein sogenanntes „pseudodiagnostisches Testen“ gegenüber (Trope & Liberman, 1996): Beim pseudodiagnostischen Testen überprüft man die vorliegende Evidenz nur auf ihre Stimmigkeit mit der zu testenden Hypothese; man vernachlässigt dabei aber drei Dinge, nämlich a) ihre Vereinbarkeit mit der Alternativhypothese, b) die mögliche Fehlerhaftigkeit von Information und c) die Basisrate. (Falls nicht alle drei Komponenten fehlen, erhält man entsprechend eine Zwischenstufe zwischen diagnostischem und pseudodiagnostischem Testen.)

Zu 2) Unter positivem Testen versteht man die Suche nach Evidenz, die man bei Zutreffen der Hypothese erwarten würde (Klayman & Ha, 1987). Die Wenn-dann-Regeln in Schritt 1 sind also alle aus der Kategorie „Wenn die Hypothese stimmt, dann ...“. Dementsprechend werden beim negativen Testen Wenn-dann-Regeln formuliert wie „Wenn die Hypothese falsch ist, dann ...“, d.h. es wird nach Evidenz gesucht, die typisch für das Nicht-Zutreffen der Hypothese ist.

Keine der hier vorgestellten Vorgehensweisen ist per se konfirmatorisch. Wenn man beispielsweise positiv testet, jedoch nur wenig erwartete Evidenz findet, so wird man nicht zur Bestätigung der Hypothese neigen. Jedoch wirkt sich eine bestimmte Kombination der Vorgehensweisen konfirmatorisch aus; konkret gemeint ist hier das pseudodiagnostische positive Testen. In diesem Fall sucht man nach Evidenz, die man als typisch für das Zutreffen der Hypothese einschätzt. Man berücksichtigt jedoch nicht, daß diese Evidenz möglicherweise fehlerhaft ist (d.h. man verläßt sich zu stark auf sie) und daß sie unter Umständen sogar mit einer bzw. der Alternativhypothese ebenso gut vereinbar ist (d.h. daß sie keinen diagnostischen Wert hat). Als Folge resultiert eine fälschliche Überzeugung, die Hypothese werde stark durch die empirische Evidenz gestützt.

Im Fall des sexuellen Mißbrauchsverdachts ist man praktisch dazu gezwungen, positiv zu testen, da die umgekehrte Strategie nicht möglich ist: Man könnte nur dann negativ testen,

wenn man über Vorstellungen verfügte, welche Evidenz für das Nicht-Vorliegen von sexuellem Mißbrauch typisch ist. Wer überprüfen will, ob kein sexueller Mißbrauch vorliegt, der ist dazu gezwungen, umgekehrt zu schauen, ob Mißbrauchskennzeichen vorliegen, und aus dem weitgehenden oder völligen Fehlen solcher Kennzeichen das Nicht-Vorliegen von sexuellem Mißbrauch zu schließen. Insofern ergibt sich notwendigerweise eine positive Teststrategie bei der Überprüfung des Mißbrauchsverdachts. (Wir haben das auch explorativ an einigen studentischen Versuchspersonen untersucht, die wir fragten, wie sie gezielt die Annahme überprüfen würden, daß kein sexueller Mißbrauch gegeben ist. Sie nannten daraufhin genau das obige Vorgehen.)

Wenn also Personen – wie sie es oft tun (Friedrich, 1993) – bei der Überprüfung des Mißbrauchsverdachts pseudodiagnostisch vorgehen, dann wirkt sich dies wegen der zugleich gegebenen Notwendigkeit zum positiven Testen konfirmatorisch aus. Man berücksichtigt nicht, daß viele der vermeintlichen Mißbrauchskennzeichen, die man findet (z.B. eigentümliches Verhalten des Kindes, mehrdeutige Äußerungen), zum einen verschiedenen Fehlerquellen unterliegen können (z.B. ungenaue Erinnerung, falls man dies aus dem Gedächtnis abrufen), zum anderen auch mit vielen Alternativhypothesen kompatibel sein können. So wählen zum Beispiel Kinder oftmals beiläufig Formulierungen, die man bei Erwachsenen für sexualisiert halten würde (z.B. „dann ist sein Ding in der Hand immer größer geworden“ – das Kind will u.U. nur das Aufblasen eines Luftballons beschreiben). Die bei pseudodiagnostischem Vorgehen fehlende Berücksichtigung der Basisrate verschärft dieses Problem: Bezieht man die Basisrate nicht mit ein, so tut man so, als wäre Mißbrauch ebenso wahrscheinlich wie Nicht-Mißbrauch. Die Schätzungen für die Prävalenz von sexuellem Mißbrauch schwanken zwar sehr stark, aber hinsichtlich einer Basisrate von unter 50% scheint doch Einigkeit zu bestehen (vgl. z.B. Endres & Scholz, 1994; Köhnken, 1999). Stellt man also fest, daß die vorliegende Evidenz mit dem Mißbrauchsverdacht kompatibel erscheint, ohne zu berücksichtigen, daß alternative Deutungen nicht nur möglich, sondern evtl. aufgrund der höheren Basisrate auch wahrscheinlicher sind, so resultiert eine Neigung zur fälschlichen Bestätigung des Mißbrauchsverdachts.

*Motivationale Prozesse.* Menschen testen Hypothesen, weil sie Fehler vermeiden möchten, die sich durch fälschliche Annahme oder Zurückweisung einer Hypothese ergeben (Trobe & Liberman, 1996). Wären Fehler völlig bedeutungslos, so könnte man auch per Münzwurf (oder auf andere, kognitiv nicht beanspruchende Art und Weise) über Zutreffen oder Nicht-Zutreffen der Hypothese entscheiden. Nun müssen beide Fehler – fälschliche Annahme vs. fälschliche Bestätigung – nicht subjektiv gleichwertig erscheinen, sondern der eine kann als schlimmer empfunden werden als der andere. Dem trägt die hypothesenprüfende Person durch Schwellenwertverschiebungen Rechnung: Erscheint ihr beispielsweise eine fälschliche Zurückweisung einer Hypothese schwerwiegender als eine fälschliche Annahme, so wird sie schon bei vergleichsweise schwacher Evidenz die Hypothese annehmen.

Werden die Fehler im Bereich des sexuellen Mißbrauchsverdachts als asymmetrisch wahrgenommen? In einer von uns durchgeführten empirischen Vorstudie gaben die Probanden im Durchschnitt an, es sei schlimmer, wenn ein Mißbrauchsfall nicht aufgedeckt werde, als wenn eine Person fälschlicherweise beschuldigt werde. Daher sollte man eher eine Falschbeschuldigung wegen sexuellen Mißbrauchs riskieren als einen realen Fall sexuellen Mißbrauchs unentdeckt lassen. Die Schwelle zur Annahme der Hypothese (hier: Mißbrauch liegt vor) ist daher vergleichsweise niedrig; uneindeutige Evidenz kann dann leichter zur Bestätigung als zur Zurückweisung des Mißbrauchsverdachts führen.

## Fazit: Die Selbstbestätigung des Mißbrauchsverdachts

Die Überprüfung des Verdachts des sexuellen Kindesmißbrauchs läßt sich als Testung einer sozialen Hypothese verstehen. Wie wir gezeigt haben, ist soziales Hypothesentesten unter bestimmten Bedingungen konfirmatorisch – d.h. die getestete Hypothese wird dann leichter bestätigt, als durch die Gesamtevidenz gerechtfertigt wäre. Solche Bedingungen sind beim Verdacht des sexuellen Kindesmißbrauchs in beinahe idealtypischer Weise gegeben. Daher gewinnen viele Personen auch angesichts schwacher Evidenz fälschlicherweise den Eindruck, der Verdacht des sexuellen Kindesmißbrauchs sei zutreffend.

Ein solcher Eindruck hat nun wiederum auch Rückwirkungen darauf, wie der Mißbrauchsverdacht bei neuer Evidenz weiter geprüft wird. Viele der vorhin dargestellten Mechanismen, die sich konfirmatorisch auf die Hypothesentestung auswirken, prägen sich um so stärker aus, je stärker man von der Richtigkeit der getesteten Hypothese überzeugt ist. Dies gilt beispielsweise für die hypothesenkonforme Enkodierung und den hypothesenkonformen Abruf von Evidenz (Ross, 1989), die hypothesenkonsistente Interpretation uneindeutiger Evidenz (Darley & Gross, 1983), die Wahl einer konfirmatorischen Teststrategie (Trope & Thompson, 1997) und die hypothesenbegünstigende Wahl der Entscheidungsschwelle (Trope & Liberman, 1996). Der Anfangsverdacht, jemand habe sexuellen Kindesmißbrauch begangen, kann insofern einen regelrechten Zirkel der Selbstbestätigung in Gang setzen; einen Teufelskreis, an dessen Ende die Überzeugung, es liege sexueller Mißbrauch vor, kaum noch zu widerlegen ist.

Unser Erklärungsansatz für Falschbeschuldigungen wegen sexuellen Kindesmißbrauchs postuliert darüber hinaus auch Gruppenprozesse, die zu einer weitergehenden Verfestigung bis hin zur Immunisierung der Überzeugung, es liege Mißbrauch vor, bei den beteiligten Personen führen können (siehe Schulz-Hardt & Köhnken, 2000). Solche Gruppenprozesse werden wirksam, wenn sich die betroffenen Eltern – wie zum Beispiel beim Montessori-Prozeß geschehen – zu Gruppen zusammenschließen, in denen sie, oft auch unter Mitwirkung von Mitarbeitern von Kinderschutzorganisationen, so etwas wie eine „kollektive Hypothesentestung“ betreiben. Wir sparen diesen Teil des Erklärungsansatzes allerdings in diesem Beitrag aus, da er – bis auf eine Pilotstudie, die bereits auf der vorigen Fachgruppentagung Rechtspsychologie in Erlangen vorgestellt wurde – bisher von uns nicht empirisch überprüft wurde.

## **Empirische Untersuchungen zu unserem Erklärungsansatz**

### Experiment 1

Aus unserem Erklärungsansatz lassen sich verschiedene Vorhersagen ableiten. Die grundlegendste ist wohl, daß die Neigung, sexuellen Mißbrauch auch ohne hinreichende Evidenz anzunehmen, nicht Personen in einer psychischen Ausnahmesituation vorbehalten ist. Vielmehr hatten wir postuliert, daß auch rational vorgehende, unvoreingenommene, nicht persönlich betroffene und um nichts als die Wahrheit bemühte Personen dazu neigen, den Mißbrauchsverdacht fälschlicherweise zu bestätigen. Eine weitere unmittelbare Folgerung aus unserem Ansatz ist, daß dies insbesondere dann auftreten sollte, wenn das empirische Datenmaterial Interpretationsspielraum läßt – konfirmatorische Prozesse wie etwa die hypothesenkonsistente Interpretation uneindeutigen Materials oder auch die konfirmatorische Teststrategie können, wie wir beschrieben haben, besser wirken, wenn ein möglichst großer Interpretationsspielraum besteht.

Beide Vorhersagen wurden von uns in einem ersten Experiment überprüft, das bereits auf der letzten Fachgruppentagung referiert wurde; deshalb soll es hier nur so weit erwähnt werden, wie es für das Verständnis der darauf aufbauenden Folgeexperimente notwendig ist. Die Versuchspersonen sahen einen ca. viertelstündigen Amateurfilm, auf dem ein Sportunterricht mit Vorschulkindern zu sehen war. Das Material war objektiv harmlos. Den Probanden in den beiden Experimentalbedingungen wurde mitgeteilt, dieses Material stamme aus einem aktuellen Ermittlungsverfahren gegen den Turnlehrer, gegen den der Vorwurf des sexuellen Mißbrauchs geäußert worden sei. Diese Information wurde entweder vor oder nach dem Ansehen des Films gegeben. Die Annahme war, daß Personen, die vorab über die zu testende Hypothese informiert werden, vergleichsweise weniger Interpretationsspielraum besitzen als Personen, die erst nachträglich von der Hypothese erfahren und nun manche Details (z.B. wo war die Hand bei der Hilfestellung genau?) nicht mehr so exakt rekonstruieren können. In einer Kontrollbedingung wurde kein Mißbrauchsverdacht induziert; den Versuchspersonen wurde vielmehr mitgeteilt, am hiesigen Institut würde zur Zeit Material aus Mißbrauchsverfahren untersucht, und um die dort erhaltenen Ergebnisse richtig einschätzen zu können, würden Vergleichswerte bei Material benötigt, das nicht aus Mißbrauchsverfahren stamme. Zu letzterer Kategorie zähle der Sportstundenfilm. (Um Mißverständnisse zu vermeiden: Damit haben wir für die Probanden noch nicht ausgeschlossen, daß auch dort Mißbrauch vorliegt, der nur nie zur Anzeige kam!).

Die Ergebnisse bestätigten unsere Erwartungen: Während nur 7% der Probanden in der Kontrollbedingung und 13% der Probanden in der Bedingung „Verdacht vorher“ glaubten, der Turnlehrer habe sexuellen Mißbrauch begangen, urteilten 53% (!) der Probanden in der Bedingung „Verdacht nachher“ so. Dies geschah wohlgernekt bei einem Videofilm, der einen absolut harmlosen Sportunterricht zeigt! Offensichtlich läßt die unklarere Datenlage bei erinnertem (im Vergleich zu direkt beobachtetem) Material genügend Interpretationsspielraum, damit konfirmatorische Mechanismen in der Hypothesentestung wirken können.

## Experiment 2

Wenn die Annahme zutrifft, daß mit zunehmendem Interpretationsspielraum konfirmatorische Mechanismen in der Hypothesentestung leichter wirksam werden können und daher eine Falschbeschuldigung wahrscheinlicher wird, dann sollte die Falschbeschuldigungsquote zunehmen, wenn ein weniger eindeutiges Stimulusmaterial als der Turnstundenfilm verwendet wird. In Experiment 2 wurde daher auf ein Tonband zurückgegriffen, auf dem ein vierjähriges Kind über eine zuvor erlebte Zaubervorstellung berichtet. Natürlich war während dieser Zaubervorstellung objektiv nichts Verdächtiges passiert. Allerdings ist Kindersprache, wie vorhin bereits ausgeführt, oft mehrdeutig und läßt weitaus mehr Interpretationsmöglichkeiten, als dies bei visuellen Informationen über die Interaktion von Turnlehrer und Kindern der Fall ist.

Der Versuchsplan wurde aus Experiment 1 übernommen. Die Ergebnisse sind wiederum in Einklang mit den Erwartungen: Insgesamt kam es zu mehr Falschbeschuldigungen als im vorigen Experiment; die Rangordnung der drei Versuchsbedingungen blieb dabei erhalten. In der Kontrollbedingung hielten 13% der Probanden den Zauberer des Mißbrauchs für schuldig, in der Bedingung „Verdacht vorher“ 33% und in der Bedingung „Verdacht nachher“ 57%.

### Experiment 3

Bei den Ergebnissen der ersten beiden Experimente kann nicht ausgeschlossen werden, daß der Hinweis, das Material stamme de facto aus einem Ermittlungsverfahren gegen den Turnlehrer bzw. den Zauberer, einen möglicherweise zu starken Einfluß ausgeübt hat. Die Tatsache, daß in der Bedingung „Verdacht nachher“ jeweils knapp über die Hälfte der Probanden die Person, gegen die angeblich ermittelt wird, für schuldig hielt, muß strenggenommen nicht auf konfirmatorischem Hypothesentesten beruhen. Denkbar wäre auch, daß die Probanden die Hypothese aus dem Gedächtnis heraus nicht zuverlässig testen konnten und sich daher sehr stark auf das Ermittlungsverfahren als Hinweisreiz stützten. Möglicherweise hatten sie die subjektive Hypothese, wenn denn schon ein Ermittlungsverfahren eröffnet werde, dann müsse es ja wohl gute Gründe dafür geben, und dann sei die Wahrscheinlichkeit, daß der Verdacht berechtigt ist, zumindest 50:50. Experiment 3 stellt daher eine Replikation von Experiment 2 dar, allerdings mit dem wichtigen Unterschied, daß der eben angesprochene Hinweisreiz deutlich abgeschwächt wurde. Diesmal wurde den Probanden in den Experimentalbedingungen mitgeteilt, für die Verwendung von Material aus realen Ermittlungsverfahren gebe es sehr strenge Auflagen. Daher müßten mehrere Tonbänder verwendet werden, von denen nur eines aus einem Ermittlungsverfahren stamme. Nicht einmal der Versuchsleiter wisse, welches der Tonbänder das sei (dies werde erst später von Dritten zugeordnet). Trotz dieser Veränderung konnten jedoch die Ergebnisse von Experiment 2 vollständig repliziert werden. Damit konnte diese Alternativerklärung für die Befunde der ersten beiden Experimente ausgeschlossen werden.

### Experiment 4

Im vierten Experiment wurde nur die „Verdacht nachher“-Bedingung aus den Experimenten 2 und 3 repliziert. Untersucht werden sollten dabei zwei Fragen: Zum sollte dann, wenn die Falschbeschuldigungen auf einer Selbstbestätigungsneigung des Mißbrauchsverdachts beim Hypothesentesten beruhen, Falschbeschuldigungen seltener bei relativ spontanen Urteilen auftreten als bei stärker elaborierten Urteilen. Die Probanden wurden daher in diesem Experiment zweimal um ein Urteil gebeten: Einmal direkt nach dem Anhören des Tonbandes und ein weiteres Mal, nachdem sie längere Zeit (angeleitet durch entsprechende Instruktionen und Fragen) über den Mißbrauchsverdacht nachgedacht hatten. Die Ergebnisse bestätigen dies: Während bei der spontanen Urteilsabfrage „nur“ ca. die Hälfte der Probanden den Zauberer für schuldig hielt, waren dies am Ende mehr als zwei Drittel (!). Dieser sehr hohe Prozentsatz ist auch in Zusammenhang mit der zweiten Frage des Experiments interessant: Wie bereits ausgeführt, schildern Kinder ihre Erlebnisse bisweilen mit Formulierungen, die man bei Erwachsenen für sexualisiert halten würde. Solche Formulierungen kamen auch in dem von uns verwendeten Tonband vor (z.B. als das Kind über den Zauberstab des Zauberers sprach). Falls jemand kindliche Sprache nicht gewöhnt ist, mag es nicht so überraschend sein, daß diese Person nicht um die Normalität solcher Äußerungen bei Kindern weiß und sie vor dem Hintergrund dieses Wissensdefizits in subjektiv richtiger Weise als diagnostisch relevante Evidenz für Mißbrauch deutet. Da die Experimente 1 bis 3 mit studentischen Probanden durchgeführt wurden, kann ein solches Wissens- und Erfahrungsdefizit nicht ausgeschlossen werden. Experiment 4 wurde daher mit einer Stichprobe durchgeführt, die in hohem Maße Erfahrung mit Kindern sowie mit kindlicher Kommunikation hat – nämlich mit angehenden sowie ausgebildeten Kinderkrankenschwestern und Kinderkrankenpflegern. Wie die oben erwähnten Ergebnisse zeigen, schützen die höhere Erfahrung und das höhere Wissen im Zusammenhang mit kindlicher Kommunikation keineswegs vor fälschlicher Bestätigung des Mißbrauchsverdachts. Eine weitere Studie, in der der Versuchsplan und Versuchsablauf des

dritten Experiments mit Eltern drei- bis sechsjähriger Kinder repliziert wird, wird momentan durchgeführt.

## Experiment 5

Bei der Erläuterung unseres Erklärungsansatzes wurde als ein Mechanismus, der konfirmatorisch auf die Testung des Mißbrauchsverdachts wirkt, eine asymmetrische Gewichtung der Fehler dargestellt. Unsere Annahme, die wir durch eine kleine Vorstudie untermauern konnten, war: Personen, die mit dem Verdacht des sexuellen Mißbrauchs konfrontiert sind, finden es schlimmer, wenn ein Fall von tatsächlichem sexuellem Kindesmißbrauch unentdeckt bleibt, als wenn eine unschuldige Person fälschlicherweise wegen sexuellen Mißbrauchs angeklagt wird. Diese unterschiedliche Fehlergewichtung kommt dadurch zustande, daß die Folgen von sexuellem Mißbrauch für die Personen salienter und schwerwiegender sind als die Folgen einer Falschanklage. Als Konsequenz hieraus wählen sie einen vergleichsweise niedrigen Schwellenwert, ab dem sie die Mißbrauchshypothese annehmen, d.h. es genügt schon eine relativ schwache Befundlage, um die Hypothese zu bestätigen.

In Experiment 5, wiederum mit angehenden sowie ausgebildeten Kinderkrankenschwestern und Kinderkrankenpflegern, haben wir daher versucht, diese Fehlergewichtung zu manipulieren. In der Kontrollbedingung wurde der Ablauf der „Verdacht nachher“-Bedingung aus Experiment 3 repliziert (d.h. es wurde wieder auf das Tonbandmaterial zurückgegriffen; und der Verdacht wurde durch den Zusatz, nur eines von mehreren Tonbändern stamme aus einem Ermittlungsverfahren, abgeschwächt). In den beiden Experimentalbedingungen bekamen die Probanden vor dem Abgeben des Urteils jeweils einen Text zu lesen. In diesem Text wurden entweder die gravierenden Folgen sexuellen Kindesmißbrauchs für das Opfer beschrieben, oder es wurden die verheerenden Konsequenzen einer Falschanklage für den zu Unrecht Beschuldigten verdeutlicht. Unsere Annahme war, daß der Mißbrauchsoffer-Text keine Wirkung im Vergleich zur Kontrollbedingung besitzen würde, da hier lediglich derjenige Fokus induziert wurde, den Personen ohnehin bei der Testung des Mißbrauchsverdachts aufweisen. Umgekehrt sollte durch den Falschanklage-Text eine Verschiebung des Fokus der Versuchspersonen und somit eine veränderte Fehlergewichtung erfolgen.

Die Ergebnisse bestätigen dies: Sowohl in der Kontrollbedingung als auch in der Mißbrauchsoffer-Text-Bedingung glaubten 92% (!) der Versuchspersonen, es liege sexueller Mißbrauch vor. Dieser Prozentsatz wurde in der Bedingung mit dem Falschanklage-Text reduziert, betrug dort aber immer noch 76%. Auf der einen Seite zeigt dieses Experiment also, daß die subjektive Gewichtung der Folgen eines übersehenen Mißbrauchs bzw. einer Falschanklage einen Einfluß auf die Testung der Mißbrauchshypothese ausübt, und daß man insofern durch eine Erhöhung der Salienz der Konsequenzen einer Falschanklage eine etwas „fairere“ Hypothesentestung ermöglichen kann. Auf der anderen Seite muß man sich aber auch verdeutlichen, daß selbst dort, wo die oftmals verheerenden Konsequenzen von Falschanklagen verdeutlicht wurden, noch drei von vier Personen bereit waren, eine ihnen unbekannt Person für einen Sexualstraftäter zu halten, nachdem sie eine kindliche Schilderung einer harmlosen Zaubervorstellung im Rahmen eines Kinderfestes gehört hatten. Viel drastischer läßt sich der dringende Forschungsbedarf in diesem Bereich eigentlich nicht illustrieren.

## Ausblick

Die hier dargestellten Untersuchungen stellen nur einen Ausschnitt aus unseren Forschungsbemühungen zu diesem Thema dar. Wie sie verdeutlicht haben sollten, sind Falschbeschuldigungen wegen sexuellen Mißbrauchs nicht zwangsläufig das Produkt von Böswilligkeit oder die Folge einer Unfähigkeit oder eines Unwillens, sich mit den Tatsachen auseinanderzusetzen. Vielmehr resultieren sie aus Gesetzmäßigkeiten, denen Menschen unterliegen, wenn sie soziale Hypothesen wie z.B. den Verdacht des sexuellen Kindesmißbrauchs testen. Soziales Hypothesentesten verläuft unter bestimmten situativen Bedingungen konfirmatorisch, d.h. es besteht dann eine erhöhte Neigung, die getestete Hypothese zu bestätigen. Wie wir gezeigt haben, liegen genau solche Bedingungen beim Verdacht des sexuellen Kindesmißbrauchs vor. Daher muß man damit rechnen, daß auch normal begabte, unvoreingenommene und nicht durch Faktoren wie übermäßige emotionale Involviertheit oder Meinungsdruck von außen getrübt Personen eine erhöhte Neigung aufweisen, aufgrund objektiv harmlosen Materials einen sexuellen Mißbrauchsverdacht als bestätigt anzusehen. „Schuldig bei Verdacht“ – diese fatale Kopplung ist, wie unsere Überlegungen und unsere Studien zeigen, im Kontext des potentiellen sexuellen Kindesmißbrauchs leider alles andere als ein Ausnahmefall.

Augenblicklich laufende sowie für die Zukunft geplante Untersuchungen beschäftigen sich zum einen damit, die postulierten vermittelnden Prozesse, die wir theoretisch dargelegt haben, noch stärker empirisch herauszuarbeiten. Zum anderen wird auch die Phase der Verfestigung einer Falschbeschuldigung näher untersucht – in diesem Rahmen sollen auch verstärkt Gruppenprozesse betrachtet werden. Und schließlich wird es auch zunehmend um die Frage gehen, welche Interventionen bereitgestellt werden können, um die Rate der fälschlichen Bestätigungen eines sexuellen Mißbrauchsverdachts zu senken, ohne damit den Anteil der fälschlichen Zurückweisungen eines solchen Verdachts zu erhöhen. Mit anderen Worten: Wie kann man Personen im Kontext des Verdachts des sexuellen Kindesmißbrauchs zu besseren Hypothesentestern machen? So lange das nicht gelingt, können sich Justizdebakel, wie wir sie etwa im Montessori-Prozeß erleben mußten, jederzeit wiederholen.

### **Abschliessender Hinweis**

Die hier dargestellte Forschung wäre nicht möglich gewesen, wenn nicht eine Vielzahl von Diplomandinnen und Diplomanden die empirischen Studien durch ihr großes Engagement, ihre Kreativität, ihren Fleiß und ihr Können realisiert hätten und immer noch realisieren. Im einzelnen waren und sind dies: Axel Amann, Anja Beer, Birgit Biemann, Thies Carstens, Berit Gerkens, Sandra Lewke, Stefanie Molck und Maike Oetjen. An der wissenschaftlichen Betreuung dieser Arbeiten hat Charlotte Möhlenbeck einen hohen Anteil. Zu danken ist darüber hinaus auch unseren vielen studentischen Hilfskräften und Forschungspraktikanten. Besonders gewürdigt und herausgehoben werden soll abschließend die Arbeit von Stefanie Bern, die als unsere Diplomandin mit bewundernswertem Eifer und großem Können unser Forschungsprojekt bereichert hat, und die leider während ihrer Diplomarbeit durch einen unverschuldeten Verkehrsunfall ums Leben kam.

### **Literatur**

- Ceci, S. J. & Bruck, M. (1993). 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, D.C.: American Psychological Association.



- Darley, J. M. & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, 44, 20-33.
- Endres, J. & Scholz, B. (1994). Sexueller Kindesmißbrauch aus psychologischer Sicht. *Neue Zeitschrift für Strafrecht*, 10, 466-473.
- Fiedler, K., Armbruster, T., Nickel, S., Walther, E. & Asbeck, J. (1996). Constructive biases in social judgment: Experiments on the self-verification of question content. *Journal of Personality and Social Psychology*, 71, 861-873.
- Friedrich, J. (1993). Primary error detection and minimization (PEDMIN) strategies in social cognition: A reinterpretation of confirmation bias phenomena. *Psychological Review*, 100, 298-319.
- Hodgins, H. S. & Zuckerman, M. (1993). Beyond selecting information: Biases in spontaneous questions and resultant conclusions. *Journal of Experimental Social Psychology*, 29, 387-407.
- Klayman, J. & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94, 211-228.
- Köhnken, G. (1997). Suggestive Prozesse in Zeugenbefragungen: Formen und theoretische Erklärungsansätze. *Monatsschrift für Kriminologie und Strafrechtsreform*, 5, 290-299.
- Köhnken, G. (1999). Sexueller Mißbrauch. In R. Lempp, G. Schütze & G. Köhnken (Hrsg.), *Forensische Psychiatrie und Psychologie des Kinder- und Jugendalters*. Darmstadt: Steinkopf.
- Newman, J. P., Wolff, W. T. & Hearst, E. (1980). The feature positive effect in adult human subjects. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 630-650.
- Ross, M. (1989). Relation of implicit theories to the construction of personal history. *Psychological Review*, 96, 341-357.
- Schulz-Hardt, S. & Köhnken, G. (2000). Wie ein Verdacht sich selbst bestätigen kann: Konfirmatorisches Hypothesentesten als Ursache von Falschbeschuldigungen wegen sexuellen Kindesmißbrauchs. *Praxis der Rechtspsychologie*, 10, Sonderheft 1, 60-88.
- Snyder, M. & Swann, W. B., Jr. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, 36, 1202-1212.
- Snyder, M. & Uranowitz, S. W. (1978). Reconstructing the past: Some cognitive consequences of person perception. *Journal of Personality and Social Psychology*, 36, 941-950.
- Srull, T. K. & Wyer, R. S., Jr. (1989). Person memory and judgment. *Psychological Review*, 96, 58-83.
- Stangor, C. & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*, 111, 42-61.
- Trope, Y. & Liberman, A. (1996). Social hypothesis testing: Cognitive and motivational mechanisms. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 239-270). New York: Guilford Press.
- Trope, Y. & Thompson, E. P. (1997). Looking for truth in all the wrong places? Asymmetric search of individuating information about stereotyped group members. *Journal of Personality and Social Psychology*, 73, 229-241.
- Volbert, R. & Pieters, V. (1996). Suggestive Beeinflussung von Kinderaussagen. *Psychologische Rundschau*, 47, 183-198.
- Wilson, T. D. & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, 116, 117-142.



# Tötung in Verbindung mit sexueller Gewalttat

*Peter Steck, Fachbereich Psychologie der Universität Konstanz*

## **1 Fragestellung**

Für Tötung in Verbindung mit sexueller Gewalttat wird traditionell das Konstrukt des sog. Deckungsmordes als Erklärungshilfe in Anspruch genommen (v. Hentig, 1956; Brückner, 1961). Nur vordergründig erscheint eine solche Tötungshandlung zur Verdeckung einer Straftat als eine durchdachte Handlungsweise, während die Tatverläufe, näher besehen, auf konflikthafte, von heftigen Affekten begleitetes Geschehen hinweisen (vgl. Steck & Pauer, 1992). Diese Beobachtung lieferte den Anlaß für die im folgenden referierte Untersuchung, die im Rahmen eines von der Deutschen Forschungsgemeinschaft geförderten Projekts mit dem Titel „Tötung als Konfliktreaktion“ (Steck, Schrader & Auchter, 1999) durchgeführt wurde. Ausgehend von der Erfahrung, daß sexuelle Gewalttaten mehrheitlich über Stunden oder Tage hin geplante Delikte darstellen (Amir, 1971; Steck & Pauer, 1992), wurde postuliert, daß Tötung im Zuge solcher Delikte nicht nur durch Handlungsbarrieren im Tatverlauf selbst, sondern auch durch Faktoren in der Planung und in der antizipierenden Handlungsregulation begründet sein können. Das Thema Tötung als Konfliktreaktion weist hier also den Doppelsinn auf, daß anhaltende Konfliktbelastung im Vorfeld der Tat ebenso wie vereitelnde Anreize im Tatverlauf selbst als Risikofaktoren zu untersuchen sind. Das bekannt hohe Gewicht einer kriminellen Karriere bei schwerer Kriminalität, namentlich auch bei Tötungsdelinquenz (s. Wulf, 1978), veranlaßt, auch persönlichkeitspsychologische Hintergründe in die Analyse der Handlungsregulation aufzunehmen.

Der Untersuchungsgegenstand wurde in Anlehnung an ein im Rahmen des Forschungsprojekts erprobtes Modell (Steck, Matthes, Wenger de Chavez & Sauter, 1997) operationalisiert. Die Handlungsanalyse orientierte sich an einem Phasenmodell des Tötungsaktes, das seinerseits eine Erweiterung des von Kahlert und Lamparter (1979) entwickelten Modells darstellt. Es unterteilt den Handlungsablauf in eine Anlaufzeit – Zeitraum von längstens vier Wochen vor der Tat – eine Tatvorszene, eine Vortat, eine Haupttat und ein Nachtatverhalten. Unter diesem Modell werden für die Tatanlaufzeit psychosoziale Belastungsfaktoren erfaßt, die als Bedingungen eines persönlich erhöhten Konfliktniveaus gelten können, sowie deliktspezifische Kognitionen wie Artikulationen der Tatabsicht und Antizipationen von Opferreaktionen. Als relevante Ereignisse der Vorszene gelten unmittelbar tatvorbereitende Verhaltensmuster, ferner konfliktverschärfende Vorgänge in den letzten Stunden vor der Tat bis zur ersten Begegnung mit dem Opfer. Die Vortat wird hier durch die Merkmale der sexuellen Gewalttat definiert, die Haupttat schließlich durch die auf eine tödliche Verletzung zielenden Bewegungen. Die Beobachtung des Nachtatverhaltens bezieht sich auf tatbezogene Reaktionen des Täters innerhalb der ersten Stunde nach der Tat.

Die Eigentümlichkeit der tödlich endenden sexuellen Gewalttat sollten durch eine Kontrastierung des so unterteilten Tatverlaufs mit sexuellen Gewalttaten ohne tödlichen Ausgang sichtbar gemacht werden. Mit Ausnahme der Haupttat selbst kann die sexuelle Gewalttat ohne Tötungsversuch in gleichnamige Phasen mit vergleichbaren Merkmalen untergliedert werden. Die Kontrastierung wird, wie oben angedeutet, auf den persönlichkeitspsychologischen Hintergrund der Täter ausgedehnt, der in Anlehnung an das Konzept der kriminellen Karriere über psychosoziale Belastungsfaktoren in der Biographie operationalisiert wird.

## **2 Die Stichprobenziehung**

Die Zielgruppe der Untersuchung setzte sich aus 22 Männern im Alter zwischen 19 und 41 Jahren zusammen, die wegen versuchter oder vollendeter vorsätzlicher Tötung im Zusammenhang mit Vergewaltigung bzw. sexueller Nötigung zu Freiheitsstrafen verurteilt worden waren. Die Vergleichsgruppe bestand aus 48 Männern im Alter zwischen 17 und 57 Jahren, die wegen gleichartiger sexueller Gewalttaten ohne Tötungsversuch zu Freiheitsstrafen verurteilt worden waren. Beide Gruppen Straffälliger waren in Vollzugsanstalten der alten Bundesländer für die Mitarbeit angeworben worden.

## **3 Datenerhebung und –verarbeitung**

Die Datenerhebung erfolgte über halbstandardisierte Interviews mit den Verurteilten zwischen Mai 1995 und März 1996. Die Interviewer hatten vor der Befragung Gelegenheit, das betreffende Strafurteil und seine Begründung einzusehen.

Grundlage der Datenverarbeitung bildeten die dichotomisierten Antworten auf 144 Standardfragen, die allen Probanden gleichlautend gestellt worden waren. Die Ergebnisse wurden diskriminanzanalytisch ausgewertet über die Berechnung von multiplen logistischen Regressionen mit der Gruppenzugehörigkeit als Kriterium und den dichotomen Antworthäufigkeiten als Prädiktoren. Die Prädiktoren wurden entsprechend ihrer psychologischen Bedeutung gruppiert, d.h. die Tatvariablen gemäß dem oben beschriebenen Phasenmodell des Tötungsdelikts in Prädiktorengruppen geordnet, die lebensgeschichtlichen Variablen gemäß ihrer Zugehörigkeit zu als jeweils bedeutsam erachteten biographischen Dimensionen. Aus diesen Vorgaben resultierten Berechnungen von 21 multiplen Regressionen auf das Kriterium der Gruppenzugehörigkeit mit drei bis 13 Prädiktoren. Der Modelltest wurde jeweils auf dem 5-Prozent-Niveau der Signifikanz beurteilt, ebenso der Regressionskoeffizient für die einzelnen Prädiktoren. Unter wahrscheinlichkeitstheoretischen Gesichtspunkten ist dieses Verfahren sicher als problematisch anzusehen, doch sollten hier alle möglichen für eine Trennung der Gruppen in Betracht kommenden Faktoren einbezogen werden; ein erhöhtes Risiko für den statistischen Einzelnachweis wurde dafür in Kauf genommen.

## **4 Ergebnisse**

Keine der multiplen Regressionen, in die biographische Daten als Prädiktoren für die Gruppenzugehörigkeit einbezogen waren, führten zu einem statistisch signifikanten Ergebnis. Die zehn hier geprüften Prädiktorengruppen erfassten (1) Belastungsfaktoren in der Herkunftsfamilie wie Gewaltanwendung in der Erziehung, frühe Trennung von den Eltern, Straffälligkeit von Familienmitgliedern, (2) Beeinträchtigung des sozialen Status wie Schulversagen, fehlende Berufsausbildung und Arbeitslosigkeit, (3) psychosoziale Auffälligkeiten wie Verhaltensstörungen von klinischem Rang, frühe dissoziale Auffälligkeiten und Suchtprobleme, (4) Zeichen sozialer Instabilität wie häufiger Wechsel der Partnerin, des Arbeitsplatzes, (5) kriminelle Vorbelastung, (6) schwere finanzielle Probleme, (7) Alkoholgewohnheiten, (8) Quellen sexueller Aufklärung, (9) Pornographiekonsum, (10) häufiger Partnerwechsel bei sexuellen Kontakten. Die beobachteten Häufigkeitsverteilungen in unserem Datenmaterial veranlassen also nicht zu der Vermutung, daß hinsichtlich dieser Merkmalsbereiche die Tötungsdelinquenten einer anderen Grundgesamtheit entstammen als die sexuellen Gewalttäter ohne tödliche Gewaltanwendung.

Demgegenüber führten sieben der elf Regressionsberechnungen mit Handlungsmerkmalen als Prädiktoren auf die Gruppenzugehörigkeit zu signifikanten (unter  $p < .05$ ) Prüfgrößen (s. Tabelle 1). Solche signifikanten Effekte fanden sich bei Merkmalen der Tatplanung, der Konfliktverschärfung durch tätliche Gegenwehr des Opfers, der Muster sexueller Gewalttaten, der vom Täter erinnerten Emotionen und Kognitionen, des Verlaufs der Gewalthandlungen und schließlich des Nachtatverhaltens. Keine signifikanten Effekte ließen sich für die erfaßten Merkmale des Vorstadiums der Tat sowie für Tatumstände wie Bewaffnung des Täters und Tatzeit, ferner für spezielle Verhaltensmuster des Opfers außer tätlicher Gegenwehr z.B. Schreien, Flucht, Drohen – nachweisen. Betrachtet man die Prädiktoren, deren Gewichte unter einem signifikanten Modelltest jeweils signifikant (unter  $p < .05$ ) geworden sind, ergibt sich folgendes Bild: Bei tödlich endenden sexuellen Gewalttaten entspringt die Gewaltanwendung seltener einem spontanen Einfall, wird also länger als Absicht verfolgt, ferner antizipiert der Täter häufiger Tatfolgen; bei tödlich endenden sexuellen Gewalttaten kommt es häufiger zu tätlicher Gegenwehr des Opfers; unter den von den Tätern für den Tathergang erinnerten Emotionen und Kognitionen trennen die Angaben über die Gefühle der Unsicherheit und der Bedrohtheit sowie die Konfrontation mit unerwarteter Reaktion des Opfers die Tätergruppen signifikant, jeweils mit einer größeren Häufung solcher Angaben bei den Tötungsdelinquenten; hinsichtlich des Tatverlaufes ergeben sich signifikante Unterschiede in der Häufigkeit der Penetration, die bei tödlich endenden sexuellen Gewalttaten seltener vollendet wird, und hinsichtlich des Auftretens einer Handlungsäsur, die von den Tötungsdelinquenten häufiger erinnert wird; im Nachtatverhalten heben sich die Tötungsdelinquenten durch eine signifikante Häufung von untätigem Verharren am Tatort sowie durch ein signifikant häufigeres Bemühen um Spurenbeseitigung von den übrigen sexuellen Gewalttätern ab. Es bleibt zu vermerken, daß alle Items, die als Indikatoren eines besonderen Erregungszustandes des Täters eingesetzt wurden, ebenso wie jene Items, die sexuell deviantes Verhalten oder besondere, über die Gefügigmachung des Opfers hinausgehende Gewalttätigkeiten anzeigen sollten, keine signifikanten Differenzen zwischen den Tätergruppen aufdeckten.

Tabelle1: Statistisch bedeutsame Häufigkeitsunterschiede ( $p < .05$ ) von Tat- und Tätermerkmalen bei sexuellen Gewalttaten mit und ohne tödlichen Ausgang

Merkmal	mit Tötung N = 22	ohne Tötung N = 48
Tatidee spontan umgesetzt	6	28
Antizipation von Tatfolgen	6	2
Tätliche Gegenwehr des Opfers	19	23
Gefühl der Unsicherheit	15	14
Gefühl der Bedrohtheit	14	5
Unerwartete Reaktion des Opfers	16	10
Vollendung der Penetration	10	44
Unterbrechung des Handlungsablaufes	17	22
Untätiges Verharren am Tatort nach der Tat	5	3
Versuch, Spuren zu beseitigen	15	5

## 5 Diskussion

Zunächst ist hervorzuheben, daß die unter dem Modell der kriminellen Karriere erfolgte Untersuchung biographischer Vorbedingungen sexueller Gewalttaten keinen Hinweis auf einen Unterschied zwischen den Tötungsdelinquenten und den anderen Gewalttätern ergeben hat. Bedenkt man, daß die Untersuchung ein relativ breites Spektrum an biographischen Dimensionen erfaßt hat, ist man wohl geneigt anzunehmen, daß beide Tätergruppen persönlichkeitspsychologisch gesehen derselben Grundgesamtheit entstammen. Prüfstatistisch beweisbar ist eine solche Feststellung zwar nicht, doch muß man bis zum Beweis des Gegenteils davon ausgehen, daß beim Täter keine lebensgeschichtlich fassbaren psychologischen Vorbedingungen speziell für tödliche Gewaltanwendung bei Vergewaltigung zu identifizieren sind.

Die Aufmerksamkeit richtet sich demzufolge auf die Bedingungen der Gewalthandlung selbst. Im Zusammenhang betrachtet zeigen die in dieser Untersuchung beobachteten signifikanten Effekte Besonderheiten tödlich endender Gewalttaten in der Tatplanung und im Tatverlauf an. Hinsichtlich Tatverlauf sind tätliche Gegenwehr der Opfers und – vermutlich in Verbindung damit – Erinnerungen der Überraschung, der Verunsicherung und der Bedrohtheit auf Seiten des Täters zu vermerken. Die ebenfalls signifikanten Faktoren der Unterbrechung des Handlungsablaufes und der nicht vollendeten Penetration mögen in den genannten Besonderheiten ihre Erklärung finden. Insgesamt deuten diese Besonderheiten an, daß der tödliche Akt unmittelbar aus der sexuellen Gewalttat erwächst, daß es keines besonderen Anreizes mehr für die Tötung unter der eskalierenden Gewaltanwendung bedarf. Für diese Vermutung spricht auch, daß verbale Äußerungen des Opfers, z.B. Drohungen, keine nachweisbare Wirkung entfaltet haben. Schwer interpretierbar scheint der Befund, wonach tödlich endende sexuelle Gewalthandlungen seltener spontanen Einfällen entspringen als die nicht tödlich endenden, also längerfristig geplant erscheinen. Vergleichbare Beobachtungen zur Tötungsdelinquenz (Simons, 1988; Volbert, 1990) lassen eher das Gegenteil vermuten. Hier ist freilich zu bedenken, daß die gedankliche Antizipation sexueller Gewaltanwendung und ihrer Folgen sich qualitativ von der Tatplanung eines Räubers oder eines Einbrechers unterscheidet. Aus dem Kontext der Befunde erschließt sich die Bedeutung nicht, die der Tatplanung – von den meisten befragten Tötungsdelinquenten wurde ein Zeitraum von ein bis zwei Tagen angegeben – zukommt. Aus handlungstheoretischer Sicht mag man lediglich in den Besonderheiten des Nachtatverhaltens – Untätigkeit, Suche der Spurenbeseitigung – den Hinweis auf Zielkonflikte in der Folge der tödlichen Gewaltanwendung erkennen.

## Literatur

- Amir, M. (1971). *Patterns in forcible rape*. Chicago: University of Chicago Press.
- Brückner, G. (1961). *Zur Kriminologie des Mordes*. Hamburg: Kriminalistik Verlag.
- Hentig, H. v. (1956). *Zur Psychologie der Einzeldelikte. II. Der Mord*. Tübingen: Mohr.
- Kahlert, T. H. & Lamparter, U. (1979). Tötungsdelikte bei Jugendlichen und Heranwachsenden. *Monatsschrift für Kriminologie und Strafrechtsreform*, 62, 206 – 217.
- Simons, D. (1988). *Tötungsdelikte als Folge mißlungender Problemlösungen*. Stuttgart: Verlag für Angewandte Psychologie.
- Steck, P. & Pauer, U. (1992). Verhaltensmuster bei Vergewaltigung in Abhängigkeit von Täter- und Situationsmerkmalen. *Monatsschrift für Kriminologie und Strafrechtsreform*, 75, 187 – 197.
- Steck, P., Matthes, B., Wenger de Chavez, C., Sauter, K. (1997). Tödlich endende Partnerkonflikte, *Monatsschrift für Kriminologie und Strafrechtsreform*, 80, 404 – 417.

Steck, P., Schrader, S., Auchter, U. (1999). *Konfliktverlauf und Verhaltensmuster bei Tötungsdelikten*. Unveröffentlichter Forschungsbericht. Konstanz: Fachgruppe Psychologie der Universität Konstanz.

Volbert, R. (1990). Lebensgeschichtliche und situative Entstehungsbedingungen von Tötungsdelikten im Rahmen von Bereicherungstaten. In D. Frey (Hrsg.), *Bericht über den 37. Kongreß der Deutschen Gesellschaft für Psychologie in Kiel 1990* (S. 153 –154). Göttingen: Hogrefe.

Wulf, B. R. (1979). *Kriminelle Karrieren von „Lebenslänglichen“*. München: Minerva.





# Wiedererkennen des Täters oder Identifizieren des Verdächtigten Experimentelle Evidenz für einen Bias in der Zeugengegenüberstellung

*Arne Weigold und Dirk Wentura*

In einem Experiment (N=70) wurde die Frage untersucht, in welchem Maße ein Tatverdächtiger in einer Wahlgegenüberstellung nur aufgrund seines Status als Verdächtiger identifiziert werden kann. Für den Verdächtigten ist die Situation – im Gegensatz zu den Distraktorpersonen – durch eine hohe persönliche Relevanz des Ausgangs der Wahlgegenüberstellung gekennzeichnet: Wird er durch einen Zeugen identifiziert, ist er schwer belastet; wird er nicht identifiziert, ist er in der Regel deutlich entlastet. Diese Relevanz kann zu Zeichen von Anspannung und erhöhtem arousal führen, die möglicherweise wahrnehmbar sind. Zum anderen führt sie zu dem Versuch, sich möglichst unauffällig zu verhalten, was ebenfalls zu wahrnehmbaren Unterschieden zu den Distraktorpersonen führen könnte. Um im Labor diese psychologischen Merkmale des Verdächtigten nachzustellen, wurde in insgesamt sechs Wahlgegenüberstellungen jeweils ein „Verdächtiger“ per Los bestimmt, der die Chance hatte, einen hohen Geldbetrag zu gewinnen, wenn er es schafft, nicht durch Beobachter identifiziert zu werden. Videoaufnahmen dieser Gegenüberstellungen wurden Versuchspersonen mit der Aufgabe präsentiert, jeweils die Person mit der Aussicht auf den Geldgewinn zu identifizieren. Es zeigte sich erwartungskonform, dass diese „Verdächtigten“ mit einer um 10% gegenüber einer Kontrollbedingung (Videoaufnahmen derselben Wahlgegenüberstellungen vor der Auslösung des „Verdächtigten“) erhöhten Wahrscheinlichkeit identifiziert werden.

## 1 Einleitung

Betrachten wir eine typische Wahlgegenüberstellung, wie sie im polizeilichen Alltag vorkommt: Der Augenzeuge sieht durch eine Einwegscheibe eine Reihe von Personen, unter denen sich ein Verdächtiger befindet, den die Polizei als Täter eines schweren Verbrechens vermutet. Der Augenzeuge soll angeben, ob er den von ihm beobachteten Täter des Verbrechens wiedererkennt. Bis zu diesem Punkt ist viel passiert. Ein Verbrechen ist geschehen und der Augenzeuge hat den Täter dabei beobachtet, die Polizei hat das Verbrechen registriert, Indizien gesichert, den Augenzeugen befragt und einen Verdächtigten ermittelt. Dann wurde mit großem Aufwand die Wahlgegenüberstellung organisiert, in der sich der Augenzeuge nun befindet. Seit der Tat sind mehrere Monate vergangen und der Augenzeuge hat nur noch eine sehr schwache Gedächtnisrepräsentation des Täters. Außerdem sehen alle ihm gegenüber gestellten Personen dem Täter relativ ähnlich, immerhin wurden sie direkt oder indirekt nach der Beschreibung des Täters ausgewählt. Nach längerem Überlegen und dem wiederholten Hinweis der Polizeibeamten „es doch einfach mal zu versuchen“ identifiziert der Augenzeuge tatsächlich auf Anhieb den Verdächtigten. Dies gilt als starkes Belastungsindiz, denn immerhin ist die Wahrscheinlichkeit, dieses Ergebnis durch Raten zu erzielen, gering. Dementsprechend werden Richter, Geschworene und die anderen Prozessteilnehmer diese Aussage bei der Urteilsfindung werten.

Doch ist diese Wahrscheinlichkeit tatsächlich so gering? Die psychologische und medizinische Forschung der letzten 30 Jahre beantwortet diese Frage mit einem klaren Nein. Tatsächlich deutet vieles darauf hin, dass die Wahlgegenüberstellung eines der unzuverlässigsten Beweismittel überhaupt ist (Devlin, 1976; zitiert nach Stadler, Fabian & Wetzels, 1992) und der größte einzelne Faktor für Fehlerurteile in strafrechtlichen Prozessen (vgl. z.B. Greene & Loftus, 1984; Wells et al., 1998). Besonders eindrucksvoll konnte dies durch nachträgliche Genanalysen nachgewiesen werden, indem in bereits mit einer Verurteilung abgeschlossenen Fällen noch vorhandene biologische Beweismittel analysiert wurden. Es zeigte sich eine beunruhigend hohe Quote von Fehlerurteilen. Insbesondere wurde

einmal mehr die Zuverlässigkeit von Augenzeugen und Wahlgegenüberstellungen in Frage gestellt (Connors, Lundregan, Miller & McEwan, 1996).

Angesichts dieser Befunde hat sich die psychologische Rechtsforschung über die letzten Jahrzehnte mit der Frage beschäftigt, warum Augenzeugen im allgemeinen und die Wahlgegenüberstellung im besonderen so schlechte Ergebnisse produzieren und wie diese Ergebnisse zu verbessern wären. Durch diese Forschung konnte eine Vielzahl von Einflussfaktoren auf die Güte von Augenzeugenaussagen identifiziert werden (vgl. z.B. Wells, 1978; für eine Übersicht siehe auch Cutler, Penrod & Martens, 1987). Ein Kernproblem ist, dass nicht unbedingt davon ausgegangen werden kann, dass der Augenzeuge über eine verlässliche Gedächtnisrepräsentation verfügt und er somit für die Identifizierung bewusst oder unbewusst andere Informationsquellen heranzieht. Da aber die Gedächtnisrepräsentation das Einzige ist, worauf sich seine Identifizierung gründen sollte, muss versucht werden, jede andere Informationsquelle, mit Hilfe derer er in der Wahlgegenüberstellung den Verdächtigten identifizieren könnte, auszuschalten. Dies hat sich praktisch als sehr schwierig erwiesen, da es tatsächlich verschiedene Informationsquellen gibt, auf die der Augenzeuge zurückgreifen kann. Das Verhalten der Polizeibeamten vor und während der Wahlgegenüberstellung, andere Augenzeugen, Fehler im Ablauf der Zeugenbefragungen und die Auswahl der Alternativpersonen seien hier beispielhaft genannt (vgl. hierzu eingehend Wells et al., 1998).

Eine bisher kaum beachtete Informationsquelle könnte aber auch im Verdächtigten selbst begründet sein. Er weiß, dass sich auf der anderen Seite der Einwegscheibe mehrere Polizeibeamte und ein Augenzeuge befinden, die ihn beobachten und über ihn reden. Sollte der Augenzeuge ihn identifizieren, stehen seine Chancen vor Gericht sehr schlecht und es drohen ihm gegebenenfalls mehrere Jahre Haft. Es ist evident, dass die Situation der Wahlgegenüberstellung beim Verdächtigten als starker Stressor wirkt (Perez, Laireiter & Baumann, 1998), mit dem entsprechende physiologische Veränderungen wie rigides Verhalten, erhöhte Muskelanspannung, Erhöhung der Herzfrequenz sowie vertiefte und beschleunigte Atmung einhergehen können (Bierbaumer & Schmidt, 1996). Im Unterschied dazu geht es für die Alternativpersonen um nichts. Werden Sie etwa fälschlicherweise durch den Zeugen benannt, hat dies keine negativen Konsequenzen.

Schon diese Unterschiede im psychologischen Status könnten zu einer Verzerrung im Identifikationsprozeß führen. Der Verdächtige ist möglicherweise als Verdächtigter zu erkennen. Diese Information wird dann eventuell zu seinen Ungunsten vom Augenzeugen genutzt (vgl. z.B. Stadler, Fabian und Wetzels, 1992). Hinzu kommt, dass der Verdächtige aufgrund der für ihn hohen Relevanz des Verfahrens dazu neigen wird, seinen Status zu verbergen. Aber eben dies könnte von Zeugen wahrgenommen werden (vgl. z.B. Ekman & O'Sullivan, 1991) und ebenfalls zu einer Verzerrung im Identifikationsprozeß beitragen. Stadler, Fabian und Wetzels (1992; vgl. auch Fabian, Stadler & Wetzels, 1995) führten ein erstes Experiment zu dieser Verzerrung – die wir im folgenden als Verdächtigten-Effekt bezeichnen wollen – durch. Sie hatten hierzu die Videoaufzeichnung einer Wahlgegenüberstellung aus einem realen Ermittlungsverfahren zur Verfügung und zeigten diese an dem Verfahren unbeteiligten Versuchspersonen mit der Bitte, den Verdächtigten zu identifizieren. Es zeigte sich ein klarer Verdächtigteneffekt, d.h. der Verdächtige wurde mit einer deutlich und signifikant über dem Rateniveau liegenden Wahrscheinlichkeit identifiziert. Problematisch für eine Generalisierung ist allerdings, dass sich der Verdächtige in diesem Fall sehr auffällig verhielt.

In der hier vorliegenden Studie soll der Verdächtigteneffekt daher unter laborexperimentellen Bedingungen erforscht werden. Die Wahlgegenüberstellungen wurden hierbei mit studentischen Probanden realisiert. Um unter diesen Bedingungen eine hohe Relevanz der Situation für den fiktiven „Verdächtigten“ herzustellen, wurde jeweils eine Person jeder Wahlgegenüberstellung ausgelost und ihr ein hoher Geldgewinn (DM 500) in Aussicht

gestellt, wenn es ihr gelänge, von unbeteiligten Personen möglichst selten als „Verdächtigter“ identifiziert zu werden. Die Situation für diese Personen war somit in mancherlei Hinsicht vergleichbar zu Verdächtigten in realen Ermittlungsverfahren. Der Ausgang der Wahlgegenüberstellung war von hoher Relevanz: Nicht identifiziert bedeutet hier einen hohen Geldbetrag zu erhalten, dort stark vom Verdacht entlastet zu sein. Identifiziert zu werden bedeutet hier, den gut erreichbaren Geldgewinn zu verpassen, dort stark zusätzlich belastet zu sein. Psychophysische Reaktionsmuster und absichtvolles Verhalten sollten vergleichbar sein. Außerdem soll explorativ untersucht werden, worauf Versuchspersonen beim Erreichen ihres Urteils achten, um hiermit eventuelle Prädiktorvariablen zu identifizieren. Wir stützen uns dabei auf Resultate aus der Lügenforschung, nach denen vor allem Körperhaltung und Stimme als mögliche Prädiktorvariablen in Frage kommen (Ekman, 1992; Ekman & O’Sullivan, 1991).

## 2 Methode

### 2.1 Stichprobe

Die Stichprobe bestand aus 70 Versuchspersonen (33 Frauen, 37 Männer). Das Alter der Versuchspersonen variierte von 15 bis 72 Jahren mit einem Mittelwert von  $M = 45$  ( $SD = 13$ ).

### 2.2 Versuchsplan

Der Versuchsplan war im wesentlichen ein einfaktorieller Messwiederholungsplan. Es wurden sechs verschiedene Wahlgegenüberstellungen mit jeweils anderen Personen durchgeführt und auf Video aufgenommen. Von jeder dieser sechs Gegenüberstellungen gab es zwei Versionen, eine Version nach der Auslösung des „Verdächtigten“ (Experimentalbedingung) und eine Version vor der Auslösung des „Verdächtigten“ (Kontrollbedingung). Diese insgesamt 2x6 Videosequenzen wurden auf zwei Videobänder im Sinne einer balancierten Versuchsplananordnung (Lateinisches Quadrat) kopiert: Jedes Band enthielt drei Experimentalversionen und drei Kontrollversionen; war eine Gegenüberstellungsgruppe auf Videoband 1 in der Experimentalversion zu sehen, war die entsprechende Kontrollversion auf Band 2 kopiert worden und vice versa. Jeder Versuchsperson wurde randomisiert entweder Videoband 1 oder 2 gezeigt.

### 2.3 Material

Für die Erstellung der sechs Wahlgegenüberstellungsgruppen wurde zunächst eine Stichprobe von  $N=49$  Personen angeworben und jeweils einzeln fotografiert. Eine weitere Stichprobe ( $N=5$ ) gruppierte die Personen aufgrund der Fotos nach Ähnlichkeit zu Gruppen von drei bis acht Personen. Die Häufigkeit der Zusammengruppierung zweier Personen wurde als Ähnlichkeitsindex in einer Clusteranalyse (Ward-Algorithmus) genutzt, so dass jeweils Gegenüberstellungsgruppen von vier hinreichend ähnlichen Teilnehmern gebildet werden konnten. Die so ausgewählten 24 ( $6 \times 4$ ) Gegenüberstellungspersonen wurden zu dem Aufnahmeterrain eingeladen. Zunächst wurden die sechs Kontrollversionen, d.h. die Wahlgegenüberstellungen vor der Auslösung des „Verdächtigten“, hergestellt. Die vier teilnehmenden Personen wurden vom Versuchsleiter aus dem Warteraum geholt und betraten den Filmraum. Vor einer Wand stellten sie sich auf markierte Plätze. Über ihren Köpfen waren die Nummern 1, 2, 3 und 4 angebracht. Den Personen gegenüber befand sich eine Einwegscheibe. Vor dieser Einwegscheibe standen die ferngesteuerte Filmkamera sowie zwei Filmscheinwerfer zur besseren Ausleuchtung. Die Wahlgegenüberstellungen verliefen wie

folgt: Nach zunächst einer Minute ruhigem Stehen wurden die Personen der Reihe nach (über eine Sprechanlage) aufgefordert, den Satz „Entschuldigung, können Sie mir sagen, wie ich zum Bahnhof komme?“ zu sagen; danach hoben sie (wieder der Reihe nach) ein vor ihnen liegendes Papiertaschentuch auf; es folgten weitere 2.5 Minuten ruhiges Stehen. Für die Endfassung wurde jede Videosequenz auf 1.75 Minuten gekürzt.

Danach wurde jeweils ein Verdächtiger pro Gegenüberstellungsgruppe ausgelost. Die Verdächtigten konnten, im Gegensatz zu den drei Alternativpersonen jeder Wahlgegenüberstellung, DM 500 gewinnen. Die Regeln für die Vergabe der DM 500 waren allen Personen in der Instruktionsphase eingehend erklärt worden. Das heißt, sie wußten, dass während der folgenden Experimentalgegenüberstellungen fünf Beobachter hinter der Einwegscheibe sitzen und versuchen würden, jeweils den Verdächtigten zu identifizieren; derjenige der sechs Verdächtigten, der am seltensten von den Beobachtern identifiziert werden würde, erhielt die DM 500. Die sechs Videoverversionen „mit Verdächtigtem“ wurden nach der Auslosung mit demselben Prozedere wie die Kontrollversionen aufgenommen. Am Ende wurden der am wenigsten identifizierten Person die DM 500 ausgehändigt.

## 2.4 Durchführung

Es wurden Standardvideorekorder und -fernseher verwendet, um die Videos zu präsentieren, wobei der Abstand der Versuchspersonen zum Fernseher und die Bildschirmgröße konstant gehalten wurden. Die Versuchspersonen wurden einzeln (56 Personen) oder in Zweiergruppen (14 Personen) vor dem Fernseher positioniert. Bei den Zweiergruppen wurde mittels einer Trennwand sichergestellt, dass sie sich nicht sehen konnten, und sie wurden vor Beginn der Videovorführung instruiert, sich vollkommen leise zu verhalten. Der Versuchsleiter befand sich während des gesamten Versuchsdurchgangs hinter den Versuchspersonen, um eventuelle Versuchsleitereffekte von vornherein auszuschließen. Nach dem Lesen der Instruktion und der Klärung von Fragen wurde den Versuchspersonen eines der beiden Videobänder (vgl. Versuchsplan) mit den sechs Gegenüberstellungsgruppen gezeigt. Den Teilnehmern war die Entstehung der Videosequenzen genau erklärt worden. D.h. sie wußten, dass jeweils eine Person aus jeder Gegenüberstellungsgruppe die Chance auf den Geldgewinn hatte. Allerdings wurden die Versuchspersonen in dem Glauben gelassen, dass dies für alle sechs von ihr zu beurteilenden Gegenüberstellungen der Fall war. Tatsächlich wurden ihnen aber drei Experimental- und drei Kontrollversionen gezeigt. Nach der Präsentation jeder Gegenüberstellung wurden die Teilnehmer um eine Einschätzung darüber gebeten, welche der Personen sie am ehesten für den Verdächtigten hielten.<sup>1</sup> Nach der Präsentation der Videosequenzen wurde den Teilnehmern ein Fragebogen zu den Verhaltensaspekten vorgelegt, die potenziell die Entscheidungen geleitet hatten. Alle diese Items hatten als Antwortformat eine Likert-Skala von 1 für „sehr unwichtig für Einschätzungen“ bis 8 für „sehr wichtig für Einschätzungen“. Mit drei Items wurde zunächst allgemein nach der Rolle von Körper, Gesicht und Sprache gefragt. Dann wurden diese drei Kategorien nochmals in spezifischere Unterkategorien unterteilt. Die Unterkategorien der Kategorie Körper waren Körperhaltung, Körperausrichtung, Stellung der Beine, Bewegungen der Beine, Position der Arme, Bewegungen der Arme und unwillkürliche Gesten. Die Unterkategorien der Kategorie Gesicht waren Blickrichtung, Gesichtsausdruck und Mimik, und die Unterkategorien von Sprache waren Sprechtempo, Lautstärke und Stimmhöhe.

## 3 Ergebnisse

Wahlentscheidungen für den Verdächtigten wurden mit 100, Wahlentscheidungen für eine der Alternativpersonen mit 0 kodiert. Für jede Versuchsperson wurde der Mittelwert dieser

Kodierungen für die drei von ihr gesehenen Experimental- und die drei Kontrollgegenüberstellungen berechnet, so dass sich diese beiden Variablen als Trefferquote (in %) interpretieren lassen. Die mittlere Trefferquote für die Experimentalversionen betrug  $M = 30.95\%$  ( $SD = 25.58\%$ ), die mittlere Trefferquote für die Kontrollversionen betrug  $M = 20.48\%$  ( $SD = 19.07\%$ ). Der Unterschied von  $\Delta = 10.47\%$  ( $SD = 27.11\%$ ) war signifikant<sup>2</sup>,  $t(69) = 3.23$ ,  $p < 0.01$ . Ein signifikanter Verdächtigteneffekt zeigt sich auch, wenn die mittlere Trefferquote der Experimentalversionen gegen die a priori Basiswahrscheinlichkeit von 25 % (da vier Personen pro Wahlgegenüberstellung teilnahmen) getestet wird,  $\Delta = 5.95\%$  ( $SD = 25.58\%$ ),  $t(69) = 1.95$ ,  $p < 0.05$ .

Um Hinweise darauf zu erhalten, welche Körper- und Verhaltensindikatoren von den erfolgreicherem Versuchspersonen genutzt wurden, wurden zunächst über die einzelnen Items des Abschlußfragebogens eine Faktorenanalyse (Hauptkomponentenanalyse) gerechnet. Vier Faktoren hatten einen Eigenwert größer eins; sie erklären zusammen 67.68% der Varianz. Nach VARIMAX-Rotation ergab sich eine Einfachstruktur, die weitgehend a priori Erwartungen entspricht: Faktor „Merkmale der Stimme“ (Sprechtempo, Stimmhöhe, Lautstärke; alle Ladungen  $> .78$ ); Faktor „Merkmale des Unterkörpers“ (Bewegungen der Beine, Stellung der Beine, Unwillkürliche Gesten; alle Ladungen  $> .68$ ); Faktor „Merkmale des Oberkörpers“ (Position der Arme, Bewegungen der Arme, Körperausrichtung, Blickrichtung; alle Ladungen  $> .55$ ) und Faktor „Ausdruck von Gesicht und Körperhaltung“ (Gesichtsausdruck, Mimik, Körperhaltung; alle Ladungen  $> .44$ ). Es wurden entsprechende Skalen durch Mittelung der genannten Items gebildet. Die Skalen und die drei summarischen Items für Körper, Gesicht und Stimme wurden mit dem Verdächtigten-Effekt (d.h. der Differenz der Trefferquote zur Basiswahrscheinlichkeit) korreliert. Die Ergebnisse sind in Tabelle 1 dargestellt.

Tabelle 1: Korrelationen des Verdächtigten-Effekts mit den Fragebogenskalen

Variable	Skala Stimme	Skala Unterkörper	Skala Oberkörper	Skala Gesicht	Stimme (Einzelitem)	Körper (Einzelitem)	Gesicht (Einzelitem)
r	.20 <sup>†</sup>	.20 <sup>†</sup>	.05	-.10	.26*	-.11	-.24*

<sup>†</sup>  $p < 0.10$ ; \*  $p < 0.05$ ;  $n = 70$

Signifikante Zusammenhänge finden sich lediglich für die Einzelitems. Bemerkenswert ist dabei, dass eine Fokussierung auf die Stimme die Trefferquote erhöht, während die Fokussierung auf das Gesicht zu einer höheren Rate von Fehlidentifikationen führt. Insgesamt sind die Zusammenhänge jedoch schwach.

#### 4 Diskussion

Es konnte gezeigt werden, dass der Verdächtige in einer Wahlgegenüberstellung verglichen mit den Alternativpersonen eine erhöhte Wahrscheinlichkeit hat, identifiziert zu werden. Der psychische Status des Verdächtigten und die damit verbundenen körperlichen Reaktionen und Verhaltensmerkmale sind also eine potenzielle Informationsquelle in dieser Situation. Mit einer Steigerung der Identifikationsrate von über 10% muss dabei von praktischen Konsequenzen dieses Ergebnisses ausgegangen werden.

Die Vorteile dieses Laborexperimentes bestehen im Wesentlichen in der Konstanthaltung beziehungsweise Standardisierung von Einflüssen auf die an der Wahlgegenüberstellung teilnehmenden Personen, der parallelisierten Erstellung mehrerer Wahlgegenüberstellungen und der Existenz einer Kontrollbedingung. Alle diese Vorteile konnten im Rahmen dieser Studie erfolgreich umgesetzt werden. Wie bei vielen Laborexperimenten wird die bessere Kontrollierbarkeit aber mit einem gewissen Verlust an Wirklichkeitsnähe erkaufte. Vor allem

bei einer so praxisrelevanten Fragestellung muss auf die Faktoren eingegangen werden, die bei einer Übertragung der Ergebnisse der Studie auf die Praxis zu beachten sind. Erstens ist die Verwendung von Videos von Wahlgegenüberstellungen anstelle echter Wahlgegenüberstellungen zu nennen. Der Nutzen der Verwendung von Videos für die praktische Durchführung des Experiments liegt auf der Hand. Abgesehen von der etwas anderen Situation des Betrachtens eines Videos im Gegensatz zum direkten Betrachten von Menschen ist das primäre Problem hier die Bildauflösung des Videos. Es lässt sich mit gängigen technischen Mitteln kein Video erzeugen, das so detailliert wie eine direkte Beobachtung ist. Das Video bietet dem Betrachter also im Allgemeinen weniger Information als eine direkte Beobachtung. Allerdings spricht dies dafür, dass der Verdächtigeneffekt eher unter- als überschätzt wurde.

Zweitens ist auf die Parallelität in der psychischen Situation von realen Verdächtigten und den Gegenüberstellungsteilnehmern dieser Studie einzugehen. Beiden Situationen ist gemein, dass der Ausgang der Wahlgegenüberstellung hoch relevant für die Personen ist, eine Identifikation negativen Charakter hat und die Nicht-Identifikation einen hohen positiven Wert darstellt. Allerdings muss man zugestehen, dass für einen realen Verdächtigten die drohende negative Sanktion im Fall einer Identifikation im Vordergrund steht, während für die Teilnehmer des Experimentes der Gewinn im Fall einer Nicht-Identifikation hervorgehoben war. Auch wenn manches dafür spricht, dass dies vergleichbare Effekte produziert, so kann hiervon jedoch nicht mit letzter Sicherheit ausgegangen werden. Ein weiterer Unterschied ist, dass obwohl DM 500 für einen Studenten ein großer Gewinn sind, sie doch sicherlich keine so lebensentscheidende Bedeutung wie ein drohender Gefängnisaufenthalt haben. Dieser Unterschied sollte allerdings lediglich zu einer konservativeren Schätzung des Effekts führen, da die an der Wahlgegenüberstellung teilnehmenden Personen somit weniger Stresssymptome zeigen sollten.

Drittens ist die unterschiedliche Situation von Augenzeugen und unseren Versuchspersonen zu nennen. Das Experiment belegt lediglich, dass der psychologische Status des Verdächtigten eine potenzielle Informationsquelle ist und damit ein unfairer Bias in das Verfahren eingebaut ist. Inwieweit Augenzeugen diese subtile Informationsquelle aber nutzen, ist damit natürlich noch nicht geklärt.

Viertens deuten die korrelativen Zusammenhänge zu den Fragebogen-Variablen darauf hin, dass möglicherweise Merkmale der Stimme eine Rolle bei dem Verdächtigeneffekt spielen. Da vermutlich in den meisten realen Wahlgegenüberstellungen sprachliche Äußerungen nicht vorkommen, ist diese Rolle in zukünftigen Studien genauer abzuklären.

## **5 Ausblick**

Ziel weiterer Forschung muss es sein, Möglichkeiten zu finden, den Verdächtigten-Effekt auszuschalten oder doch zumindest auf ein Minimum zu reduzieren, um die Gerechtigkeit von Wahlgegenüberstellungen zu verbessern. Um dieses Ziel zu erreichen, müsste zunächst die Ursache des Verdächtigten-Effekts näher untersucht werden. In dieser Studie wurde davon ausgegangen, dass der Verdächtigten-Effekt durch Stresssymptome und/oder den Versuch, sich absichtlich unauffällig zu verhalten, hervorgerufen wird. Interessant wäre es hier, diesen Stress psychologisch und physiologisch zu messen und ihn mit dem Normalzustand des Verdächtigten, vor allem aber mit dem Stressniveau der Alternativpersonen zu vergleichen. Eine verwandte Fragestellung ist, ob mit der Stärke der potenziellen positiven Sanktion beziehungsweise mit der Stärke einer negativen Sanktion auch die Größe des Verdächtigten-Effekts und die Stärke der ihm zugrunde liegenden psychologischen und physiologischen Reaktionen steigen. Ferner sollte untersucht werden, ob tatsächlich die Betonung positiver Sanktionen im Nicht-Identifikationsfall (also auf den Geldgewinn in der vorliegenden Studie)

ähnliche psychologische und physiologische Folgen wie die Fokussierung auf negative Sanktionen im Identifikationsfall (wie in realen Gegenüberstellungen) nach sich zieht. Dies wäre wichtig, weil es voraussichtlich schwierig bleiben wird, mit starken negativen Sanktionen in Laborexperimenten zu arbeiten.

Ein zusätzlicher Ansatzpunkt für weitere Forschungen ergibt sich daraus, dass in der hier vorgelegten Studie nur der potenzielle Bias durch den Status des Verdächtigten belegt wurde. Es sollte untersucht werden, ob der Verdächtigten-Effekt auch im Falle einer vorhandenen Gedächtnisrepräsentation Auswirkungen auf die Entscheidung des Augenzeugen hat. Für den Fall, dass der Verdächtige tatsächlich der beobachtete Täter ist, könnte der Verdächtigten-Effekt die korrekte Identifizierungsrate noch steigern. Für den Fall, dass der Verdächtige nicht der Täter ist, könnte der Verdächtigten-Effekt zu einem Anstieg der falschen Identifizierungen des Verdächtigten führen. Auch wenn vermutet werden könnte, dass bei einer starken Gedächtnisrepräsentation die Auswirkungen des Verdächtigten-Effekts entsprechend gering sind, so könnten sie doch bei einer schwachen Gedächtnisrepräsentation eine Rolle spielen.

Letztlich sollte untersucht werden, ob bei alternativen Arten der Wahlgegenüberstellung der Verdächtigteneffekt ebenfalls auftritt. Interessant könnte hier vor allem die von vielen Forschern schon lange geforderte sequenzielle Wahlgegenüberstellung sein, bei der die an der Wahlgegenüberstellung teilnehmenden Personen nacheinander und nicht gleichzeitig präsentiert werden (Lindsay & Wells, 1985; für eine Übersicht über weitere Formen der Gegenüberstellung siehe Lindsay & Pozzulo, 1999). Ferner könnte ein vermutlich stark vermindertes Auftreten bei Photo-Wahlgegenüberstellungen untersucht werden.

Zusammenfassend lässt sich also konstatieren, dass die Größe des Verdächtigten-Effekts für die Praxis der Wahlgegenüberstellungen Konsequenzen hat, die vermieden werden sollten. Doch ist es zu diesem Zeitpunkt noch zu früh, um Empfehlungen für die Praxis zu geben, da zwar die Existenz des Verdächtigten-Effekts nachgewiesen werden konnte, Methoden zu seiner Vermeidung aber erst noch durch weitere Forschung gefunden werden müssen.

## Literatur

- Bierbaumer, N. & Schmidt, R. F. (1996). *Biologische Psychologie*. Berlin: Springer.
- Connors, E., Lundregan, T., Miller, N. & McEwan, T. (1996). *Convicted by juries, exonerated by science: Case studies in the use of DNA evidence to establish innocence after trial*. Alexandria, VA: National Institute of Justice.
- Cutler, B. L., Penrod, S. D. & Martens, T. K. (1987). The reliability of eyewitness identifications: The role of system and estimator variables. *Law and Human Behavior*, 11, 233-258.
- Devlin, Lord (1976). *Report to the Secretary of State for the Departmental Committee on Evidence of Identification in Criminal Cases*. London: Her Majesty's Stationary Office.
- Ekman, P. (1992). *Telling Lies. Clues to deceit in the marketplace, politics, and marriage*. New York: W. W. Norton & Company.
- Ekman, P. & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist*, 46, 913-920.
- Fabian, T., Stadler, M. & Wetzels, P. (1995). The „authenticity error“ in real lineup procedures - Effects of suspect-status and corresponding psychological dissimilarities between target person and distractors: An experimental study. In G. Davies, S. Lloyd-Bostock, M. McMurrin & C. Wilson (Hrsg.), *Psychology, Law, and Criminal Justice* (S. 29-38). Berlin: Walter de Gruyter.
- Greene, E. & Loftus, E. F. (1984). Solving the eyewitness problem. *Behavioral Sciences and the Law*, 2, 395-406.

- Lindsay, R. C. L. & Pozzulo, J. D. (1999). Sources of eyewitness identification error. *International Journal of Law and Psychiatry*, 22, 347-360.
- Lindsay, R. C. L. & Wells, G. L. (1985). Improving eyewitness identifications from lineups: Simultaneous versus sequential lineup presentation. *Journal of Applied Psychology*, 70, 556-564.
- Perrez, M., Laireiter, A.-R. & Baumann, U. (1998). Streß und Coping als Einflußfaktoren. In U. Baumann & M. Perrez (Hrsg.), *Lehrbuch Klinische Psychologie – Psychotherapie* (S. 277-305). Bern: Verlag Hans Huber.
- Stadler, M., Fabian, T. & Wetzels, P. (1992). *Wiedererkennen des Täters oder Identifizieren des Beschuldigten?* (*Bremer Beiträge zur Psychologie 1/92*). Bremen: Universität Bremen, SG Psychologie.
- Wells, G. L. (1978). Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology*, 36, 1546-1557.
- Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M. & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior*, 22, 603-647.

<sup>1</sup> Faktisch wurden die Teilnehmer gebeten, zu jeder Person der Wahlgegenüberstellung eine Wahrscheinlichkeitseinschätzung (in %) unter Einhaltung der folgenden Regeln zu geben: (a) Es mußte immer genau eine Person geben, die einen höchsten Wert erhielt; (b) die Prozentangaben mußten sich zu 100 % ergänzen. Dieses Vorgehen erlaubt sowohl die Auswertung als Wahlentscheidung (d.h. die Person mit dem höchsten Prozentwert wird als Wahl des Teilnehmers gewertet) als auch eine differenziertere Auswertung auf der Basis der Prozentwerte. Da die Auswertung keine wesentlichen Unterschiede erbrachte, beschränken wir uns hier auf die Wahlentscheidungen.

<sup>2</sup> Für diesen Test wurde eine Differenzvariable gebildet, indem von der Trefferquote für die Experimentalgegenüberstellungen die zugehörige Basiswahrscheinlichkeit abgezogen wurde (d.h. für die Experimentalsequenzen des Videobandes 1 die mittlere Trefferquote der Kontrollgegenüberstellungen des Bandes 2 und vice versa), deren Mittelwert dann gegen Null getestet wurde. Der etwas testschwächere Vergleich der Trefferquoten von Experimental- und Kontrollbedingung ergab aber ebenfalls einen signifikanten Unterschied.