

Forensisch-Psychiatrische Klinik

Abschlussbericht für das  
Bundesamt für Justiz

# Die Basler MIPS- Studie

Prognostisch und therapeutisch relevante messbare Indikatoren bei pädosexuellen  
Sexualstraftätern (MIPS)

29.04.2016 / Revision: 17.01.2017

Studienleitung:	Dr. phil. Marlon O. Pflüger
Studienoberarzt:	Dr. med. Thorsten Spielmann
Laborleitung Neurophysiologie:	Gunnar Deuring
Computervisualisierung:	Dr. Patrick Lemoine
Wissenschaftliche Mitarbeiterin:	Coralie Boillat
Forschungsleitung:	PD Dr. phil. Timm Rosburg
Laborleitung:	Prof. Dr. Ralph Mager
Forschungsdirektor:	Prof. Dr. Marc Graf



## Inhaltsverzeichnis

Abstract.....	5
Dank.....	7
1. Einleitung .....	7
2. Methode .....	10
2.1 Ein-/Ausschlusskriterien.....	10
2.2 Rekrutierung .....	11
2.3 Stichprobe.....	12
2.3.1 Demographische Daten.....	13
2.3.2 Familienstand und Geschwister.....	15
2.3.3 Suchtmittel- und Alkoholkonsum.....	15
2.3.4 Sexuelle Orientierung .....	16
2.4 Ablauf .....	19
2.5 Analysestrategie .....	20
2.6 Klassifikationsalgorithmus.....	21
2.7 Konzeption des Berichts .....	22
3. Resultate.....	23
3.1 Klinik .....	23
3.1.1 Compliance während der Untersuchung.....	23
3.1.2 Kritische Lebensereignisse.....	23
3.1.3 Erlebter Missbrauch und Vernachlässigung im Kindesalter .....	23
3.1.4 Krankheitsanamnese.....	24
3.1.5 Persönlichkeitsprofil.....	24
3.1.6 Persönlichkeitsstörungen.....	25
3.1.7 Psychopathie.....	27
3.1.8 Zusammenfassung.....	27
3.2 Neuropsychologie .....	28
3.2.1 Diskrimination von CTL und CSO auf Basis der Neuropsychologie .....	29
3.2.2 Diskrimination von CSEM und CSA auf Basis der Neuropsychologie .....	30

3.3 Neurophysiologie .....	32
3.3.1 Go/NoGo-Paradigma .....	32
3.4 Implizite Verfahren .....	38
3.4.1 Semantic Misattribution-Paradigma (SMP) .....	39
3.4.2 Impliziter Assoziationstest (IAT).....	44
3.4.3 Betrachtungszeit (Viewing Time, VT) und Bildbewertung.....	47
3.4.4 Choice Reaction Time-Aufgabe (CRT) .....	50
3.4.5 Virtuelle Exposition .....	51
3.4.5.1 Verhaltensdaten .....	54
3.4.5.2 Elektrodermale Aktivität .....	55
3.4.5.3 Kontrolle der Müdigkeit (EEG) .....	57
3.4.5.4 Habituation .....	57
3.4.5.5 Blickverhalten (Fixationen) .....	58
3.4.5.6 Alternative Auswertungsmethode (Musterextraktion) .....	63
3.5 Gesamtklassifikation .....	66
3.5.1 Vorhersage der sexuellen Präferenz (Klassifikation von CTL und CSO).....	66
3.5.2 Vorhersage des Risikoverhaltens (Klassifikation von CSEM und CSA) .....	68
4. Diskussion.....	70
4.1 Schwierigkeiten bei der Versuchsdurchführung.....	70
4.2 Diskussion der Resultate .....	71
4.3 Hypothesenprüfung.....	78
4.4 Praktische Relevanz der Resultate .....	79
Literatur .....	86
Anhang.....	93
Abkürzungsverzeichnis .....	93
Glossar .....	95
Verwendete Tests .....	99
Kongressbeiträge und Publikationen über die MIPS-Studie .....	101

## **Abstract**

Messbare Indikatoren für die Beurteilung von Sexualstraftätern mit einer pädophilen Orientierung fehlen weitgehend. Trotz bisheriger Forschungsbemühungen ist die Befundlage zur Neurobiologie und Ontogenese der Pädosexualität noch unbefriedigend und weitere Forschung, sowohl zur Klärung der Grundlagen als auch für die klinische Anwendung in der Begutachtung und Therapie, sind dringend angezeigt. Dies gilt insbesondere, da pädophile Präferenz als wichtiger Risikofaktor für wiederholten Kindesmissbrauch gilt.

Um Rückschlüsse auf die pädosexuelle Präferenz ziehen zu können, hatte die MIPS-Studie das Ziel Straftäter mit pädosexueller Ausrichtung psychiatrisch, neuropsychologisch und neurophysiologisch zu charakterisieren und so festzustellen in welchen Eigenschaften sich diese Tätergruppe von nicht-straffälligen Kontrollen mit adultsexueller Präferenz (CTL) unterscheidet. Darüber hinaus wurde nach Unterschieden zwischen Kontakt-Straftätern (Kindesmissbrauch, CSA) und Nicht-Kontakt-Straftätern (Konsum illegaler Internetpornographie, CSEM) gesucht, um (Risiko-) Indikatoren für Missbrauchsverhalten zu identifizieren.

In unserer Studie wurden 21 CTL und 43 Straftäter (22 CSA, 21 CSEM) untersucht. Neben gängigen neuropsychologischen und neurophysiologischen Verfahren wurden auch sogenannte implizite Verfahren und ein eigens konstruiertes Expositionsparadigma (VE) eingesetzt, das in eine virtuelle Umgebung implementiert war.

Die Resultate unserer Studie zeigen, dass CSA-Probanden im Vergleich zu den CTL-Probanden durch Minderleistungen charakterisiert waren, die auf ein hirnorganisches Entwicklungsdefizit hinweisen. CSEM-Probanden boten neuropsychologisch und neurophysiologisch ebenfalls ein defizitäres Bild, jedoch in geringerer Ausprägung und Konsistenz als die CSA-Gruppe. Die Vorhersage der sexuellen Präferenz gelang mit ähnlicher Präzision wie in vergleichbaren Studien, jedoch konnte keine Kombination von Verfahren die Vorhersagbarkeit darüber hinaus steigern. Mittels einer Kombination aus einem kognitiven Test und elektrodermalen Mustern aus der VE konnten hingegen sehr gute Vorhersagen des Risikoverhaltens gewonnen werden. Die dafür kritischen virtuellen Szenarien enthielten Merkmale, die mit spezifischen Charakteristika der Sexualstraftätergruppen in Verbindung gebracht werden konnten. Damit konnte nachgewiesen werden, dass die VE ein grosses Potential für die Vorhersage der pädosexuellen Präferenz und des Risikoverhaltens hat.

In Hinblick auf das Ziel Risiko- und Therapie-Verlaufsbeurteilungen zu verbessern, können neuropsychologische Untersuchungen zur Bestimmung des Straftätertyps und implizite Verfahren zur Bestimmung der sexuellen Präferenz empfohlen werden. Die Entwicklung portabler Untersuchungssysteme ist dazu notwendig.

## **Abstract (en français)**

Des indicateurs objectifs mesurables pour l'évaluation de délinquants sexuels avec une orientation pédophile manquent largement. Malgré des efforts de recherches scientifiques antérieures, l'état des connaissances en neurobiologie et ontogenèse des pédosexuels reste encore insatisfaisant; d'autres recherches scientifiques sont nécessaires afin de clarifier les bases et pour l'utilisation clinique dans le cadre d'expertises et de thérapies. Cela est particulièrement vrai compte tenu du fait que la préférence pédophile constitue un facteur de risque important pour les abus répétés d'enfants.

Pour être en mesure de tirer des conclusions sur la préférence pédosexuels, l'étude MIPS a eu pour objectif de caractériser d'un point de vue psychiatrique, neuropsychologique et neurophysiologique les pédosexuels délinquants et ainsi de déterminer quelles propriétés différencient ce groupe de délinquants du groupe-contrôle de non-délinquants ayant une préférence sexuelle adulte (CTL). En outre, on a cherché les différences entre les délinquants qui ont un contact direct (abus sur mineurs, CSA) et indirect (pornographie illicite sur Internet, CSEM) pour identifier des indicateurs de risque d'abus sexuel sur mineur.

Dans notre étude, 21 personnes de contrôle (CTL) et 43 délinquants (22 CSA, 21 CSEM) ont été examinés. En plus des méthodes neuropsychologiques et neurophysiologiques standards, des méthodes dites implicites et un paradigme d'exposition au travers d'un environnement virtuel conçu spécialement ont été mis en œuvre.

Les résultats de notre étude montrent que les sujets du groupe CSA ont été caractérisés par des performances réduites par rapport aux sujets-contrôles, ce qui indique un déficit de développement du cerveau. Les personnes du groupe CSEM présentaient du point de vue neuropsychologique et neurophysiologique également un profil déficient toutefois dans une plus faible mesure et cohérence par rapport à ceux du groupe CSA. La prédiction des préférences sexuelles a réussi avec une précision semblable à des études similaires effectuées, mais aucune combinaison de méthodes n'aurait en outre augmenté la prévisibilité. Grâce à une combinaison d'un test cognitif et de mesures galvaniques durant l'exposition virtuelle, une très bonne prédiction de comportements à risque a pu se faire. Les scènes de réalité virtuelle contenaient des caractéristiques qui pouvaient être associées à des caractéristiques spécifiques des groupes de délinquants sexuels. Il est ainsi prouvé que le dispositif d'exposition virtuelle a un grand potentiel de prédiction des préférences pédophiles et des comportements pédosexuels.

Concernant l'objectif d'amélioration de l'évaluation des risques et des évolutions thérapeutiques, des examens neuropsychologiques pour déterminer le type de délinquant et des méthodes implicites pour déterminer les préférences sexuelles peuvent être recommandés. Le développement de systèmes d'examen portables est cependant nécessaire.

## **Dank**

An dieser Stelle möchte sich das gesamte Team der MIPS-Studie ganz herzlich bei allen Probanden bedanken.

Uns war stets bewusst, dass es gerade mit dem heiklen Forschungsschwerpunkt der Pädophilie nicht selbstverständlich war, diese wertvolle Unterstützung zu erhalten.

Für das grosse Vertrauen, welches es seitens der Teilnehmer braucht, um sich in eine Studie zu begeben und gegenüber fremden Menschen sein Leben mit bisweilen sehr persönlichen Erfahrungen zu offenbaren, wollen wir uns ganz herzlich bedanken.

Uns war es ein Anliegen eine Studie durchzuführen, welche grosse Teile der Schweiz einbezieht und verschiedene Bereiche des Strafrechts berücksichtigt. Deshalb auch unseren herzlichen Dank an die Staatsanwaltschaften, die Vollzugsbehörden, die Therapiezentren und die niedergelassenen Psychiater und Psychologinnen, welche uns geholfen haben, Teilnehmer für die Studie zu rekrutieren und die Studie auf diese Weise erfolgreich durchzuführen.

## **1. Einleitung**

Sexueller Kindesmissbrauch ist ein weltweit verbreitetes Problem mit schwerwiegenden Folgen für die Betroffenen. Eine 2009 erschienene Metaanalyse (Pereda, Guilera, Forns, & Gómez-Benito, 2009) hat versucht, die Prävalenz von sexuellem Kindesmissbrauch zu erfassen. Diese Analyse, zusammengefasst aus 65 epidemiologischen Studien durchgeführt in 22 Ländern, ergab eine mittlere globale Prävalenz von 7.9% unter Männern und 19.7% unter Frauen, die vor dem Erreichen des 18. Lebensjahrs einmal oder häufiger sexuell missbraucht worden sind. In der Schweiz kam es von 1984 bis 2012 gemäss den Zahlen des Bundesamts für Statistik (BFS) zu 9'975 Verurteilungen wegen sexuellen Handlungen mit Kindern (Art. 187 StGB) und zu 10'161 Verurteilungen wegen illegaler Internetpornografie (Art. 197 StGB). Dabei hat laut KOBK (Koordinationsstelle zur Bekämpfung der Internetkriminalität) der Vertrieb von illegaler Internetkinderpornografie in den vergangenen Jahren massiv zugenommen. Gerichtliche Entscheide über Anordnungen von Verwahrungen und therapeutischen Massnahmen für Straftäter stützen sich immer auf forensisch-psychiatrische Gutachten. Diese Gutachten, welche sich auch zur Prognose äussern, basieren grösstenteils auf subjektiven Verhaltens- und Therapiebeobachtungen sowie auf aktuarischen Prognoseinventaren, wie beispielsweise dem Static-99 (Harris, Phenix, Thornton, & Hanson, 2003).

Über messbare und somit objektivierbare Indikatoren bei Gewaltstraftätern wurde ausführlich publiziert (Volavka, 2008). So fanden sich zum Beispiel vermindertes 5-HIAA (Serotonin-Metabolit) in der zerebrospinalen Flüssigkeit (Linnoila et al., 1983), oder reduzierte präfrontale graue Substanz bei Personen mit hohem „psychopathy“-Score in der Psychopathie Checkliste-Revidierte Version (PCL-R) (Raine A., Lencz T., Bihle S., LaCasse L., & Colletti P.,

2000). Vergleichbare Befunde für Sexualstraftäter, insbesondere für Sexualstraftäter mit einer pädophilen Orientierung fehlen weitgehend oder die Resultate sind heterogen. Trotz weltweit grossem Forschungsinteresse an der Operationalisierung von Pädophilie ist eine weitere Verbesserung der Befundlage, was die neurobiologischen und psychopathologischen Grundlagen dieser sexuellen Orientierungsstörung anbetrifft, wünschenswert. Insbesondere in Anbetracht dessen, dass eine pädophile Präferenz als wichtiger Risikofaktor für wiederholten Kindsmisbrauch gilt (Hanson & Morton-Bourgon, 2005) was nahelegt, dass eine Pädophilie auch das Risiko für erstmaligen Kindsmisbrauch erhöht (Mokros, Osterheider, & Nitschke, 2012).

Dabei zeigen sich in der Forschung unterschiedliche methodische Ansatzpunkte die in der Direktheit des Ansatzes variieren. So finden sich Studien, die mittels Phallometrie (physikalische Messung der Erektion bei visueller oder auditiver Stimulation) (Fisher, Cross, & Zuch, 1965) oder der Messung von peripher-physiologischen Reaktionen während einer Befragung (Lykken, 1984) versuchen, pädosexuelle Neigungen direkt zu erfassen. Andere Ansätze haben zum Ziel, die Präferenz über indirektere Methoden wie die Messung von Bildbetrachtungszeiten (Rosenzweig, 1942) oder Reaktionszeiten (Greenwald, McGhee, & Schwartz, 1998) zu erheben. Durch die Untersuchung von neurokognitiven Defiziten (Joyal, Black, & Dassylva, 2007; Kruger & Schiffer, 2011; Schiffer et al., 2008) wird versucht eher begleitende oder möglicherweise auch ursächliche Faktoren der Pädosexualität zu erfassen.

Übersichtsarbeiten (Fagan, Wise, Schmidt, & Berlin, 2002; Seto, 2008) zeigen, dass es sinnvoll und notwendig ist, verschiedene methodische Ansatzpunkte für die Erforschung und somit auch für die Prognoseeinschätzung von pädophilen Sexualstraftätern zu kombinieren.

Die MIPS-Studie ist eine Querschnittsstudie, in der nach Indikatoren zur Unterscheidung der sexuellen Präferenz in drei verschiedenen Gruppen gesucht wird. Die drei Gruppen bestehen aus pädophilen Sexualstraftätern, Konsumenten von Internetkinderpornographie und einer heterosexuellen Kontrollgruppe. Im Rahmen dieser Studie wurden nach einer psychiatrischen Exploration mittels verschiedenen Tests elektrophysiologische und neuropsychologische Daten erhoben, wobei explizite und implizite Testverfahren zur Anwendung kommen. Explizite Verfahren messen Verhalten, welches dem Probanden bewusst und willkürlich steuerbar ist. Implizite Verfahren beruhen auf Messungen, welche dem Bewusstsein weniger zugänglich sind, da sie automatisiertes Verhalten oder den Einfluss einer vergangenen Erfahrung auf eine darauf folgende Leistung messen. Der Vorteil, im Vergleich zu expliziten Tests, liegt in der geringen Verfälschbarkeit von implizitem Verhalten und den daraus resultierenden Befunden durch die getestete Person. In der MIPS-Studie wurden einerseits be-

kannte implizite Testverfahren, wie der „Implicit Association Test“ (IAT) und das „Semantic Misattribution“-Paradigma (SMP) angewendet, andererseits kam ein eigens konstruiertes Paradigma in Form eines „Strandspaziergangs“ in einer virtuellen Umgebung zum Einsatz.

Die Studie hat zum Ziel, Sexualstraftäter mit pädosexueller Ausrichtung (CSO = Child Sexual Offender) psychiatrisch, neuropsychologisch und elektrophysiologisch zu charakterisieren und festzustellen, hinsichtlich welcher Eigenschaften sie sich von nicht-straffälligen Kontrollen (CTL = Controls) unterscheiden. Innerhalb der CSO-Gruppe wird weiterhin zwischen Kontakt-Sexualstraftätern (CSA = Child Sexual Assaulter, Kindsmisbrauchler) und Nicht-Kontakt-Sexualstraftätern (CSEM = Child Sexual Exploitation Material, Konsum von Internetkinderpornographie) unterschieden, wobei angenommen wird, dass die CSA-Gruppe stärker von der Kontrollgruppe abweicht als die CSEM-Gruppe. Diese Gruppencharakteristika sollen einerseits inhaltlich bewertet werden, andererseits sollen diese Indikatoren genutzt werden, um die Gruppen zu diskriminieren und Risikofaktoren zu identifizieren. Hierbei wird insbesondere durch eine sogenannte Random Forest-Analyse geprüft, ob die Kombination verschiedener Indikatoren zu einer zuverlässigeren Diskriminierbarkeit führt. Die Relevanz des Messbarmachens der pädosexuellen Neigung liegt darin, die entdeckten Indikatoren legalprognostisch anzuwenden. Das primäre Ziel der MIPS-Studie ist folglich die Erforschung messbarer Indikatoren, um die Prognose in Bezug auf Therapieerfolg und Rückfallrisiko von pädosexuellen Straftätern nicht nur grob einschätzen sondern besser objektivieren zu können.

## 2. Methode

### 2.1 Ein-/Ausschlusskriterien

<b>Allgemeine Einschlusskriterien:</b>		
<ul style="list-style-type: none"> <li>- Männliches Geschlecht.</li> <li>- ≥ 18 – 55 Lebensjahre.</li> <li>- Ein IQ von &gt; 70, um die Fähigkeit zur korrekten Durchführung der unterschiedlichen Tests sicherzustellen.</li> <li>- Deutsch als Muttersprache oder sehr gute Deutschkenntnisse, um das Verständnis der Instruktionen sowie die Durchführung der verbalen Tests zu gewährleisten.</li> <li>- Eine uneingeschränkte Einwilligungsfähigkeit.</li> </ul>		
Einschlusskriterien CTL:	Einschlusskriterien CSA:	
<ul style="list-style-type: none"> <li>- Eine heterosexuelle Präferenz.</li> </ul>	<ul style="list-style-type: none"> <li>- Das Vorhandensein einer „relevanten pädophilen Ansprechbarkeit“. Dies erlaubt uns auch Teilnehmer zu untersuchen, bei welchen die Therapeuten eine pädophile Präferenz bejahen, jedoch offiziell (noch) keine Diagnose nach dem DSM-IV/5 gestellt wurde.</li> <li>- Eine Verurteilung oder konkrete Anhaltspunkte für einen Kindesmissbrauch (Art. 187 StGB), um Teilnehmer welche geständig sind, jedoch (noch) nicht verurteilt wurden miteinschliessen zu können.</li> <li>- Einen Low-Risk-Bereich, d.h. ein niedriges unmittelbares Risiko für sexuellen Kindesmissbrauch innerhalb des nächsten halben Jahres, um die Studie mit der notwendigen Sicherheit ambulant durchführen zu können.</li> </ul>	
	Einschlusskriterien CSEM:	
	<ul style="list-style-type: none"> <li>- Eine strafrechtliche Verurteilung nach Art. 197 StGB oder konkrete Anhaltspunkte für den Konsum von Kinderinternetpornographie.</li> <li>- Eine adult heterosexuelle Präferenz, um eine möglichst homogene Stichprobe zu erhalten.</li> </ul>	
<b>Allgemeine Ausschlusskriterien:</b>		
<ul style="list-style-type: none"> <li>- (Andere) akute psychiatrische Störungen nach DSM-IV/ICD-10.</li> <li>- Systemische/körperliche Erkrankungen oder Medikamente, welche die Leistungsfähigkeit für die Untersuchung beeinträchtigen.</li> <li>- Neurologische Erkrankungen im letzten halben Jahr.</li> <li>- Aktueller schädlicher Konsum psychotroper Substanzen (ausser Nikotin).</li> </ul>		
Ausschlusskriterien CTL:	Ausschlusskriterien CSA und CSEM:	
<ul style="list-style-type: none"> <li>- Sexuelle Vorlieben für Kinder.</li> <li>- Vorstrafen bezüglich sexueller Übergriffe oder Gewaltdelikten.</li> </ul>	<ul style="list-style-type: none"> <li>- Eine aktuelle Inhaftierung, ohne Möglichkeit polizeilicher Begleitung, da die Studiendurchführung nur in unseren Räumlichkeiten möglich war.</li> <li>- Eine antiandrogene Behandlung.</li> </ul>	

## 2.2 Rekrutierung

Die Pädophilie oder pädophile Interessen sind aktuell ein hitzig diskutiertes Thema, bei dem die gesellschaftliche Haltung allgemein mit einer eindeutig negativen emotionalen Komponente einhergeht. Aus diesem Grund war zu erwarten, dass die Rekrutierung der Indexgruppen, aber auch der Kontrollgruppe, sich eher schwierig gestalten würde. Wir versuchten dem mittels einer sogenannten Coverstory entgegenzuwirken, welche das Studienziel partiell verblindete. Diese Coverstory informierte darüber, dass das Forschungsprojekt die unterschiedliche Wahrnehmung von Männern, Frauen und Kindern in unserer Gesellschaft untersucht. Die Probanden wurden zusätzlich darüber informiert, dass es zu Beginn aus methodischen Gründen nicht möglich sein würde, sie vollständig über die Studienziele zu informieren.

Die Kontrollgruppe wurde mittels zwei Zeitungsinseraten (20 Minuten, Basler Zeitung) im Mai 2013 rekrutiert und innerhalb der vier Folgemonate untersucht. Für die Indexgruppen stützte sich unser interkantonal angelegtes Rekrutierungskonzept neben klinikinternen Aufrufen auf eine breit angelegte Information von Institutionen, Staatsanwaltschaften, Vollzugsbehörden, Bewährungsdiensten, Vollzugseinrichtungen und niedergelassenen Fachkollegen (beginnend im September 2013). Unsere Vernetzung ermöglichte es uns, zahlreiche Adressaten persönlich anzusprechen. Die Rückmeldungen waren insgesamt positiv, jedoch erwies sich die Rekrutierung der Indexgruppe trotzdem als schleppend und zeitintensiv: die einzelnen Einrichtungen und Fachkollegen mussten in regelmässigen Zeitabständen erneut kontaktiert und an unsere Studie erinnert werden. Bis zu einer tatsächlichen Terminierung waren (aus Gründen wie Vorbehalte gegenüber der Coverstory, gering frequentierte Therapiesitzungen, wechselnde Ansprechpartner, etc.) weiterhin eine Vielzahl an Telefon- und Mailkontakte vonnöten, mit bisweilen zahlreichen Gesprächen.

Dank schriftlichen und telefonischen Erinnerungen versäumten lediglich drei Kontrollprobanden und ein Indexproband den Untersuchungstermin, wobei mit dem betreffenden Indexprobanden ein neuer Termin vereinbart werden konnte. Tabelle 1 dokumentiert den Verlauf der rekrutierten Studienteilnehmer.

Die MIPS-Studie wurde am 24.02.2014 von PD Dr. Elmar Habermeyer und PD Dr. Andreas Mokros (beide Psychiatrische Universitätsklinik Zürich, Klinik für Forensische Psychiatrie) einer wissenschaftlichen Begutachtung (Audit) unterzogen. Alle Verfahren der Studie und das zugrundeliegende Rationale wurden dabei geprüft und positiv bewertet. Die wichtigsten Empfehlungen der Gutachter waren:

1. Die Grösse der drei Probandengruppen von  $N = 20$  zu priorisieren und bei Nichterreichen dieser Probandenzahlen in den Gruppen CSEM und CSA ein Oversampling in der CTL-

Gruppe vorzunehmen, d.h. diese bis zur Summe der beiden anderen Gruppen nach zu rekrutieren.

2. Zur Erhöhung der Teilnehmerzahl in den Indexgruppen die homosexuelle Orientierung als Ausschlusskriterium für die CSEM-Gruppe aufzuheben, und die geforderte Pädophilie-Diagnose nach DSM-IV-TR für die Probanden der CSA-Gruppe in eine „relevante pädophile Ansprechbarkeit“ zu lockern.

3. Die eingesetzten Fremdbeurteilungsverfahren (SKID-II, PCL:SV) einem Test der Inter-Rater-Reliabilität zu unterziehen.

Diese und andere Empfehlungen bezüglich der eingesetzten Verfahren und deren Auswertung wurden in der Folgezeit umgesetzt und formten die endgültige Fassung der Studie.

Tab. 1: Anzahl teilgenommener Probanden über den Zeitraum der Datenerhebung.

	Juni 2013	November 2013	Juni 2014	November 2014	Mai 2015	Dezember 2015
CTL	0	22	23	23	23	23
CSA	0	2	7	14	16	22
CSEM	0	3	7	10	10	21

### 2.3 Stichprobe

Angelegt war die Untersuchung mit einer Stichprobengrösse von  $n = 20$  pro Versuchsgruppe. Diese Zielgrösse wurde für alle drei Gruppen leicht übertroffen (CSA = 22, CSEM = 21, CTL = 21). Zwei weitere Kontrollprobanden wurden aus den Analysen ausgeschlossen und hier nicht berichtet, weil sie die Einschlusskriterien nicht erfüllten (Hinweise auf Polytoxikomanie bzw. Verdacht auf ein hirnorganisches Defizit). Ein CSEM-Proband brach die Studienteilnahme aus persönlichen Gründen nach dem ersten Vormittag ab. Seine Daten wurden gemäss dem Studienprotokoll gelöscht.

In der Regel nahmen alle Probanden an allen geplanten Untersuchungen teil. Bei einem CSEM-Probanden musste die Untersuchung nach dem ersten Tag wegen fehlender Medikation und darauf resultierender Incompliance (starker Unruhezustand) abgebrochen werden. Für diesen Probanden wurden einzig die Daten aus der allgemeinen psychiatrischen Exploration und aus den Fragebögen in die Analyse aufgenommen. Für einen CSA-Probanden

fehlen die Daten aus den Fragebögen, da er diese am Untersuchungstag nicht mit sich brachte und er wenige Tage nach der Untersuchung inhaftiert wurde.

### 2.3.1 Demographische Daten

Tab. 2: Demographische Daten: Medianwerte mit den jeweiligen Minimal- und Maximalwerten pro Gruppe und Unterschiede zwischen den drei Gruppen.

Demographie	CSA (N = 22)	CSEM (N = 21)	CTL (N = 21)	Statistik <sup>1</sup>	
				CSO vs. CTL	CSA vs. CSEM
	Md (range)	Md (range)	Md (range)		
Alter	36.5 (18-55)	36.0 (21-53)	26.0 (18-49)	$\chi^2 = 4.25^*$	$\chi^2 = 0.20$
Ausbildungsjahre	12.0 (10-17)	13.0 (10-23)	13.0 (10-21)	$\chi^2 = 0.63$	$\chi^2 = 0.60$
Berufsprestige	39.2 (23.8-64.4)	38.2 (23.8-64.4)	48.4 (29.4-53.5)	$\chi^2 = 2.10$	$\chi^2 = 0.85$
Händigkeit	86.7 (-100-100)	100.0 (71.4-100)	100.0 (-100-100)	$\chi^2 = 0.30$	$\chi^2 = 7.86^{**}$
Verbaler IQ	102.0 (91-130)	112.0 (64-136)	104.0 (89-130)	$\chi^2 = 0.23$	$\chi^2 = 1.80$
Nonverbaler IQ	116.0 (94-135)	116.0 (66-138)	114.0 (102-131)	$\chi^2 = 0.21$	$\chi^2 = 0.17$
Nationalitäten	N	N	N		
Schweiz	19	15	14		
Deutschland	1	5	5		
Andere	1	1	2		

<sup>1</sup>Kruskal-Wallis Test, df = 1

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Die erhobenen demographischen Variablen zeigen signifikante Unterschiede im Alter beim Vergleich der Indexgruppen mit der Kontrollgruppe und in der Händigkeit beim Vergleich aller drei Gruppen. Der mittlere **Alterswert** der beiden Indexgruppen liegt zehn Jahre über dem mittleren Alterswert der Kontrollgruppe. Somit unterscheiden sich die beiden Gruppen der CSO hinsichtlich des Alters knapp signifikant von der Kontrollgruppe ( $p = .039$ ). Hinsichtlich der **Händigkeit** unterscheiden sich die drei Gruppen signifikant voneinander, was darauf zurückzuführen ist, dass sich in der Gruppe der CSA mehr Personen mit einer verminderten Dominanz der rechten gegenüber der linken Hand befinden.

Die **Intelligenz** der drei Gruppen ist vergleichbar und unterscheidet sich nicht signifikant. Die Werte aller drei Gruppen liegen im Normbereich. Die eher hohen Werte v.a. des nonverbalen Intelligenzquotienten (IQ) sind auf den Flynn-Effekt zurückzuführen. Der Flynn-Effekt bezeichnet die Tatsache, dass der gemessene durchschnittliche IQ, insbesondere der nonverbale IQ, seit Beginn des 20. Jahrhunderts um 5 bis 25 IQ-Punkte pro Generation gestiegen ist (Flynn, 1987). Die Eichung der Normstichproben der verwendeten Intelligenztests ist jedoch veraltet und wurde seitens der Herausgeber dieser Testverfahren noch nicht angepasst. Hinsichtlich der Nationalität ist die die Mehrheit der Studienteilnehmer Schweizer Nationalität. Die **Ausbildungsdauer** aller drei Gruppen ist vergleichbar.

Das **Ansehen für den Berufsstand** wurde anhand der National Opinion Research Center (NORC)-Liste (Featherman, Sobel, & Dickens, 1975) eingeschätzt, welche Prestigewerte für einzelne Berufsgruppen vergibt. Gemessen wurde der Berufsstatus zum Zeitpunkt der Untersuchung. Die Berufe (bei aktueller Arbeitslosigkeit der letzte ausgeübte bzw. erlernte Beruf) der CSEM-Probanden erreichten hierbei den tiefsten, die der Kontrollprobanden den höchsten Prestigewert ohne signifikante Unterschiede zwischen den drei Gruppen. Zwei Kontrollprobanden, sechs Probanden aus der CSEM-Gruppe und neun Probanden aus der CSA-Gruppe waren zum Zeitpunkt der Untersuchung arbeitslos oder befanden sich im Strafvollzug. Bezogen auf die Arbeitslosigkeit unterscheiden sich die drei Gruppen nicht signifikant voneinander.

### 2.3.2 Familienstand und Geschwister

In Tabelle 3 ist der **Beziehungsstatus** zum Zeitpunkt der Untersuchung abgebildet. Die drei Gruppen unterschieden sich nicht signifikant.

Tab. 3: Prozentangaben für den Beziehungsstatus zum Untersuchungszeitpunkt.

Beziehungsstatus	CSA	CSEM	CTL
	(N = 22)	(N = 21)	(N = 21)
	N (%)	N (%)	N (%)
Single	11 (50.0)	13 (61.9)	7 (33.3)
Partnerschaft	4 (18.2)	4 (19.0)	6 (28.6)
Verheiratet	3 (13.6)	2 (9.5)	8 (38.1)
Geschieden, Single	2 (9.1)	2 (9.5)	0 (0.0)
Geschieden, Partnerschaft	2 (9.1)	0 (0.0)	0 (0.0)

Mindestens ein **eigenes Kind** haben sechs Kontroll-, vier CSEM- und sieben CSA- Probanden, wobei die Anzahlen nicht signifikant voneinander abweichen. In der Kontrollgruppe sind sechs, in der CSEM-Gruppe zwei und in der CSA-Gruppe vier der Probanden Einzelkinder.

### 2.3.3 Suchtmittel- und Alkoholkonsum

Im Rahmen der klinischen Datenerhebung wurden Angaben zum aktuellen sowie zum **Suchtmittelkonsum** in der Lebensgeschichte erhoben. Bei der Untersuchung der Kontrollgruppe zeigte sich, dass neben der anamnestischen Erfragung zum Umgang mit Suchtmitteln ein Drogen-Urintest nötig war, um einen objektiven Wert für den aktuellen Konsum und dessen möglichen Einfluss auf die Testergebnisse zu erhalten. Eine Urinprobe wurde ab dem siebten Kontrollprobanden abgenommen (d.h. für die ersten sechs Kontrollprobanden liegt kein Drogen-Urintest vor. Darüber hinaus fehlen Urinproben bei vier CSA- und einem CSEM-Probanden). Die Urinprobe wurde auf die Substanzen Cannabis, Benzodiazepine, Amphetamine, Halluzinogene und Opiate getestet.

Bei vier Kontroll-, zwei CSA- und sechs CSEM-Probanden waren die Urinproben positiv für Cannabis und bei drei CSEM-Probanden positiv für Benzodiazepine. Hinweise auf akute Intoxikationen oder Entzugssymptome ergaben sich bei keinem Versuchsteilnehmer. In der Befragung hinsichtlich des Konsumverhaltens über die Lebenszeit gaben zehn Kontroll-, 15 CSEM-, sowie neun CSA-Probanden an, Drogen (Cannabis, Kokain/Speed, Ecstasy, andere

Amphetamine, Halluzinogene und/oder Opiate) einmalig probiert oder mehrmals konsumiert zu haben. Einen vergangenen Cannabiskonsum bejahten acht Kontroll-, 14 CSEM- und neun CSA-Probanden. Einen Kokain-/Speedkonsum bejahten zwei Kontroll-, sechs CSEM-, ein CSA-Proband, Ecstasykonsum zwei Kontroll-, ein CSEM-, zwei CSA-Probanden, der Konsum anderer Amphetamine drei Kontroll-, sieben CSEM- und zwei CSA-Proband und einen Konsum von Halluzinogenen zwei Kontroll-, drei CSEM-, sowie zwei CSA-Probanden. Opiatkonsum wurde von keinem Teilnehmer bejaht.

Über eine aktuelle **Medikamenteneinnahme** berichteten zwei Kontroll-, zehn CSEM- und neun CSA-Probanden. Bei den Kontrollprobanden handelte es sich dabei um Antiallergika (2), bei den CSEM-Probanden um Antidepressiva (6), Methylphenidat (2), Benzodiazepin (1) und Neuroleptika (1), bei den CSA-Probanden um Antidepressiva (3), Methylphenidat (1), Neuroleptika (3), Antiepileptikum (1), Sonstige (z.B. Antihypertensiva) (2).

Der allgemeine und aktuelle Umgang mit **Alkohol** wurde mit dem AUDIT-Fragebogen (Saunders et al., 1993) erhoben. Der Alkoholkonsum unterscheidet sich nicht zwischen den drei Gruppen. Ein kritischer Wert von acht Punkten, welcher auf einen problematischen Alkoholkonsum hinweist, wird bei vier Kontroll-, drei CSA- und vier CSEM-Probanden überschritten.

#### **2.3.4 Sexuelle Orientierung**

Zur Abbildung der **sexuellen Ausrichtung** (homo-, hetero, bisexuell) verwendeten wir die Kinsey-Skala (Kinsey, von Eckardt-Jaffé, von Eckardt-Jaffé, Martin, & Pomeroy, 1966). Dabei handelt es sich um eine siebenstufige Grafik, in der sich die Probanden selbst hinsichtlich ihrer sexuellen Interessen einstufen (0 = ausschliesslich heterosexuell; 6=ausschliesslich homosexuell). Die Kontrollgruppe gab dabei, bis auf einen Probanden mit einer Ausprägung von 1, eine ausschliesslich heterosexuelle Präferenz an. In der Gruppe der CSA gaben 15 Probanden an, ausschliesslich heterosexuell zu sein und sechs ein relevantes homosexuelles Interesse (Ausprägung  $\geq 3$ ) zu haben. Zwölf CSEM-Probanden berichteten ein ausschliesslich heterosexuelles Interesse, fünf eine relevante homosexuelle Ansprechbarkeit (Ausprägung  $\geq 3$ ).

Zur Erfassung **psychosexueller Merkmale und Verhaltensauffälligkeiten** verwendeten wir das Multiphasic Sex Inventory (MSI). Dieser Fragebogen wurde in den 80er Jahren zur Untersuchung von Sexualstraftätern entwickelt und seither nicht mehr aktualisiert. Neuere Formen sexueller Phänomene (z.B. Internetpornographie) werden demnach nicht erfasst. Da es sich jedoch um einen etablierten Fragebogen handelt, der weite Bereiche der Sexualität er-

fasst und er in vielen bisherigen Studien verwendet wurde, setzten wir ihn in unserer Studie ein.

Wie zu erwarten ergeben sich für die CSA-Gruppe in Bezug zu den anderen Gruppen die höchsten Werte in der Skala *Sexueller Missbrauch an Kindern*. Die Skala beinhaltet fünf Subskalen, welche verschiedene Phasen des Annäherungsverhaltens abbilden. Während die CSA- und CSEM-Gruppen in der Subskala für *Sexuelle Fantasien mit Kindern* einen identischen Wert erreichten, gab die Kontrollgruppe keine solchen Fantasien an ( $p < .001$ ). Hinsichtlich der Subskala *Suchverhalten* im Sinne von Herumschleichen, Anschleichen und Überredungstaktik erreichte die CSA-Gruppe den höchsten Wert und somit einen signifikant höheren Wert als die Kontrollgruppe ( $p = .033$ ). Zur CSEM-Gruppe unterscheidet sie sich nicht signifikant, ebenso wenig wie sich die CSEM- von der Kontrollgruppe unterscheidet. In den Subskalen *Sexueller Angriff auf Kinder*, *Erschwerter sexueller Angriff* und *Inzest* erzielte einzig die CSA-Gruppe erhöhte Werte, wohingegen die Werte der CSEM- und der Kontrollgruppe bei null liegen ( $p < .001$ ). Bezüglich des präferierten *Geschlechts* von Kindern zeigen die CSA- und CSEM-Gruppen im Vergleich zu der Kontrollgruppe ein signifikant höheres Interesse für Mädchen ( $p < .001$ ). Auch das Interesse für Jungen ist in den Indexgruppen signifikant höher als in der Kontrollgruppe, welche keinerlei sexuelles Interesse für Jungen und Mädchen zeigt. Die positiven Werte für die Skalen Missbrauch an Mädchen/Jungen bei den CSEM- und Kontrollprobanden ergeben sich daraus, dass diese Skalen auch Items wie „herumalbern, kitzeln, sich sexuell angezogen fühlen“ beinhalten. Hinsichtlich den im Test erfragten *Paraphilien* (*Fetischismus, Voyeurismus, obszöne Telefonanrufe, Fesselung und Züchtigung, Sadomasochismus*) zeigen sich für alle drei Gruppen keine relevant erhöhten Werte. Die Skala welche *Sozial erwünschtes Antwortverhalten* erfasst zeigt keine Gruppenunterschiede. Die Auswertung der *Lügenskala* zeigt, dass die Kontrollprobanden den Fragebogen ausschliesslich ehrlich beantwortet haben. Die Gruppe der CSA erreicht auf der Skala exakt den Wert, der im Manual für nicht behandelte Sexualstraftäter erwartet wird. Die CSEM-Gruppe kann nicht mit einem erwarteten Wert verglichen werden, da Internetpornographie im Fragebogen nicht erfasst wird, jedoch unterscheidet sie sich in Bezug auf die Lügenskala nicht signifikant von der CSA-Gruppe. Das Ausmass der *Kognitiven Verzerrung* ist bei beiden Indexgruppen vergleichbar: Kritische Items, welche als Signal für eine starke kognitive Verzerrung angesehen werden müssten, wurden keine bejaht.

Tab. 4: Multiphasic Sex Inventory: Medianwerte mit den jeweiligen Minimal- und Maximalwerten pro Gruppe und Unterschiede zwischen den drei Gruppen. Bemerkung: Scores sind zu Prozentwerten umgerechnet (% des maximalen Itemscores).

MSI	CSA		CSEM		CTL		Statistik <sup>1</sup>		
	(N = 22)		(N = 20)		(N = 21)		CSO vs. CTL	CSA vs. CSEM	
	Md	(range)	Md	(range)	Md	(range)			
Soziale Sexualerwünschtheit	75.5	(49-97)	83.0	(43-100)	83.0	(49-100)	$\chi^2 = 1.26$	$\chi^2 = 1.10$	
Lügenskala: Sexueller Missbrauch	38.5	(8-100)	65.5	(15-100)	100.0	(85-100)	$\chi^2 = 33.70^{***}$	$\chi^2 = 1.57$	
Kognitive Verzerrung und Unreife	27.5	(0-65)	25.0	(5-75)	10.0	(0-35)	$\chi^2 = 15.14^{***}$	$\chi^2 = 0.06$	
Sexueller Missbrauch an Kindern									
Phantasie	60.0	(0-90)	60.0	(10-100)	0.0	(0-10)	$\chi^2 = 28.95^{***}$	$\chi^2 = 0.53$	
Suchverhalten	30.0	(0-80)	20.0	(0-60)	20.0	(0-40)	$\chi^2 = 3.66$	$\chi^2 = 2.74$	
Sexueller Angriff/ Übergriff	61.5	(11-89)	0.0	(0-67)	0.0	(0-11)	$\chi^2 = 20.29^{***}$	$\chi^2 = 20.89^{***}$	
Erschwerter sexueller Angriff	33.0	(0-83)	0.0	(0-50)	0.0	(0-0)	$\chi^2 = 17.80^{***}$	$\chi^2 = 19.15^{***}$	
Inzest	37.5	(0-100)	0.0	(0-75)	0.0	(0-0)	$\chi^2 = 13.82^{***}$	$\chi^2 = 15.15^{***}$	
Total	44.0	(8-72)	19.5	(5-62)	5.0	(0-10)	$\chi^2 = 38.75^{***}$	$\chi^2 = 4.50^*$	
Missbrauch an Mädchen	50.0	(0-100)	33.0	(0-67)	0.0	(0-33)	$\chi^2 = 14.02^{***}$	$\chi^2 = 1.62$	
Missbrauch an Jungen	33.0	(0-100)	0.0	(0-100)	0.0	(0-33)	$\chi^2 = 10.49^{**}$	$\chi^2 = 0.95$	
Paraphilien									
Fetischismus	0.0	(0-56)	5.5	(0-44)	0.0	(0-22)	$\chi^2 = 0.62$	$\chi^2 = 1.46$	
Voyeurismus	11.0	(0-56)	5.5	(0-56)	0.0	(0-11)	$\chi^2 = 7.74^{**}$	$\chi^2 = 0.35$	
Obszöne Telefonanrufe	0.0	(0-50)	25.0	(0-50)	0.0	(0-75)	$\chi^2 = 1.32$	$\chi^2 = 10.62^{**}$	
Fesselung/ Züchtigung	0.0	(0-67)	17.0	(0-100)	17.0	(0-100)	$\chi^2 = 3.21$	$\chi^2 = 10.20^{**}$	
Sado-Masochismus	0.0	(0-20)	0.0	(0-50)	0.0	(0-60)	$\chi^2 = 2.64$	$\chi^2 = 4.89^*$	
Kritische Items	0.0	(0-12)	0.0	(0-50)	0.0	(0-50)	$\chi^2 = 0.98$	$\chi^2 = 1.81$	
Sexuelle Dysfunktionen									
Sexuelle Unzulänglichkeit	12.0	(0-62)	12.0	(0-50)	0.0	(0-12)	$\chi^2 = 11.41^{***}$	$\chi^2 = 0.21$	
Vorzeitige Ejakulation	25.0	(0-100)	0.0	(0-25)	0.0	(0-100)	$\chi^2 = 0.55$	$\chi^2 = 9.37^{**}$	
Physische Behinderung	0.0	(0-12)	0.0	(0-25)	0.0	(0-12)	$\chi^2 = 0.15$	$\chi^2 = 1.16$	
Impotenz	0.0	(0-25)	8.0	(0-50)	0.0	(0-8)	$\chi^2 = 12.05^{***}$	$\chi^2 = 7.31^{**}$	
Kritische Items	0.0	(0-12)	0.0	(0-25)	0.0	(0-25)	$\chi^2 = 0.02$	$\chi^2 = 0.00$	

<sup>1</sup>Kruskal-Wallis Test, df = 1

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

Für die Quantifizierung von **sexuellem Interesse** bei übergriffig gewordenen Sexualstraftätern verwendeten wir die revidierte Screening Scale for Pedophilic Interests (SSPI-2) (Seto, Stephens, Lalumière, & Cantor, 2015). Bei Bejahung einer der fünf Fragen wird jeweils der Wert 1 (ausgenommen Item 1 der Wert 2) und bei Verneinung der Wert 0 vergeben. Positive Werte sprechen für ein stärkeres Reagieren auf kindliche als auf erwachsene Stimuli (Seto & Lalumière, 2001). Die CSA-Gruppe erreicht hierbei folgende Mittelwerte (m): Item 1: „Offender has male victim.“ m = 0.61. Item 2: „Offender has more than one victim “ m = 0.43. Item 3: „Offender has a victim aged eleven or younger.“ m = 0.70 . Item 4: „Offender has an unrelated victim.“ m = 0.43. Item 5: „Offender admitted to or was charged with a child pornography offence.“ m = 0.48.

## 2.4 Ablauf

Tab. 5: Ablauf der zweitägigen Untersuchung.

Tag 1	Tag 2
Untersuchung 1	Untersuchung 1
Klinische Exploration:	
<ul style="list-style-type: none"> <li>- Demographische Daten</li> <li>- Familie</li> <li>- Krankheiten</li> <li>- Suchtmittel</li> <li>- Sexuelle Ausrichtung</li> <li>- Kritische Lebensereignisse</li> <li>- Persönlichkeitsprofil/-störung</li> <li>- Psychopathie</li> </ul>	Neurophysiologisches Verfahren: <ul style="list-style-type: none"> <li>- Impulsivität-/Antwortunterdrückung</li> </ul> Neurophysiologische/Implizite Verfahren: <ul style="list-style-type: none"> <li>- Impliziter Assoziationstest</li> <li>- Semantischer Misattributionstest</li> <li>- Virtuelle Exposition</li> </ul>
Untersuchung 2	Untersuchung 2
Neuropsychologische Untersuchung:	
<ul style="list-style-type: none"> <li>- Aufmerksamkeit</li> <li>- Arbeitsgedächtnis</li> <li>- Gedächtnis</li> <li>- Exekutive Funktionen</li> <li>- Risikoverhalten</li> <li>- Intelligenz</li> <li>- Motorik</li> </ul>	<ul style="list-style-type: none"> <li>- Aufdeckung der Coverstory</li> <li>- Bildbetrachtungszeitmessungen</li> </ul> Explizite Verfahren: <ul style="list-style-type: none"> <li>- Beurteilung der im Verlauf der Untersuchung wiederholt dargestellten Bilder (Erwachsene und Kinder)</li> <li>- Missbrauch und Vernachlässigung im Kindesalter</li> <li>- Psychosexuelle Merkmale und Verhaltensauffälligkeiten</li> </ul>

## 2.5 Analysestrategie

Die Datenanalyse erstreckt sich auf vier Analyseeinheiten, die jeweils gesondert dargestellt und auf ihrer Diskriminationsfähigkeit überprüft werden sollen. Diese Analyseeinheiten sind:

- a) klinische Daten
- b) neuropsychologische Daten
- c) neurophysiologischen Daten
- d) durch implizite Verfahren gewonnen Daten.

Jede Analyseeinheit gliedert sich dann in einen deskriptiven Abschnitt, indem die Gruppen verglichen und ihre Besonderheiten herausgestellt werden sollen und einen klassifikatorischen Abschnitt, in dem diejenigen Variablen selektiert werden, die eine optimale Trennung der Gruppen ermöglichen. Davon ausgenommen blieben aus methodologischen Gründen die klinischen Daten. Sowohl die Deskription als auch die Klassifikation erfolgen im Hinblick auf (1) einen Vergleich der Kontrollprobanden und einer gepoolten Gruppe der Indexprobanden (CSEM & CSA gemeinsam) und (2) einen Vergleich der Indexgruppen untereinander (CSEM versus CSA). Dieses Vorgehen folgt dem Ziel, Variablen zu extrahieren, die einerseits im Zusammenhang mit der sexuellen Präferenz stehen und andererseits zu prüfen, inwiefern Unterschiede zwischen Kontakt-Straftätern und Nichtkontakt-Straftäter herausgearbeitet und für eine Klassifikation nutzbar gemacht werden können. In einem abschließenden Schritt erfolgt die Gesamtklassifikation, in dem die jeweils optimalen Prädiktoren kombiniert einer Klassifikationsanalyse unterzogen werden und damit die wichtigsten Variablen des MIPS-Projektes identifiziert werden können (vgl. Abb. 1).

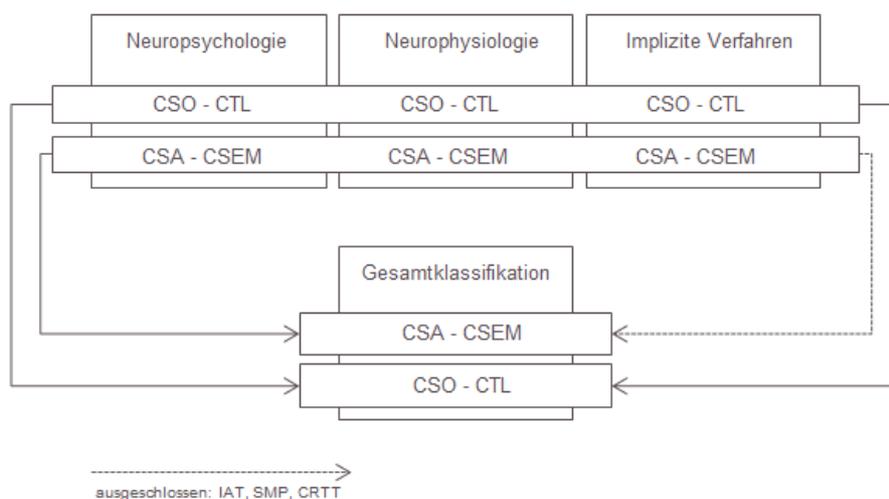


Abb. 1: Die Klassifikationsresultate der einzelnen Analyseeinheiten gehen in die abschließende Gesamtklassifikation ein.

## 2.6 Klassifikationsalgorithmus

Die Methoden zur Klassifikation beruhen im Kern auf modernen Maschinenlernalgorithmen, die sich das statistische Prinzip der grossen Zahlen zu eigen machen und in hohem Masse auf wiederkehrend gezogenen Zufallsstichproben beruhen. Dies macht diese Verfahren zwar enorm rechenintensiv, weshalb sie erst in jüngerer Zeit verstärkt genutzt werden, aber sie bieten eine Fülle von Vorteilen gegenüber herkömmlichen Verfahren. Im Vergleich zu traditionellen Klassifikationsverfahren in der Statistik stellen sie so gut wie keine Anforderung an die Beschaffenheit der zu analysierenden Daten. Sie sind nichtparametrisch, d.h. verteilungsfrei. Darüber hinaus sind sie geeignet, komplexeste Zusammenhänge zwischen den Daten zu identifizieren (nicht-lineare Zusammenhänge, Interaktionen höherer Ordnung) und schliesslich sind sie auch für hoch-dimensionale Datensätze (Datensätze mit deutlich mehr Variablen als Personen) die Methode der Wahl.

Der in der MIPS-Studie zur Anwendung kommende Klassifikationsalgorithmus beruht im Wesentlichen auf dem Verfahren des auf binären Entscheidungsbäumen beruhenden "*Random Forest*" (Breiman, 2001; Breiman, Friedman, Stone, & Olshen, 1984). Entscheidungsbäume sind geordnete und gerichtete Bäume, die in ihren Knoten Entscheidungsregeln beinhalten, die hierarchisch abgearbeitet werden und an deren Ende die Entscheidung für die Zugehörigkeit zu einer bestimmten Gruppe steht. Ein "*Random Forest*" ist nun ein ganzer Wald solcher Entscheidungsbäume, die jeweils nur aus einer kleinen zufällig selektionierten Anzahl von Variablen des Datensatzes und etwa 2/3 der Fälle konstruiert sind. Um nun zu prüfen, wie geeignet ein solcher Wald zur Klassifikation ist, werden die verbliebenen 1/3 Fälle klassifiziert, in dem jeder Baum ein Votum für die Zugehörigkeit zu einer Klasse abgibt. Das Mehrheitsvotum des gesamten Waldes ist dann das Gesamtklassifikationsergebnis. Neben dem Hauptergebnis der Klassifikation liefert der "*Random Forest*"-Algorithmus auch ein Mass für die Wichtigkeit der jeweiligen Variablen für die Klassifikation. In der MIPS-Studie wurde dieses Mass nutzbar gemacht und ein zweistufiges Verfahren entwickelt, um zu einem optimalen Klassifikationsergebnis zu kommen.

In einem ersten Schritt wurde der "*Random Forest*"-Algorithmus fünfzig Mal auf einen vorhandenen Datensatz angewandt und die Variablen nach ihrer durchschnittlichen Wichtigkeit absteigend geordnet. Dann wurden die unwichtigsten Variablen nach einem empirischen Kriterium gelöscht. In einem zweiten Schritt wurde dann durch eine gleichermassen fünfzigfache Berechnung jede einzelne verbliebene Variable danach geprüft, ob sie geeignet ist, den Klassifikationsfehler zu minimieren. Wenn sie dieses Kriterium nicht erreicht hatte, wurde sie eliminiert, so dass schliesslich mit einem minimalen Set von Variablen eine optimale Klassifikation erreichen werden kann (Genuer, Poggi, & Tuleau-Malot, 2010).

## **2.7 Konzeption des Berichts**

Im Vergleich zu einer rein wissenschaftlichen Veröffentlichung verzichten wir in diesem Bericht in der Regel auf eine Beschreibung von Effekten, die für die Differenzierung der Indexgruppe und Kontrollgruppe bzw. für die Differenzierung innerhalb der Indexgruppe nicht von Bedeutung sind. Zur besseren Lesbarkeit verzichten wir auch auf eine komplette Beschreibung statistischer Kennwerte. Daneben werden bestimmte technische Aspekte der Datenverarbeitung nur ansatzweise dargestellt.

Wissenschaftlich interessierte Leser werden daher ausdrücklich auf zukünftige Publikationen der Arbeitsgruppe verwiesen. Es ist möglich, dass in dem Abschlussbericht genannte statistische Kennwerte von denen, die in solchen Publikationen genannt werden, abweichen, insbesondere wenn andere statistische Verfahren angewandt werden. Daneben möchten wir darauf verweisen, dass Kennwerte biologischer Daten (wie EKPs oder die EDA) auf unterschiedliche Art und Weise extrahiert werden können, und sich so Abweichungen zu den hier berichteten Daten ergeben können.

### **3. Resultate**

Der hier vorgelegte Bericht beschreibt den Stand der Auswertung zum 30.04.2016.

#### **3.1 Klinik**

##### **3.1.1 Compliance während der Untersuchung**

Insgesamt war die Compliance seitens der Versuchsteilnehmer bei allen Teilen der Untersuchung gut. Bei einzelnen Probanden zeigte sich im Laufe der Untersuchung ein wachsendes Misstrauen bezüglich des Studienziels. Trotz anfänglichen Hinweises, dass einige Tests womöglich undurchsichtig erscheinen würden, reagierten einzelne ärgerlich oder verunsichert und waren besorgt um die Vertraulichkeit der Daten. Indem, soweit es die Coverstory zuließ, auf die Fragen eingegangen wurde und auf ein bevorstehendes Aufklärungsgespräch verwiesen wurde, konnte die Motivation und Konzentration der Probanden aufrechterhalten werden.

##### **3.1.2 Kritische Lebensereignisse**

Soziale Faktoren und persönliche Umstände wurden wiederholt mit Kindsmisbrauch in Verbindung gebracht (Ward & Beech, 2006). Stressoren aufgrund kritischer Lebensereignisse wurden anhand der Social Readjustment Rating Scale (SRRS; Holmes & Rahe, 1967) erhoben. In dieser Skala werden 43 Ereignisse wie Umschulung, Scheidung, Tod eines nahen Angehörigen, aber auch Hochzeit und persönliche Erfolgserlebnisse erfragt. Obschon die Gruppe der CSA den höchsten und die Gruppe der CSEM den tiefsten Stresswert aufweisen, unterscheiden sich die drei Gruppen nicht signifikant voneinander.

##### **3.1.3 Erlebter Missbrauch und Vernachlässigung im Kindesalter**

Die Literatur zeigt, dass sexuelle Missbrauchserfahrung in der Kindheit ein Risikofaktor für die Entstehung einer pädophilen Präferenz darstellt (Jespersen, Lalumière, & Seto, 2009). Der Childhood Trauma Questionnaire (CTQ) ermöglicht die retrospektive Erfassung von in der Kindheit erlebtem Missbrauch und erfahrener Vernachlässigung. Die drei untersuchten Gruppen weisen in den Rohwerten vergleichbare Werte in den einzelnen Subskalen (*emotionaler/physischer/sexueller Missbrauch, emotionale/physische Vernachlässigung, Lüge*) auf. Beim Zugrunde legen eines cut-off-Wertes für einen mindestens geringfügigen Missbrauch/Vernachlässigung ergeben sich jedoch signifikante Gruppenunterschiede in den kategorialen Häufigkeiten (erlebter vs. nicht-erlebter Missbrauch oder Vernachlässigung). Die beiden Indexgruppen berichten einen signifikant höheren Anteil an emotionalen ( $p = .010$ ) und physischen ( $p = .011$ ) Missbrauchserfahrungen, sowie einen signifikant höheren Anteil an physischer Vernachlässigung ( $p = .040$ ) als die Kontrollgruppe, wobei sich die CSA- nicht signifikant von der CSEM-Gruppe unterscheidet. Auffallend ist ein hoher prozentualer Anteil

emotionaler Vernachlässigung aller drei Gruppen ohne signifikante Unterschiede. Auch hinsichtlich eigener erfahrener sexueller Übergriffe zeigen sich keine signifikanten Unterschiede. Die Lügenskala weist in allen drei Gruppen einen Wert von Null auf, was für eine ehrliche Beantwortung der Fragen spricht.

Tab. 6: Childhood Trauma Questionnaire: Prozentualer Anteil derer, die mind. einen geringen Missbrauch/Vernachlässigung erfahren haben.

CTQ	CSA (N = 22)	CSEM (N = 20)	CTL (N = 21)	Statistik <sup>1</sup>	
				CSO vs CTL	CSA vs CSEM
Emotionaler Missbrauch	36.4%	42.1%	4.8%	$\chi^2 = 6.56^*$	$\chi^2 = 0.00$
Physischer Missbrauch	54.5%	35.0%	9.5%	$\chi^2 = 6.51^*$	$\chi^2 = 0.92$
Sexueller Missbrauch	40.9%	30.0%	19.0%	$\chi^2 = 1.14$	$\chi^2 = 0.17$
Emotionale Vernachlässigung	68.2%	65.0%	57.1%	$\chi^2 = 0.21$	$\chi^2 = 0.00$
Physische Vernachlässigung	59.1%	60.0%	28.6%	$\chi^2 = 4.20^*$	$\chi^2 = 0.00$

<sup>1</sup>Chi-Quadrat Test, df = 1

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

### 3.1.4 Krankheitsanamnese

Studien zeigen einen Zusammenhang einer pädosexuellen Präferenz mit Kopfverletzungen (Blanchard et al., 2002, 2003). Die Erhebung der Häufigkeit erfahrener Schädelhirntraumata ergab keine signifikanten Unterschiede zwischen den Gruppen. Als relevant wurde ein Schädelhirntrauma eingestuft bei einer Bewusstlosigkeit von mehreren Minuten oder wenn es mit neurologischen und/oder neuropsychologischen Symptomen einherging. Die Kontrollgruppe weist mit vier Probanden mit einem einmaligen Schädelhirntrauma den höchsten Wert auf (im Vgl. zu der CSEM-Gruppe mit zwei Probanden mit je einem Trauma und einem Probanden mit zwei Traumata, sowie der CSA -Gruppe mit zwei Probanden mit je einem Schädelhirntrauma in der Lebensgeschichte). Auch hinsichtlich allgemeiner schwerer Krankheiten oder Verletzungen in der Lebensgeschichte welche zu längeren Spitalaufenthalten führten unterscheiden sich die drei Gruppen nicht.

### 3.1.5 Persönlichkeitsprofil

Mögliche Auffälligkeiten im Persönlichkeitsprofil (Langevin, Paitich, Freeman, Mann, & Handy, 1978) wurden mit dem NEO-Persönlichkeitsinventar-Revidierte Version (NEO-PI-R) er-

hoben. Der NEO-PI-R ist ein Fragebogen zur Erfassung des Fünf-Faktoren-Modells der Persönlichkeit. Demnach lässt sich jede Persönlichkeit durch fünf Hauptdimensionen beschreiben: *Verträglichkeit*, *Gewissenhaftigkeit*, *Extraversion*, *Neurotizismus* und *Offenheit für Erfahrungen*. Die um das Alter bereinigten Daten zeigen, dass sich die beiden Indexgruppen signifikant auf den Skalen *Gewissenhaftigkeit* ( $p = .001$ ) und *Neurotizismus* ( $p = .001$ ) von der Kontrollgruppe unterscheiden. Hierbei weichen jeweils fünf Probanden der CSA- und CSEM-Gruppe auf der Skala Gewissenhaftigkeit ( $T < 40$ ), sowie sechs Probanden der CSA-Gruppe, vier Probanden der CSEM-Gruppe und ein Kontrollproband auf der Skala Neurotizismus vom Normbereich der Eichstichprobe des Fragebogens ab ( $T > 60$ ).

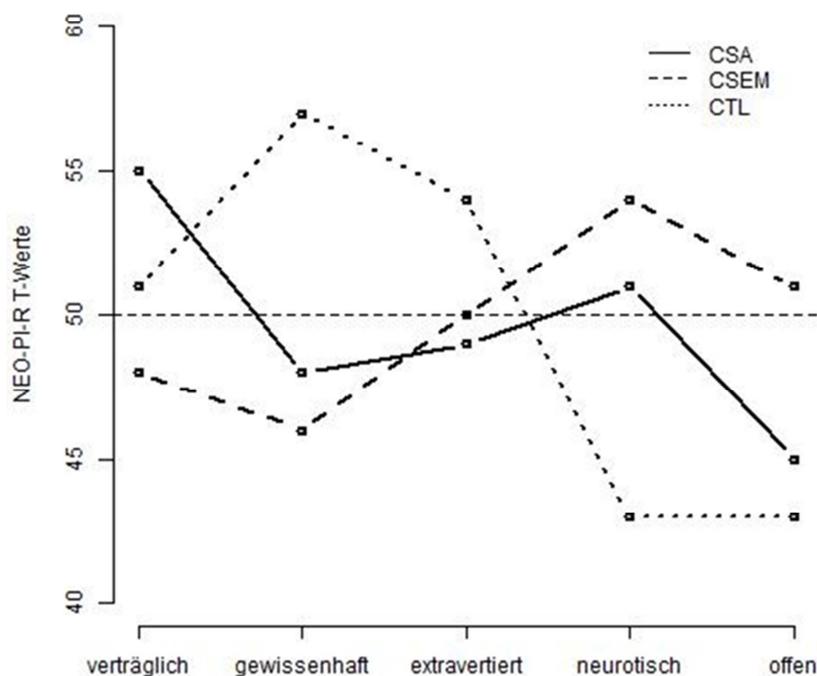


Abb.2: Alterskorrigierte Profile der drei Gruppen in den fünf Persönlichkeitsfaktoren des NEO-PI-R. Aufgeführt sind die T-transformierten Gruppenmittelwerte.

Die Kurzversion des Inventars der Persönlichkeit (IPO) misst den Schweregrad der strukturellen Beeinträchtigung in den Bereichen Identität, Abwehr und Realitätsprüfung. Die Medianwerte zeigen keine signifikanten Unterschiede zwischen den drei Gruppen.

### 3.1.6 Persönlichkeitsstörungen

Die Literatur zeigt einen Zusammenhang von Pädophilie mit Auffälligkeiten in der Persönlichkeit (Cohen et al., 2002). Beim Strukturierten Klinischen Interview für DSM-IV (SKID) Achse-II handelt es sich um einen Selbstbeurteilungsfragebogen der als Screening für

Merkmale einer möglichen Persönlichkeitsstörung dient. Persönlichkeitsbereiche, welche im Fragebogen einen cut-off-Wert überschreiten, werden in einem nachfolgenden strukturierten Interview genauer überprüft. Insgesamt erreichen sieben Probanden Werte, die nach der genaueren Abklärung in einem Interview auf eine Persönlichkeitsstörung hindeuten. Ein CSA-Proband erfüllt hierbei die Kriterien für drei (zwanghaft, paranoid, borderline) und ein weiterer CSA-Proband für zwei (borderline, antisozial) Persönlichkeitsstörungen. Die Kriterien für eine antisoziale Persönlichkeit erfüllen vier CSA-Probanden und zwei CSEM-Probanden. Innerhalb der Kontrollgruppe erreicht hingegen kein Proband Werte für eine Persönlichkeitsstörung.

Der SCL-90 ist ein Fragebogen, welcher anhand von 90 Fragen die subjektiv empfundene körperliche und psychische Beeinträchtigung innerhalb der letzten sieben Tage misst. Auf neun Skalen werden folgende Bereiche abgebildet: *Aggressivität/Feindseligkeit*, *Ängstlichkeit*, *Depressivität*, *Paranoides Denken*, *Phobische Angst*, *Psychotizismus*, *Somatisierung*, *Unsicherheit im sozialen Kontakt* und *Zwanghaftigkeit*. Ein Vergleich unserer Daten mit den Normwerten des Test-Handbuches zeigt, dass kein Proband gleichviel oder mehr Beschwerden angibt als 5% bzw. 10% der Normstichprobe, welche die meisten Beschwerden berichten. Erst bei einem Ausdehnen der Schwelle auf 15% geben acht Probanden gleichviel oder mehr Beschwerden an (4 CTL-, 3 CSEM-, 1 CSA-Proband). Drei Kontroll-, 14 CSEM- und neun CSA-Probanden erreichen auf mindestens einer Skala auffällige, d.h. signifikant ( $p < .05$ ) von der Norm abweichende, Werte. Die CSA-Gruppe erzielt hierbei die meisten Nennungen in den Skalen *Depressivität*, *Ängstlichkeit*, *Unsicherheit im sozialen Kontakt* und *Psychotizismus*. Die CSEM-Gruppe erreicht die höchsten Werte in den Skalen *Depressivität*, *Paranoides Denken* und *Phobische Angst*. Bei einem CSA-Probanden sind dabei alle und bei einem CSEM-Probanden acht der neun Skalen auffällig.

Beim Zusammenfassen der einzelnen Skalenwerte zu einem globalen Kennwert, welcher die durchschnittliche subjektiv empfundene körperliche und psychische Beeinträchtigung angibt, erreichen die beiden Indexgruppen signifikant höhere Werte als die Kontrollgruppe ( $p = .005$ ), wobei die CSEM-Gruppe am meisten belastende Symptome angibt. Untereinander unterscheiden sich die Indexgruppen nicht.

Tab. 7: Symptom-Check-Liste-90: Medianwerte mit den jeweiligen Minimal- und Maximalwerten pro Gruppe und Unterschiede zwischen den drei Gruppen.

SCL-90	CSA (N = 21)	CSEM (N = 21)	CTL (N = 21)	Statistik <sup>1</sup>			
				CSO CTL	vs.	CSA CSEM	vs.
	Md (range)	Md (range)	Md (range)				
Global Severity Index	0.28 (0.02-1.26)	0.37 (0.02-1.39)	0.17 (0.01-0.80)	$\chi^2 = 7.88^{**}$		$\chi^2 = 0.02$	

<sup>1</sup>Kruskal-Wallis Test, df = 1

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

### 3.1.7 Psychopathie

Studien zeigen einen Zusammenhang von pädosexuellem Missbrauchsverhalten mit Defiziten in der Empathiefähigkeit (Covell & Scalora, 2002) und psychopathischen Persönlichkeitszügen (Millon, Simonsen, Birket-Smith, & Davis, 2003). Die Psychopathie Checkliste: Screening Version (PCL:SV) dient der Diagnose einer psychopathischen Persönlichkeitsstörung. Eine Psychopathie zeichnet sich unter anderem aus durch Mangel an Empathiefähigkeit und Schuldgefühlen, promiske Sexualität, impulsives und manipulatives Verhalten, übersteigertes Selbstwertgefühl, pathologisches Lügen, polymorphe Kriminalität und/oder jugendliche Delinquenz. Der Medianwert über alle Skalen hinweg unterscheidet sich signifikant zwischen den beiden Indexgruppen und der Kontrollgruppe ( $p = .046$ ). Er befindet sich jeweils in einem sehr niedrigen Bereich, unterhalb eines kritischen Grenzwertes. Auf folgenden Skalen erreichen mindestens 30% oder mehr der Probanden aus den entsprechenden Gruppen einen mindestens leicht erhöhten Wert: *Übernimmt keine Verantwortung* (CSA), *Impulsiv* und *Schlechte Verhaltenssteuerung* (CSEM), *Fehlende Lebensziele*, *Verantwortungslos*, *Antisozial in der Adoleszenz* (CSA und CSEM).

### 3.1.8 Zusammenfassung

Die Gruppe der CSA berichtet am meisten kritische Lebensereignisse, wobei sich die Anzahl nicht relevant von den anderen beiden Gruppen unterscheidet. Beim Gruppieren der beiden Indexgruppen zeigt sich allerdings, dass diese signifikant häufiger Opfer von emotionalem und physischem Missbrauch sowie physischer Vernachlässigung in der Kindheit geworden sind, als die Kontrollgruppe. Die Betrachtung der Daten auf Ebene der Persönlichkeit kennzeichnet die beiden Indexgruppen als weniger gewissenhaft aber neurotischer als die Kontrollgruppe. Bei fünf CSA-Probanden und zwei CSEM-Probanden weisen die Daten auf mindestens eine Persönlichkeitsstörung hin, wohingegen kein Proband aus der Kontrollgruppe eine Persönlichkeitsstörung aufweist. Auffällig zeigen sich die beiden Indexgruppen auch

hinsichtlich ihrer subjektiv empfundenen körperlichen und psychischen Beeinträchtigung, die Kontrollgruppe bleibt hier unauffällig. Werte, welche auf eine psychopathische Persönlichkeit hinwiesen, werden von keinem Studienteilnehmer erreicht. Die einzelnen Werte befinden sich in einem sehr tiefen Bereich, weit von einem kritischen Grenzwert entfernt. Insgesamt zeigen die beiden Indexgruppen über alle Skalen hinweg aber signifikant höhere Werte als die Kontrollgruppe. Hinsichtlich erfahrener Schädelhirntraumata und/oder schwerer Krankheiten unterscheiden sich die drei Gruppen nicht.

Zusammenfassend ist festzuhalten, dass sich die beiden Indexgruppen zusammengenommen in einigen der erhobenen klinischen Parametern systematisch von der Kontrollgruppe unterscheiden, dass aber die Ausprägung dieser Eigenschaften in den Indexgruppen nicht auf ausgeprägte Störungen in anderen Bereichen als der Pädosexualität verweisen, sondern auch im Bezug zur Gesamtpopulation eher diskrete Abweichungen beschreiben. Zudem liess sich der Einfluss einiger in der Literatur diskutierter Faktoren, die als begünstigend für das Entwickeln einer Pädosexualität angesehen werden (eigener erfahrener sexueller Missbrauch während der Kindheit oder erfahrene Kopfverletzungen), nicht nachweisen. Unsere Annahme, dass die CSA-Gruppe stärker von der Kontrollgruppe abweicht als die CSEM-Gruppe, lässt sich anhand der klinischen Daten nicht bestätigen.

### **3.2 Neuropsychologie**

Die Neuropsychologie ist eine relativ junge wissenschaftliche Teildisziplin der Psychologie, deren Gegenstand das Verhältnis von Verhalten und Erleben zu den Strukturen und Funktionen des Gehirns ist. Dabei macht sie Gebrauch von nach psychometrischen Gütekriterien entwickelten Tests, die geeignet sind, distinkte Hirnfunktionen zu evaluieren und in Leistungskennwerte zu übersetzen.

Die neuropsychologische Untersuchung im Rahmen der MIPS-Studie fokussiert auf die Integrität von Hirnfunktionen im Bereich Aufmerksamkeit, Gedächtnis und Handlungsregulation (exekutive Funktionen). Ein besonderer Schwerpunkt bildet das sogenannte „Risk-Taking Behavior“ (Risikoverhalten) und die Impulskontrolle (die eigentlich gleichermassen Teilfunktionen der exekutiven Funktionen sind). Die Auswahl erfolgt vor dem Hintergrund der besonderen Vulnerabilität dieser Funktionsbereiche und bestehender Forschungsliteratur, die über Störungen in diesen Funktionsbereichen bei mit unseren Studienteilnehmern vergleichbaren Sexualstraftätern berichtet (Cantor et al., 2004; Kruger & Schiffer, 2011b; Schiffer & Vonlaufen, 2011). Wir erwarten dass die Impulsivität und Risikobereitschaft im Entscheidungsverhalten eine besondere Rolle im Zusammenhang mit den kognitiven Funktionen der CSO- und insbesondere der CSA-Gruppe spielt, da dies häufig auch im forensischen Kontext und lange schon in der Fachliteratur thematisiert wird (Finkelhor & Araji, 1986). Eine detaillierte

Auflistung der neuropsychologischen Verfahren kann dem Anhang: Verwendete Tests entnommen werden.

### 3.2.1 Diskrimination von CTL und CSO auf Basis der Neuropsychologie

Abb. 3 zeigt die neuropsychologischen Leistungsdaten der Gruppe der CSO aggregiert nach generellen Hirnfunktionen. Dabei wurden die einzelnen Werte an den Daten der Kontrollgruppe normiert und auf eine gemeinsame Skala transformiert, sodass die Leistungsfähigkeit in den unterschiedlichen Hirnfunktionen vergleichbar wird. Wie aus der Abbildung ersichtlich ist die Leistungsfähigkeit der Indexgruppe in den meisten Bereichen geringer als die der Kontrollgruppe. Jedoch sind nur die Gedächtnisleistungen und die exekutiven Funktionen signifikant davon betroffen (das 95%-Konfidenzintervall umschließt nicht die Null).

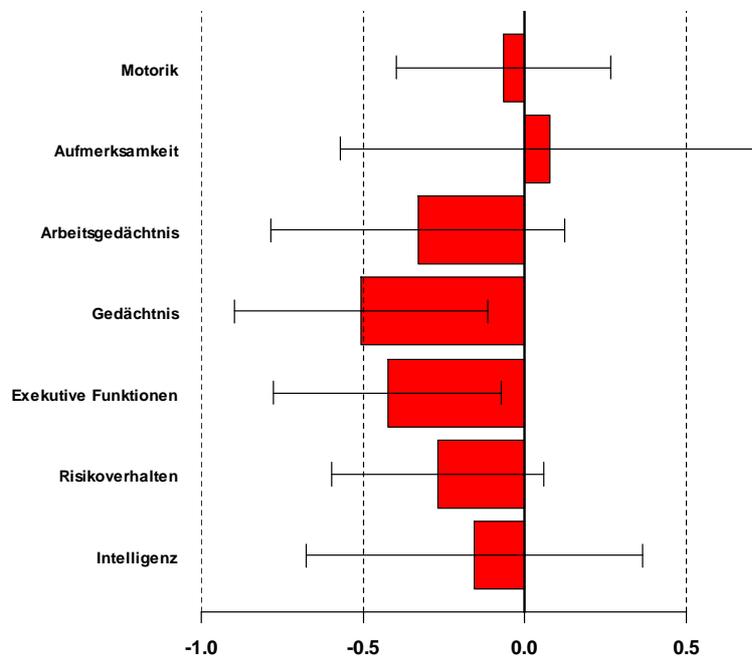


Abb. 3: Die Balken repräsentieren Abweichungen der Gruppe der CSO von Kontrollprobanden in Einheiten der Standardabweichung. Das 95%-Konfidenzintervall indiziert Gedächtnis und exekutive Funktionen als signifikant vermindert.

Die Klassifikationsanalyse indiziert fünf Variablen, die in 83% der Fällen eine korrekte Klassifikation von Kontrollgruppe und Indexgruppen ermöglichen. Die wichtigste Rolle nimmt dabei die Impulsivität im Rahmen einer Risikoentscheidung ein, gefolgt vom Abrufvermögen unmittelbar zuvor gelernter Wörter, zweier Variablen zum Arbeitsgedächtnis (Kurzzeitgedächtnis) und der Fähigkeit zum schnellen verschieben des Aufmerksamkeitsfokus. In zwei der fünf Bereiche (Arbeitsgedächtnis Kapazität, verschieben des Aufmerksamkeitsfokus) übertreffen die Indexgruppen die Kontrollgruppe, in den

verbleibenden drei anderen sind sie hingegen deutlich gemindert. Insgesamt entspricht damit das kognitive Profil unseren apriori getroffenen Annahmen und stimmt weitestgehend mit der Fachliteratur überein. Darüber hinaus zeigt die durch die Klassifikationsanalyse getroffene Auswahl der Impulsivität ebenfalls die Korrektheit der Hypothesen. Die Beobachtung, wonach nicht ausschliesslich Variablen ausgewählt wurden hinsichtlich derer die Indexgruppen Minderleistungen erbringen lässt darauf schliessen, dass die Zusammensetzung der Stichprobe (CSO-Gruppe) hinsichtlich der Hirnleistungsfähigkeit sehr heterogen ist.

### **3.2.2 Diskrimination von CSEM und CSA auf Basis der Neuropsychologie**

Eine getrennte Analyse der neuropsychologischen Leistungsfähigkeit nach CSEM- und CSA-Gruppe zeigt dass die Minderleistungen keineswegs gleichförmig sind sondern dass CSEM-Probanden entgegen der CSA-Probanden signifikant verminderte Gedächtnisleistungen aufweisen, während bei der Gruppe der CSA eher das Risikoverhalten im Vordergrund steht (vgl. 95%-Konfidenzintervalle). Bei den meisten anderen Funktionsbereichen zeigt sich ebenfalls eine Tendenz zu Minderleistungen. Insbesondere im Bereich des Arbeitsgedächtnisses, einem Kurzzeitspeicher und der exekutiven Funktionen, welche für eine zielgerichtete Verhaltensregulation im Allgemeinen erforderlich sind. Hier wird jedoch die Signifikanzgrenze nicht erreicht.

Aus der Klassifikationsanalyse gehen drei Variablen hervor, die es ermöglichen die Gruppen mit etwa mit 82%iger Genauigkeit korrekt zu klassifizieren. Zum einen handelt es sich hierbei um eine Variable, die für Irritierbarkeit bzw. das Unterdrücken von starken Verhaltenstendenzen kodiert. Zum anderen um eine Variable, die im Zusammenhang mit fehlerhaften Reaktionen unter hohen selektiven Aufmerksamkeitsanforderungen steht und schliesslich einer Variablen, die für die Neigung zu risikoreichen Entscheidungen kodiert. In allen drei Bereichen zeigen die CSA- eine geringere Leistungsfähigkeit als die CSEM-Probanden.

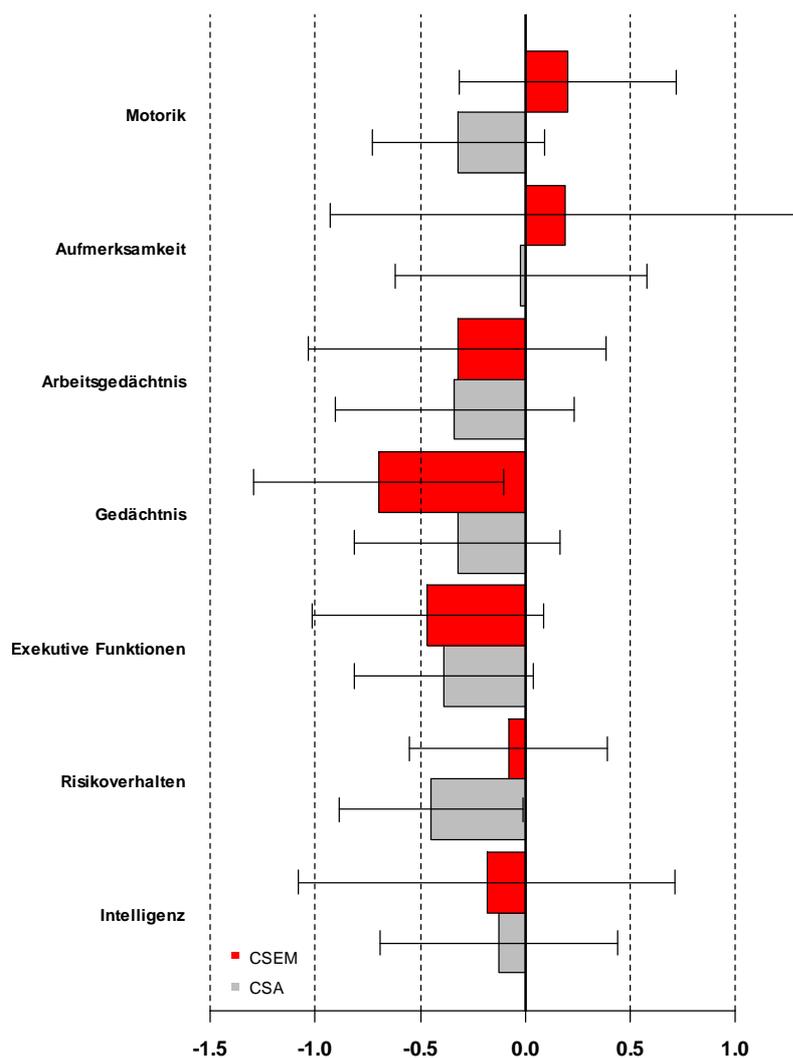


Abb. 4: Die Balken repräsentieren Abweichungen der CSEM- und CSA-Gruppe von Kontrollprobanden in Einheiten der Standardabweichung. Das 95%-Konfidenzintervall indiziert signifikante Verminderungen der Gedächtnisleistung der CSEM- und ein signifikant erhöhtes Risikoverhalten der CSA-Gruppe.

Abschliessend lässt sich sagen, dass in Übereinstimmung mit der Literatur kognitive (neuropsychologische) Minderleistungen ein charakteristisches Merkmal von Sexualstraftätern im Massnahmenvollzug sind, die auf entwicklungsbedingte Störungen hinweisen könnten. Wie von uns erwartet sind auch das Risikoverhalten und Impulsivität von grosser Bedeutung für die Verhaltensregulation der Indexgruppen.

### 3.3 Neurophysiologie

#### 3.3.1 Go/NoGo-Paradigma

##### Beschreibung der Untersuchung

Das sogenannte Go/NoGo-Paradigma ist ein klassisches Verfahren um Verhaltenskontrolle im Sinne von Impulsivität und Antwortunterdrückung zu erfassen. Die Probanden sollen in diesem Verfahren auf einen bestimmten visuellen Reiz, der am Bildschirm präsentiert wird, so schnell wie möglich mit einer motorischen Antwort – Tastendruck – reagieren („Go“). Dieser Reiz wird häufig gezeigt (80%), was zur Folge hat, dass der Proband eine Tendenz entwickelt, auf einen visuellen Reiz mit einer motorischen Antwort zu reagieren. In seltenen Fällen (20% der Durchgänge) wird aber ein anderer Reiz gezeigt und die motorische Antwort muss in diesem Fall unterdrückt werden („NoGo“). Abbildung 5 veranschaulicht den Ablauf des Verfahrens. Im vorliegenden Fall war der häufige „Go“-Reiz ein **M**, der seltene „NoGo“-Reiz ein **W**. Als motorische Reaktion war eine Taste mit der rechten Hand zu drücken. Die Anzahl der *entgegen* der Instruktion erfolgten Antworten (Tastendruck in einem „NoGo“-Durchgang = Fehlalarm) gilt als Mass dafür, wie gut jemand Impulskontrolle ausüben kann. Das häufige Ausbleiben von Antworten in „Go“-Durchgängen würde hingegen auf eine mangelnde Mitarbeit des Probanden schliessen lassen.

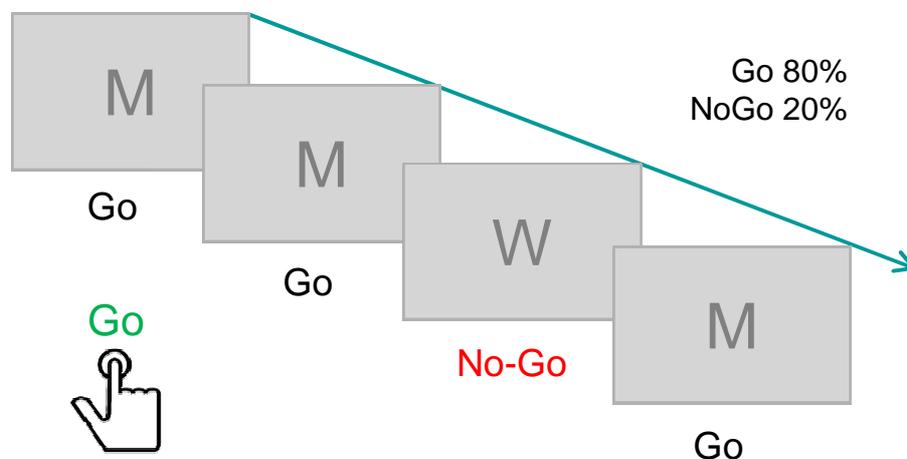


Abb. 5: Ablauf des Go/NoGo-Verfahrens. Der häufige „Go“-Reiz (M) und der seltene „NoGo“-Reiz (W) werden in einer quasirandomisierten Reihenfolge mit der Häufigkeit von 80% „Go“ und 20% „NoGo“ am Bildschirm präsentiert. Der Proband muss in der „Go“-Bedingung eine Taste drücken und die sich aufbauende Reaktionstendenz in der „NoGo“-Bedingung unterdrücken. Bei einem Auslassungsfehler (nicht gedrückt bei „Go“), wurde nach 500 ms ein unangenehmer Ton als Rückmeldung gegeben. Bei Fehlalarmen (gedrückt bei „NoGo“) gab es keine Rückmeldung des Fehlers.

Neben Verhaltensmassen (Reaktionszeiten und Treffer bzw. Fehlerraten) wurden während des Experiments ereigniskorrelierte Potentiale (EKPs) aufgezeichnet. Unter EKPs versteht

man den typischen Verlauf der elektrischen Spannung (Potenzial) an einer EEG-Elektrode (EEG = Elektroenzephalographie = Hirnstromableitung), bezogen auf den Zeitpunkt der Reizpräsentation oder den Zeitpunkt der motorischen Antwort. EKPs erlauben Rückschlüsse auf neuronale und damit assoziierte kognitive Prozesse, die während dieser Aufgabe eine Rolle spielen; speziell beim Go/NoGo-Paradigma Impulskontrolle und Fehlerverarbeitung.

Unsere Hypothesen in dieser Untersuchung waren, dass die Indexgruppen, insbesondere die CSA-Gruppe, in ihrer Impulskontrolle vermindert ist und daher erhöhte Fehlerraten im Vergleich zu Kontrollprobanden zu finden sind. Zudem wurde angenommen, dass bestimmte EKP-Komponenten, nämlich solche die mit Fehlererkennung, Feedbackverarbeitung und Kontextupdate assoziiert sind, in ihrer Amplitude vermindert sind.

## Ergebnis

Verhaltensdaten: Entsprechend der Hypothese, dass verminderte Impulskontrolle ein bezeichnendes Merkmal der Indexgruppen sei, wurde zunächst die Fehlerrate (Fehlalarm) in den ‚NoGo‘-Durchgängen ausgewertet. Die Gruppengrößen für die Auswertung waren wie folgt: CTL: N = 21; CSEM: N = 17; CSA: N = 21. Es zeigte sich, dass die Probanden der CSO-Gruppe signifikant mehr Fehlalarme aufwiesen (35.3%) gegenüber der CTL-Gruppe (26.9%) ( $F_{1,57} = 5.579, p = .02$ ), siehe Abbildung 6, links.

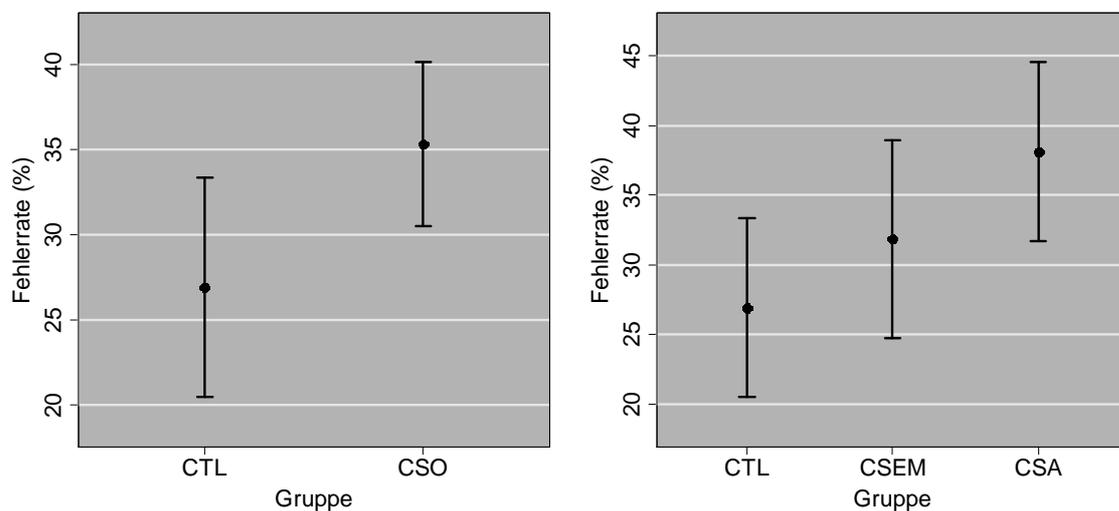


Abb. 6: Fehlerraten unter der ‚NoGo‘-Bedingung in den beiden Gruppen CSO und CTL (links), sowie aufgeschlüsselt in den Gruppen CTL, CSEM und CSA (rechts). Die Fehlerbalken beschreiben das 95% Konfidenzintervall.

Ein Blick auf die Kontraste der Gruppen CSEM und CSA gegenüber der Kontrollgruppe (Abbildung 6, rechts) offenbarte, dass vor allem die Probanden der CSA-Gruppe mit einer Feh-

lerrate von 38.1% für die Gruppendifferenz verantwortlich waren (CTL vs. CSA:  $t = 2.47$ ,  $p = .02$ ), wohingegen die CSEM-Gruppe mit 31.8% im Mittel dazwischen angesiedelt war und sich statistisch weder von der Kontrollgruppe noch der CSA-Gruppe unterschied.

Weiterhin wurden die Reaktionszeiten bei richtigen Reaktionen in der ‚Go‘-Bedingung und bei inkorrekten Reaktionen in der ‚NoGo‘-Bedingung (Fehlalarm) analysiert. Die Reaktionszeiten bei den Fehlalarmen gegenüber den ‚Go‘-Treffer-Durchgängen war signifikant verkürzt ( $F_{1,114} = 45.97$ ,  $p < .001$ ), was durchaus zu erwarten war, denn diese Fehlleistung ist in der Regel eine vorschnelle Reaktion, im Sinne einer nicht erfolgten Impulskontrolle, zurückzuführen. Die Reaktionszeiten unterschieden sich dabei in beiden Bedingungen jedoch weder zwischen den Gruppen CTL und CSO noch zwischen den Gruppen CSEM und CSA.

EKP-Daten: Abbildung 7 zeigt den typischen Verlauf eines EKPs, wie es unter der ‚Go‘-Bedingung bei richtiger Reaktion (‚NoGo‘-Treffer) gemessen werden kann. Die Form des Potenzialverlaufs bei einem einzigen Durchgang unterliegt situationsbedingten Störeinflüssen und individuellen Variationen. Um ein EKP zu erhalten werden die Potenzialverläufe vieler Durchgänge unter identischen Bedingungen (‚Go‘-Reiz gefolgt durch eine ‚Go‘-Antwort oder ‚NoGo‘-Reiz gefolgt durch eine ‚Go‘-Antwort) gemittelt. Die Beschreibung des EKPs erfolgt durch sogenannte Komponenten, d.h. über spezifische Ausschläge des Potenzials im negativen (‚N‘) oder positiven (‚P‘) Bereich in bestimmten Zeitbereichen („Latenzen“). Für den Vergleich zwischen den Gruppen wurden die Amplituden (in Mikrovolt,  $\mu\text{V}$ ) und Latenzzeiten (in ms) für die einzelnen Komponenten erfasst.

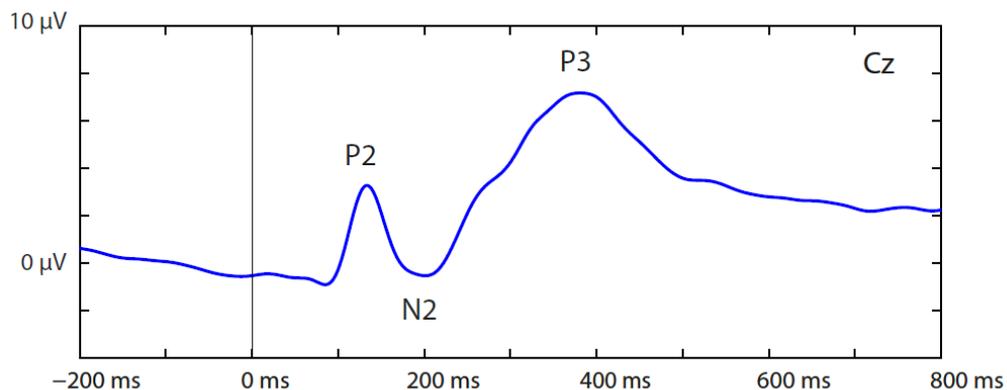


Abb. 7: Evoziertes Potenzial von allen Probanden der CTL-Gruppe, gemessen an der Elektrode Cz (Vertex) unter der ‚Go‘-Treffer-Bedingung. Deutlich unterscheidbar die Komponenten P2, N2 und P3.

In der Literatur werden die Komponenten N2 und P3 bei den Durchgängen mit richtiger Reaktion (‚Go‘ und ‚NoGo‘) im Zusammenhang mit den psychologischen Konstrukten der Reaktionshemmung und Impulskontrolle diskutiert. Die Analyse der Komponenten N2 und P3 un-

ter den Bedingungen mit richtiger Reaktion zeigte jedoch keine Gruppenunterschiede hinsichtlich der N2- und P3-Amplituden sowohl unter der ‚Go‘- als auch unter der ‚NoGo‘-Bedingung. Auch die N2-Latenzen in beiden Bedingungen waren nicht verschieden zwischen allen Gruppen. In der P3-Latenz unter der ‚Go‘-Treffer-Bedingung (nicht aber unter der ‚NoGo‘-Fehlalarm-Bedingung) fand sich ein Gruppeneffekt ( $F_{1,58} = 6.869$ ,  $p = .01$ ), der eine kürzere P3-Latenz in der Kontrollgruppe als in den Indexgruppen kennzeichnete, siehe Abbildung 8 (links). Abbildung 8 (rechts) illustriert die Positionierung der Indexgruppen untereinander und gegenüber der CTL-Gruppe. Es ergab sich ein ähnliches Bild wie bei den Fehlalarmen: Die P3-Latenz der CSA-Gruppe (404 ms) war signifikant länger als die der CTL-Gruppe (381 ms) ( $t = 2.7$ ,  $p = .01$ ). Die CSEM-Gruppe lag mit 397 ms dazwischen, der Unterschied der CSEM-Gruppe zu den beiden anderen war nicht signifikant.

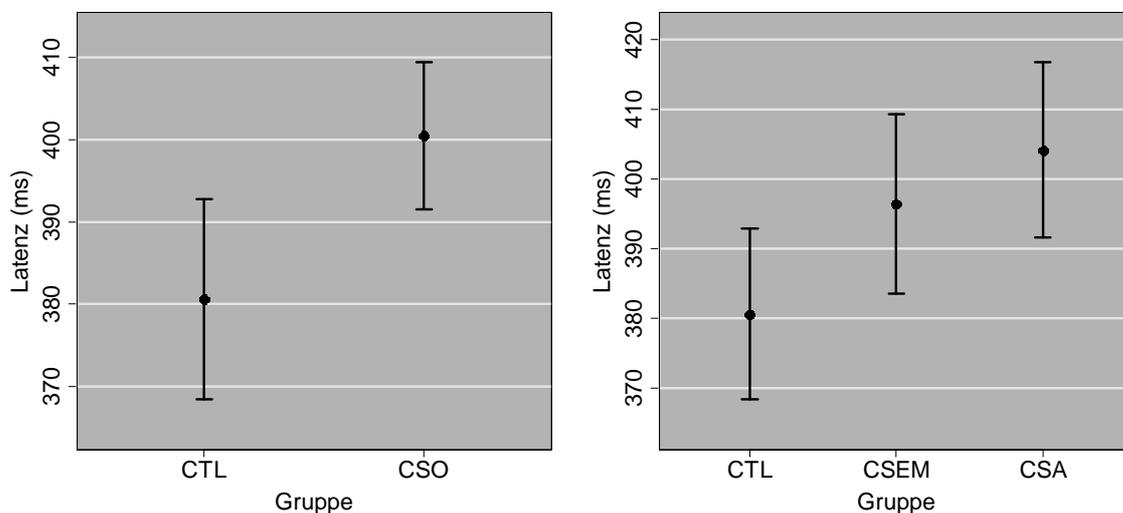


Abb. 8: P3-Latenz (ms) des EKPs unter der ‚Go‘-Treffer-Bedingung in den beiden Gruppen CSO und CTL (links), sowie aufgeschlüsselt in den Gruppen CTL, CSEM und CSA (rechts). Die Fehlerbalken beschreiben das 95% Konfidenzintervall.

Weiterhin gibt es nach einer falschen Reaktion in der ‚NoGo‘-Bedingung, bei welcher die Taste entgegen der Anweisung gedrückt wurde, die Komponenten Ne und Pe (‚-e‘ für ‚error‘), die auf die Fehlerverarbeitung nach dem erkannten Fehlalarm schliessen lassen. Diese Komponenten werden durch die Mittelung auf die erfolgte Antwort hin (und nicht, wie weiter oben berichtet, auf den präsentierten Reiz) berechnet. Es fanden sich keine Gruppeneffekte hinsichtlich der Latenzen von Ne oder Pe. Auch die Ne-Amplitude diskriminierte nicht zwischen den Gruppen CTL und CSO. Die Pe-Amplitude war jedoch mit 9.9  $\mu\text{V}$  in der CSO-Gruppe signifikant kleiner als in der CTL-Gruppe (13.7  $\mu\text{V}$ ) ( $F_{1,51} = 5.113$ ,  $p = .03$ ), siehe Abbildungen 9 (links) und 10. Ein weiteres Mal zeigt sich hier das Muster, nach dem die CSA-Gruppe einen Grossteil des Effekts der CSO-Gruppe bedingte, wie in Abbildung 9 (rechts)

dargestellt, sogar etwas akzentuierter. So war die mittlere Pe-Amplitude der CSA-Gruppe mit 8.1  $\mu\text{V}$  signifikant kleiner als die der CTL-Gruppe ( $t = 2.994$ ,  $p = .004$ ); auch der Unterschied zwischen der CSA-Gruppe und CSEM-Gruppe (11.9  $\mu\text{V}$ ) war marginal signifikant ( $t = 1.998$ ,  $p = .051$ ). Kein Unterschied fand sich hingegen zwischen den Gruppen CSEM und CTL. Das Alter der Probanden hatte keinen oder einen die Gruppeneffekte verstärkenden Einfluss auf alle für das Go/NoGo-Paradigma weiter oben beschriebenen Analysen.

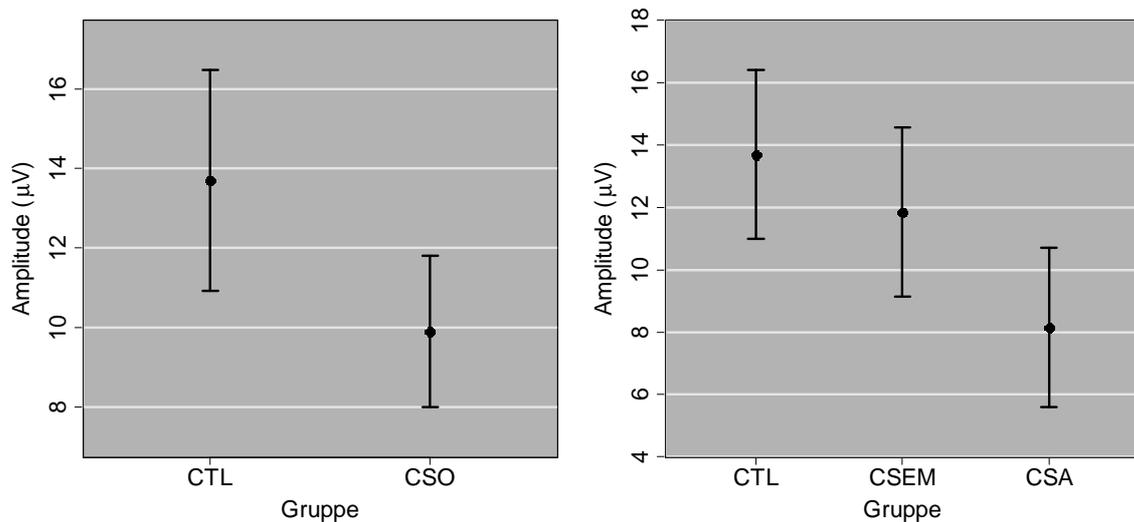


Abb. 9: Die Amplitude der Pe-Komponente bei falscher Reaktion in der ‚NoGo‘-Bedingung in den Gruppen CSO und CTL. Die Fehlerbalken beschreiben das 95% Konfidenzintervall.

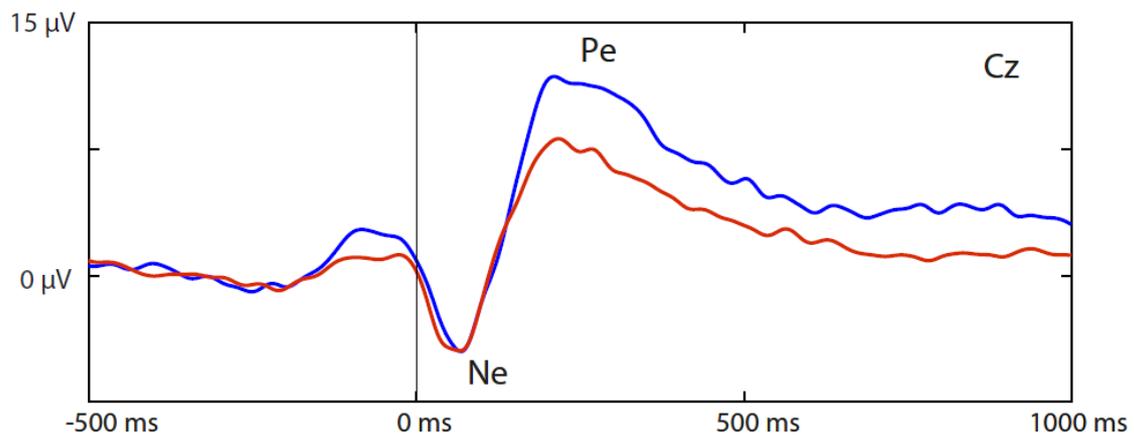


Abb. 10: Pe-Komponente in der ‚NoGo‘-Fehler-Bedingung bei den Gruppen CTL (blau) und CSO (rot).

### Interpretation

Entsprechend unserer Hypothesen fanden wir in den Indexgruppen eine erhöhte Fehlerrate auf NoGo-Stimuli, was als Kennzeichen einer verminderten Impulskontrolle gewertet werden kann. Dieser Verhaltenseffekt war stark genug, um die CSO-Gruppe als Ganze von der CTL-Gruppe zu diskriminieren, allerdings nicht ausreichend, um selbiges bei der Diskriminierung

der CSA- von der CSEM-Gruppe, deren Fehlerrate geringer war, zu erreichen. Der Verhaltensbefund einer Fehlerrate auf ‚NoGo‘-Stimuli wird elektrophysiologisch ergänzt durch eine verlängerte P3-Latenz in der CSO-Gruppe (und hier insbesondere in der CSA-Gruppe), was sich in das Bild einer neurokognitiven Beeinträchtigung bei diesen Probanden fügt. Die P3-Latenz wird als kennzeichnend für die Geschwindigkeit angesehen, mit der ein Stimulus klassifiziert und evaluiert wird (Polich, 2007). Die P3-Latenz nimmt beispielsweise mit zunehmendem Alter, aber auch im Rahmen bestimmter neurodegenerativer Erkrankungen zu (Polich, 2007). Somit scheint das Defizit in der Fehlerverarbeitung nicht allein auf eine erhöhte Impulsivität zurückführbar, sondern auch durch eine diskrete kognitive Verlangsamung charakterisiert. Dabei deutet die ebenfalls verminderte Pe-Komponente an, dass das Bewusstsein für begangene Fehler (Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001) in der CSA-Gruppe geringer ausgeprägt ist, als in der Kontrollgruppe. Verminderungen der Pe finden sich beispielsweise auch nach schweren Schädelhirntraumata (Larson & Perlstein, 2009). Zusammenfassend lassen sich somit die Befunde dahingehend interpretieren, dass die Indexgruppe durch eine erhöhte Impulsivität (erhöhte Fehlerrate), ein vermindertes Fehlerbewusstsein (verminderte Pe) und eine diskrete kognitive Verlangsamung (erhöhte P3-Latenz) gekennzeichnet ist.

### 3.4 Implizite Verfahren

Die sogenannten „impliziten“ Testverfahren stützen sich auf Theorien aus der neuropsychologischen und der sozialpsychologischen Kognitionsforschung. Diese Verfahren haben den Vorteil, dass sie auf Verhaltensmechanismen beruhen, die dem Bewusstsein des Probanden begrenzt zugänglich sind, und damit nur schwer absichtlich verfälscht werden können.

Menschliche Verhaltensregulation ist in einem expliziten und einem impliziten System organisiert, deren Haupteigenschaften in Tabelle 8 einander gegenübergestellt sind.

Tab. 8: Die Eigenschaften der beiden Systeme der Verhaltensregulation.

Implizites System	Explizites System
<b>unbewusst</b>	<b>bewusst</b>
<b>schnell</b>	<b>langsam</b>
<b>unflexibel-stereotyp</b>	<b>flexibel</b>
<b>robust</b>	<b>störanfällig</b>
prozedural	deklarativ
parallel-automatisch	seriell-kontrolliert
intuitiv	rational
holistisch	analytisch
heuristisch-assoziativ	logisch
anstrengungslos	anstrengend
konkret	abstrakt

Das explizite System steuert bewusstes, planvolles Verhalten. Es unterliegt einer kontrollierten Informationsverarbeitung mit seriell und langsam ablaufenden Prozessen, hat eine begrenzte Kapazität und benötigt Aufmerksamkeit. Das implizite System hingegen reguliert Fertigkeiten, die beim Lösen komplexer Aufgaben inzidentell angeeignet wurden. Darunter fallen z.B. erlernte motorische Handlungsmuster oder stark routiniertes Verhalten, welches Gewohnheiten und Präferenzen widerspiegelt. Prozesse des impliziten Systems erfordern keine Aufmerksamkeit und laufen schnell und parallel ab, sind dafür aber wenig flexibel. Die sexuelle Präferenz trägt inhärent die Charakteristika des impliziten Systems in Form von stark handlungsbestimmenden Vorlieben, die sich in Gewohnheiten und unbewussten Reaktions- und Handlungsmustern zeigen.

Implizite Test- oder Messmethoden nutzen die geringe Flexibilität und den hohen Grad der Automatisiertheit des impliziten Systems, um auf Präferenzen der Probanden zu schliessen. Ein Umschalten der Verhaltensregulation auf das langsamere explizite System, z.B. im Falle

einer neuen oder ungewohnten Aufgabe, zeigt sich etwa durch längere Reaktionszeiten, was bei sogenannten **Impliziten Assoziationstests** (IATs) genutzt wird. Spontanes Blickverhalten ist besonders durch gewohnheitsmäßige Muster geprägt, die z.B. bei der freien Exploration von Bildmaterial dadurch zu Tage treten, dass attraktive Bilder länger betrachtet werden als weniger attraktive. Auf diesem Effekt beruhen die klassischen **Viewing Time** (VT) Verfahren. Bei den **Choice Reaction Time** (CRT)-Verfahren wird der erhöhte Bedarf an Aufmerksamkeitsleistung beim Umschalten vom impliziten auf das explizite System als Reaktionslatenz bei der Ablösung des Blicks von einem attraktiven Bild gemessen, um die implizite Attraktivität eines Bildes gegenüber anderen, weniger attraktiven Bildern zu erfassen. Des Weiteren dominieren implizite Einstellungen das menschliche Urteilsvermögen, besonders in Situationen, in denen eine schnelle und spontane Entscheidung gefragt ist. So beruht das **Semantic Misattribution-Paradigma** (SMP) darauf, dass sich bestimmte Aspekte eines nur kurz gezeigten Bildes auf eine nachfolgende Beurteilungsaufgabe auswirken und so die implizite Einstellung des Probanden gegenüber dem Bild offenbaren. Auch emotionale Reaktionsmuster werden durch das implizite System gesteuert. Diese können bei der realitätsnahen **Virtuellen Exposition** (VE) mit einem Stimulus, vermittelt durch physiologische Erregung (elektrodermale Aktivität, Herzfrequenz), messbar gemacht werden und so Aufschluss über individuelle Präferenzen gewähren.

### **3.4.1 Semantic Misattribution-Paradigma (SMP)**

#### **Beschreibung der Untersuchung**

Das Semantic Misattribution-Paradigma (SMP) macht sich den Umstand zunutze, dass sehr kurz, unterhalb der Wahrnehmungsschwelle dargebotene Bilder, das Verhalten von Probanden beeinflussen können. Wir verwendeten Bilder aus dem standardisierten Not-Real-People Bilderset (NRP). Dieses ist ein Satz von stückweise zusammengesetzten, fotorealistischen Bildern von Erwachsenen und Kindern beiderlei Geschlechts, welches eigens für die Forschung entwickelt wurde (Pacific Behavioral Assessment Corporation, 2004). Jede dargestellte Person ist von einem Expertengremium nach dem Tannerstadium der sexuellen Entwicklung eingestuft (1 = Kleinkind bis 5 = Erwachsen). Jede der insgesamt 40 Personen ist in einer unbekleideten Variante und einer mit Badekleidung angezogenen Variante vorhanden. In unserem SMP wurden Bilder von nackten Erwachsenen oder nackten Kindern für eine kurze Spanne auf einem Bildschirm dargeboten, gefolgt durch einen Maskierungsreiz (siehe Abb. 11). Bei einem langen zeitlichen Abstand (Stimulus Onset-Asymmetrie, SOA = 77 ms) zwischen dem Bild und dem Maskierungsreiz kann das Bild bewusst wahrgenommen werden (supraliminale Bedingung); bei einem kurzen Abstand (SOA = 28 ms) kann das Bild nicht oder nur eingeschränkt bewusst wahrgenommen werden (subliminale Bedingung).

Unmittelbar nachfolgend wurde ein chinesisches Schriftzeichen eingeblendet, welches für Mitteleuropäer ohne entsprechende Kenntnisse der chinesischen Sprache keine semantische Bedeutung hat. Gleichwohl wurde der Proband aufgefordert, durch Drücken einer Taste anzuzeigen, ob er dem Zeichen eine sexuelle Bedeutung zuschreibt. Neben diesem Verhaltensmass wurden, während des Experiments, ereigniskorrelierte Potentiale (EKPs) aufgezeichnet.

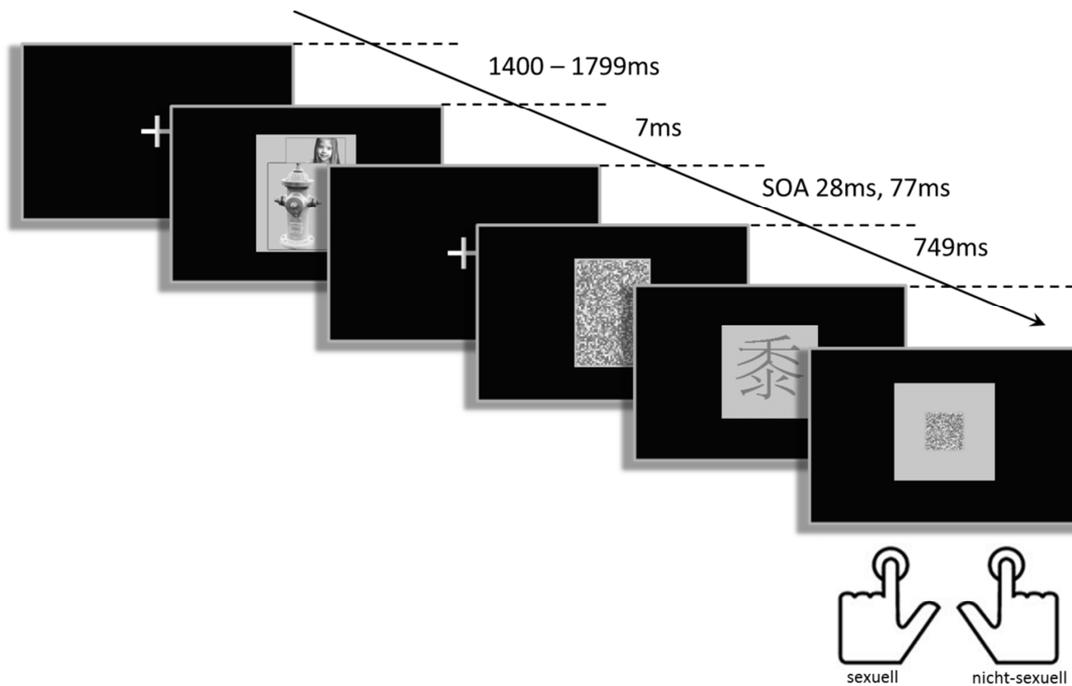


Abb. 11: Die „Semantic Misattribution Task“. Das chinesische Schriftzeichen soll nach seiner sexuellen Bedeutung beurteilt werden (Tastendruck).

Das hier benutzte SMP ist an das von Imhoff et al. (2011) verwendete Verfahren angelehnt, die zeigten, dass heterosexuelle Männer einen deutlichen höheren Proporz an bejahenden „Sexuell“-Antworten nach weiblichen Bildern als nach männlichen Bildern zeigten, während bei homosexuellen Männern ein genau umgekehrtes Bild mit einem grösseren Proporz an bejahenden „Sexuell“-Antworten nach männlichen Bildern als nach weiblichen Bildern zu finden war. Zudem zeigte sich dieser erhöhte Anteil an bejahenden „Sexuell“-Antworten jeweils am deutlichsten für Bilder des Tannerstadiums 5 (Imhoff et al., 2011), was in Einklang mit einer adultsexuellen Präferenz steht.

Unsere Hypothesen für die hier durchgeführte Untersuchung waren, dass das Entscheidungsverhalten durch den Bildreiz in Abhängigkeit von der sexuellen Präferenz bzw. Pädosexualität des Probanden moduliert wird. Konkret erwarteten wir einen höheren Proporz von bejahenden „Sexuell“-Antworten nach Bildern von Kindern in der Gruppe von Pädosexuellen im Vergleich zur Kontrollgruppe, insbesondere für die subliminale Bedingung, da hier eine

Verfälschungstendenz, seitens des Probanden, weitgehend ausgeschlossen werden kann. Für die EKPs erwarteten wir analog verstärkte Antworten auf selbige Reize.

## Ergebnis

Verhaltensdaten: Generell zeigten die Probanden, nach weiblichen Personen, eine signifikant grössere Tendenz das Schriftzeichen als „Sexuell“ zu konnotieren, als nach männlichen Personen ( $F_{3,738} = 9.955, p < .0001$ ). Diese Tendenz war in der CSA-Gruppe am stärksten ausgeprägt ( $t = 4.359, p < .0001$ ) (Abb. 12). Die Antwortpräferenzen wurden, wider Erwarten, weder in der CSO-Gruppe noch in den einzelnen Indexgruppen durch das Tannerstadium des Bildreizes moduliert. Auch die Präsentationsbedingung (subliminal vs. supraliminal) hatte keinen wesentlichen Einfluss auf das Antwortverhalten.

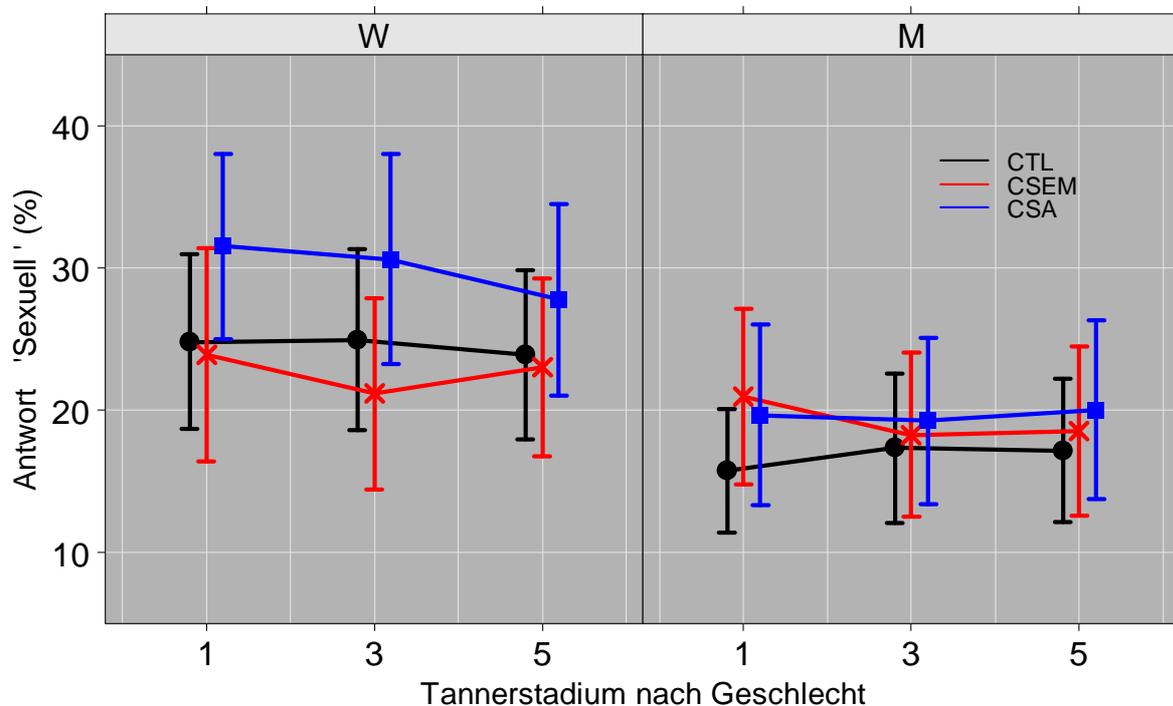


Abb. 12: Die über die Bedingungen (subliminal / supraliminal) hinweg gemittelte Wahrscheinlichkeit auf ein Schriftzeichen eine „Sexuell“-Antwort zu drücken. Die Probanden der CSA-Gruppe (blau) gaben deutlich häufiger nach weiblichen Bildern (links) eine solche Antwort als die CSEM-Gruppe (rot) und die Kontrollgruppe (schwarz). Der Effekt ist unabhängig vom Tannerstadium des Bildreizes. Die Fehlerbalken beschreiben das 95% Konfidenzintervall.

EKP-Daten: Für die Auswertung der EKP-Daten konnten die Daten von vier Probanden (1 CSEM, 3 CSA) aufgrund einer technischen Fehlfunktion bei der Präsentation nicht berücksichtigt werden. Die Abfolge der untersuchten EKPs wird in Abbildung 13 beispielhaft für die Gesamtgruppe in der subliminalen Bedingung gezeigt. Bei diesen EKPs wurden bereits die Anteile, die allein auf die Präsentation der Maske zurückzuführen sind, subtrahiert.

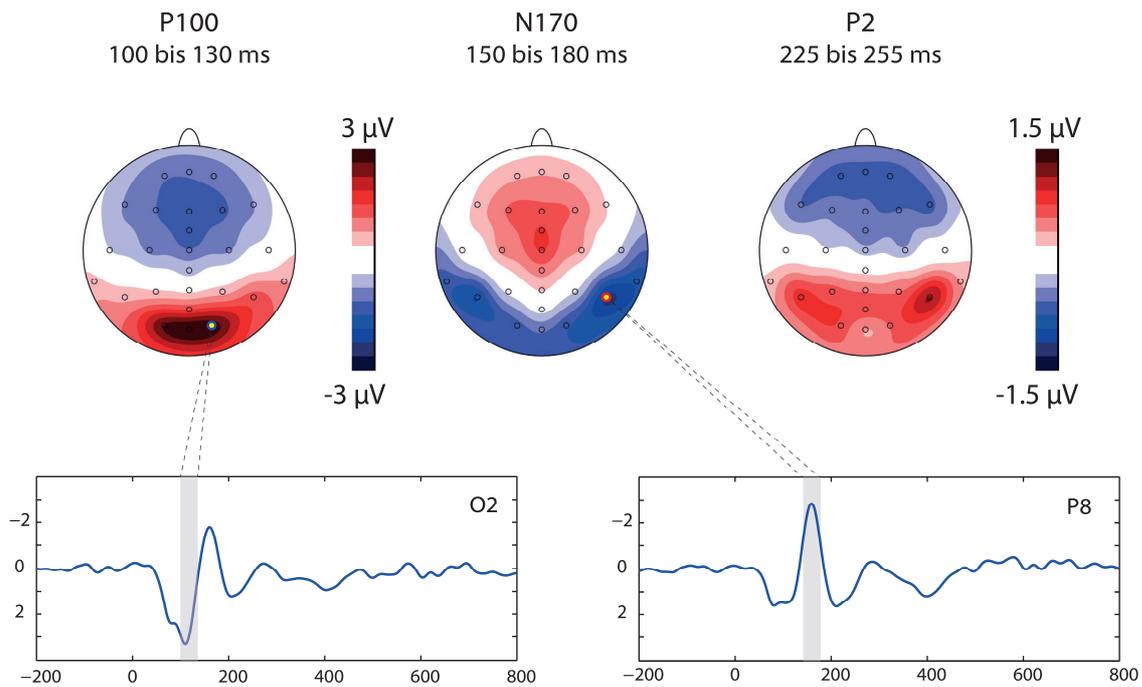


Abb. 13: EKP-Komponenten werden durch die Potentialausrichtung und Latenz spezifiziert. Die für das SMP relevanten EKP-Komponenten werden in ihrer Topographie (Potentialverteilung über die einzelnen Elektroden hinweg) bzw. an ausgewählten Elektroden im zeitlichen Verlauf gezeigt.

Die Indexgruppen zeigten unerwartet eine generell verminderte frühe EKP-Antwort („P100“, ein positives über dem Hinterhaupt gelegenes Potential, Haupteffekt GRUPPE:  $F_{1,57} = 8.596$ ,  $p = .005$ ). Für die subliminale Darbietung wurde dieser Unterschied durch eine Interaktion mit dem Tannerstadium und dem Geschlecht der abgebildeten Person moduliert (Interaktion GRUPPE x TANNER x GESCHLECHT:  $F_{2,114} = 6.306$ ,  $p = .003$ ), aber diese Interaktion spiegelt lediglich wider, dass die Unterschiede für solche Bilder am stärksten ausfielen, für die die Kontrollprobanden die stärkste P100 zeigten. Bei den Kontrollprobanden liesen sich für die P100 signifikante Unterschiede in Abhängigkeit vom Prime finden: die P100 für männliche Kinder (M\_T1) und für erwachsenen Frauen (W\_T5) war am stärksten ausgeprägt (Abb. 14 links). In der supraliminalen Bedingung fand sich ebenfalls eine generell verminderte P100 in den Indexgruppen im Vergleich zur Kontrollgruppe (Abb. 14 rechts), wobei dieser Unterschied statistisch weniger ausgeprägt war als in der subliminalen Bedingung (GROUP:  $F_{1,57} = 3.691$ ,  $p = .060$ ) und hier nicht durch andere Faktoren moduliert wurde.

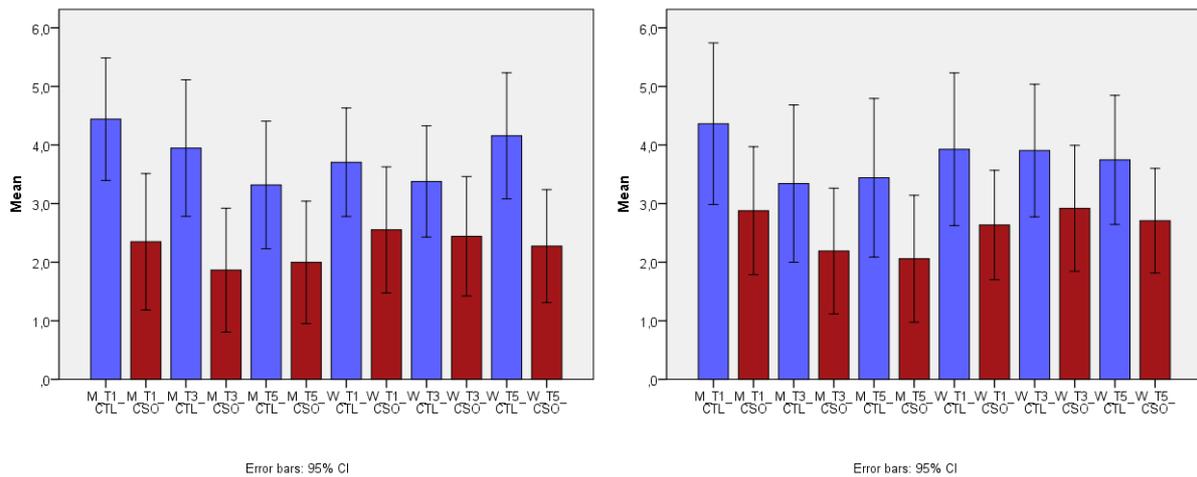


Abb. 14: An den Elektroden O1 und O2 gemittelte Amplituden der P100 in mikroV für die Kontrollgruppe CTL (blau) und kombinierte Indexgruppen CSO (rot), separat für männliche und weibliche Bildreize und Tannerstadien; links die subliminale Darbietung, rechts die supraliminale Darbietung. Deutlich sind die durchgängig verminderten P100-Amplituden in den Indexgruppen. Man beachte zudem, dass die P100 für die Kontrollgruppe bei männlichen Bildern (jeweils die drei blauen Balken links) für das Tannerstadium 1 (T1, jeweils ganz links) am grössten sind.

Innerhalb der Indexgruppen wurden für die subliminale und supraliminale Bedingung keine Gruppenunterschiede (CSEM vs. CSA) für die P100-Amplitude gefunden. Die P100-Amplitude zeigte einen negativen Zusammenhang zum Alter: Innerhalb der Indexgruppen (CSEM & CSA) wie auch innerhalb der Kontrollgruppe (CTL) nahmen die P100-Amplituden mit zunehmendem Alter ab. Die weiter oben beschriebenen Gruppenunterschiede (CSO vs. CTL) blieben jedoch signifikant wenn das Alter als Kovariate berücksichtigt wurde.

Spätere EKP-Komponenten (N170, P2) unterschieden an den hierzu üblicherweise untersuchten Elektroden geringfügig (N170) oder gar nicht (P2) zwischen den Indexgruppen und der Kontrollgruppe. Ebenso liessen sich keine Unterschiede innerhalb der Indexgruppen aufzeigen. Auf eine detaillierte Darstellung der Ergebnisse für diese späteren EKP-Komponenten (N170, P2) wurde daher an dieser Stelle verzichtet.

### Interpretation

Verhaltensdaten: Das beobachtete Antwortmuster ist nicht konform mit unseren Hypothesen. Es fand sich in der CSO-Gruppe keine erhöhte Antwortquote „Sexuell“ nach Bildreizen mit Kindern. Auch fand sich in der Kontrollgruppe unerwartet keine Modulation des Antwortmusters durch das Tannerstadium der abgebildeten Person. Die relativ hohe Rate an „Sexuell“-Antworten nach männlichen Bildern in den heterosexuellen Kontrollpersonen stellt möglicherweise einen Hinweis dar, dass die Primes (also die Bilder nackter Personen) insgesamt nur einen geringen Einfluss auf das Entscheidungsverhalten hatten und die Verhaltensvari-

anz, die auf die Primes zurückgeht, zu gering ist, um sich in Gruppenunterschieden niederzuschlagen. Somit scheint das SMP, in der Form wie es hier umgesetzt wurde, nicht geeignet um Aussagen über die sexuelle Präferenz der Versuchsteilnehmer zu machen.

EKP-Daten: Die generell verminderte P100-Antwort in der Indexgruppe ist ein überraschender Befund. Hierbei ist zu betonen, dass nicht anzunehmen ist, dass die Ausprägung der P100 durch die sexuelle Präferenz moduliert wird. Wie erwähnt, waren in der Kontrollgruppe die P100-Antworten auf erwachsene Frauen und männliche Kinder am stärksten, was kaum mit einer solchen Annahme in Einklang zu bringen ist. Wir führen diese systematische und signifikante P100-Variation in der Kontrollgruppe auf physikalische Charakteristika des Reizmaterials zurück (Grösse der abgebildeten Personen, Kontraste, Farben). Die Modulation des Gruppenunterschieds durch das Tannerstadium und Geschlecht der abgebildeten Person kennzeichnet somit *nicht* pädosexuelle Neigung, sondern deutet an, dass die Unterschiede zwischen der Index- und Kontrollgruppe für solche Reize am grössten sind, die starke P100-Antworten generieren. Die generell verminderte P100 in den Indexgruppen könnte auf ein neurokognitives Defizit in der Verarbeitung visueller Reize hinweisen, was insbesondere dann zu Tage tritt, wenn solche Reize nur kurzzeitig dargeboten werden und könnte auf eine Störung in der Entwicklung des neuronalen Nervensystems in der Indexgruppe hindeuten.

### **3.4.2 Impliziter Assoziationstest (IAT)**

IATs finden in der sozialen Einstellungsforschung schon weithin Anwendung. Ganz allgemein prüft ein IAT Einstellungen/Neigungen gegenüber einem vordefinierten Konzept. Zur Evaluation der sexuellen Präferenz, wie in unserem Fall, wird das Konzept (Erwachsen vs. Kind) durch verschiedene Bilder von Erwachsenen und Kindern und die Einstellungen durch sexuell vs. nicht-sexuell konnotierte Attributwörter repräsentiert. Die Bilder und Attributwörter werden abwechselnd auf einem Bildschirm angezeigt. Die Versuchspersonen werden gebeten die Bilder und Eigenschaftswörter möglichst schnell und fehlerfrei durch Tastendruck mit der linken oder rechten Hand den Konzepten zuzuordnen, die jeweils der linken oder rechten Seite zugeordnet sind, wobei die Reaktionszeiten gemessen werden, vgl. Abbildung 15.

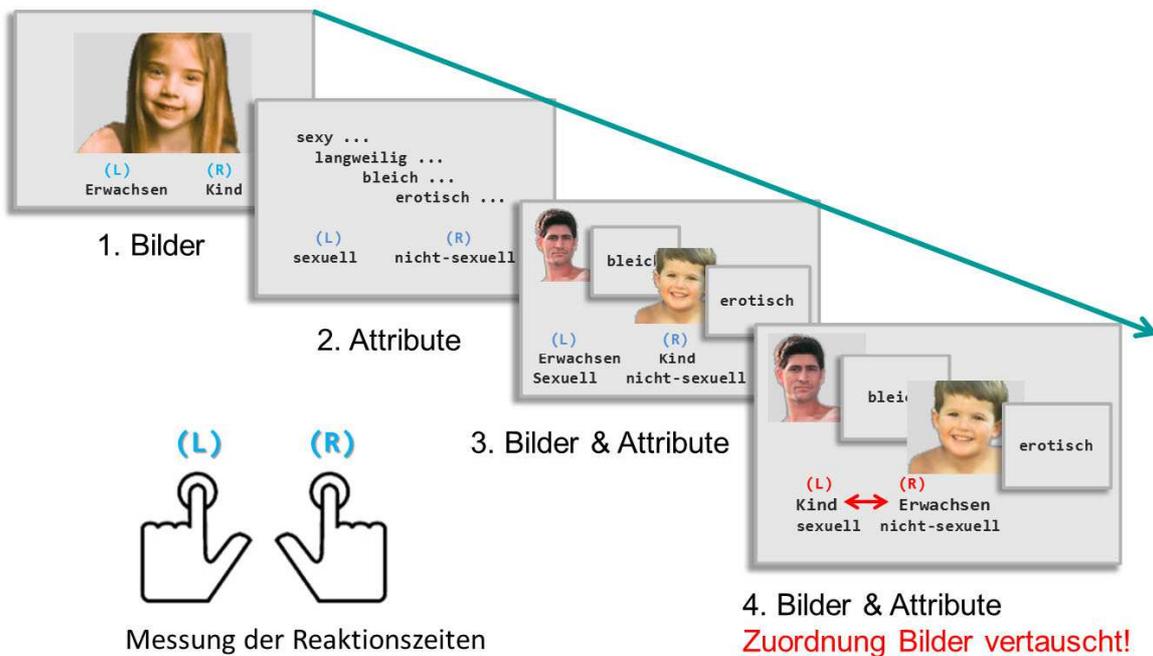


Abbildung 15: Schematischer Ablauf des Impliziten Assoziationstests.

In Abhängigkeit von der Neigung des Probanden ergeben sich dabei zwei Untersuchungsbedingungen: Es ist für ihn jeweils eine der Paarungen kongruent, die andere nicht kongruent, wie in der Abbildung 15 unter Bilder & Attribute dargestellt. Bei Annahme einer kurzen Reaktionszeit bei kongruenter Bild-Attribut Paarung, da diese vom schnellen impliziten System verarbeitet werden kann, ist es somit möglich, aus der Differenz der Reaktionszeiten zwischen den beiden Bedingungen auf die sexuelle Präferenz zu schließen. Der IAT, den wir hier eingesetzt haben, liefert negative Werte für die IAT-Effektgröße bei einer pädosexuellen und positive Werte bei einer adultsexuellen Präferenz. Der Betrag der IAT-Effektgröße gibt dabei die Ausprägung des Zusammenhangs an, d.h. je größer der absolute Wert ist, umso stärker ist die Präferenz ausgeprägt. Es muss hier allerdings beachtet werden, dass das Lösen der IAT-Aufgabe vom individuellen Mass an Reaktionsflexibilität beeinflusst werden kann. Die IAT-Effektgröße fällt bei flexiblen Individuen geringer aus, da ihnen die Ablösung von automatischen Reaktionstendenzen natürlicherweise leichter fällt. Bei der vorliegenden Auswertung wurde daher ein Mass für Flexibilität berücksichtigt, die „Reversal Learning Capacity“ (RLC), welches in der neuropsychologischen Testung erfasst wurde.

## Ergebnis

Aufgrund unzureichender Performanz (weniger als 400 gültige Durchgänge bei insgesamt 540 Durchgängen) mussten sechs Probanden von der Auswertung ausgeschlossen werden, die Gruppengrößen betragen demnach CTL = 20, CSEM = 14, CSA = 20.

Die CTL-Gruppe wies, entsprechend unserer Erwartung, eine mittlere positive IAT-Effektgrösse von 0.37 auf, siehe Abbildung 16, was im Einklang mit einer adultsexuellen Präferenz steht. Die Indexgruppen zeigten hingegen kein so eindeutiges Resultat. Mit mittleren IAT-Effektgrösse von -0.05 (CSA) und -0.03 (CSEM) liegen sie zwar, wie für diese Gruppen erwartet, im negativen Bereich, sind aber deutlich geringer im negativen Bereich, als auf der Grundlage der CTL-Gruppe prädiziert. Die für die CSA-Gruppe erhobenen SSPI-Werte korrelierten nur zu  $\rho = .21$  mit den IAT-Effektgrössen, was der Erwartung eines umgekehrt proportionalen Zusammenhangs - je grösser der Pädophilie-Index desto kleiner die IAT-Effektgrösse - widerspricht. Damit ist die Erklärung der geringen Differenzierung des IAT zwischen den Indexgruppen durch eine uneindeutige sexuelle Präferenz eher unwahrscheinlich.

Trotzdem sehen wir ein klares Potenzial, auf der Grundlage der IAT-Effektgrösse, die Gruppenzugehörigkeit (CTL vs. CSO) zu prädizieren, da sich die beiden Gruppen statistisch signifikant unterscheiden lassen ( $F_{3,50} = 5.888$ ,  $p = .002$ ). Innerhalb der Indexgruppen war aber keine Differenzierung möglich. Der signifikante Einfluss der RLC ( $p = .02$ ) auf die IAT-Effektgrösse wurde bei der Teststatistik berücksichtigt.

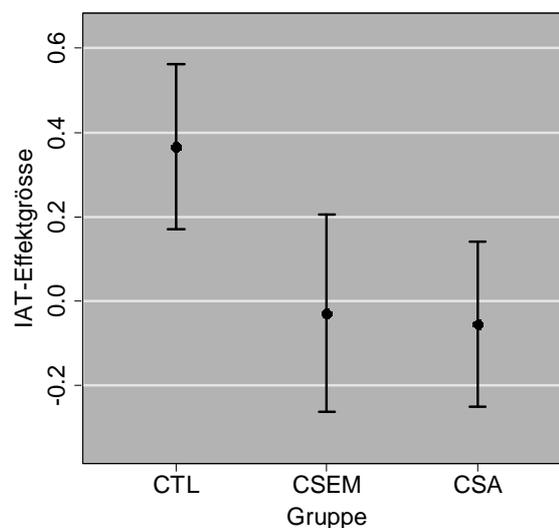


Abbildung 16: Die IAT-Effektgrössen in den drei Gruppen CTL, CSEM und CSA. Die Fehlerbalken beschreiben das 95% Konfidenzintervall.

Eine „Receiver-Operating-Characteristic“-Analyse (ROC) der IAT-Effektgrössen, zunächst unbereinigt um den Einfluss der RLC, ergab eine passable Diskriminabilität der Gruppen CTL und CSO (AUC = 76.8). Die RLC war in allen Gruppen ähnlich verteilt und wirkte sich daher nicht stark auf die Gruppenunterschiede aus. Das für RLC adjustierte Modell ergab dabei eine geringfügig geringere Diskriminabilität (AUC = 73.8).

### 3.4.3 Betrachtungszeit (Viewing Time, VT) und Bildbewertung

In einer freien Explorationssituation werden Bilder mit sexuell attraktiven Personen länger betrachtet als Bilder mit sexuell weniger attraktiven Personen, woraus auf die sexuelle Präferenz der Probanden geschlossen werden kann. Die Studienteilnehmer bekamen bei dem vorliegenden VT-Verfahren nacheinander alle Bilder auf dem Bildschirm angezeigt, die in den verschiedenen Verfahren der MIPS-Studie verwendet wurden und sollten deren Attraktivität auf einer Skala von 1 = „sehr unattraktiv“ bis 6 = „sehr attraktiv“ einstufen. Die Zeit, die die Probanden für die Bewertung jedes Bildes benötigten, wurde registriert. Das Bildmaterial entstammte dem NRP-Bilderset (Tannerstadien 1 bis 5). Unserer Hypothese zufolge sollte sich die unterschiedliche sexuelle Präferenz der drei Gruppen so abbilden, dass die CTL-Gruppe die höheren Tannerstadien länger betrachtet als die niedrigeren, und umgekehrt die Indexgruppen die niedrigeren Tannerstadien länger als die höheren. Das Geschlecht und die Bekleidung (bekleidet vs. unbekleidet) wurden dabei nicht berücksichtigt.

#### Ergebnis

Ein Proband der CSEM-Gruppe ging nicht in die Analyse ein, damit waren die Gruppengrößen wie folgt: CTL = 21, CSEM = 20, CSA = 22. Die verschiedenen Tannerstadien wurden unterschiedlich lang betrachtet und dieser Effekt unterschied sich auch zwischen den Gruppen (TANNER x GRUPPE Interaktion:  $F_{10,238} = 3.257$ ,  $p < .001$ ), allerdings ist die Natur dieser Interaktion komplexer als erwartet: Gruppenübergreifend wurden die niedrigen Tannerstadien kürzer betrachtet als die höheren, der Grad der Unterschiedlichkeit war in der CTL-Gruppe am grössten, darauf folgte die CSA-Gruppe und dann die CSEM-Gruppe. Die Analyse von neuropsychologischen Kovariaten mit Einfluss auf die Betrachtungszeit stellte den verbalen Intelligenzquotienten (vIQ) als bedeutsam heraus. Es war ein differentieller Einfluss von Gruppenzugehörigkeit mit der verbalen Intelligenz zu beobachten, der besonders in der CTL-Gruppe zutage trat (vIQ x GRUPPE Interaktion:  $F_{2,59} = 6.252$ ,  $p = .003$ ). Weniger intelligente Probanden der CTL-Gruppe hatten sehr kurze Betrachtungszeiten und eine geringere Dynamik zwischen den Tannerstadien. In den Gruppen CSEM und CSA ist dieser Einfluss auch vorhanden, kommt dort aber weniger zum Tragen. In Abbildung 17 sind die beschriebenen Zusammenhänge verdeutlicht. Dort sind die vorhergesagten Betrachtungszeiten auf die Tannerstadien (I bis V) in den drei Gruppen abgebildet, und zwar für jede Gruppe unter drei Abstufungen des verbalen IQs (IQ 85, IQ 100 und IQ 115). Besonders der Einfluss des vIQs in der CTL-Gruppe sticht heraus. Eine mögliche Erklärung hierfür wäre der Einsatz von Vermeidungsverhalten bei den weniger intelligenten Probanden der CTL-Gruppe.

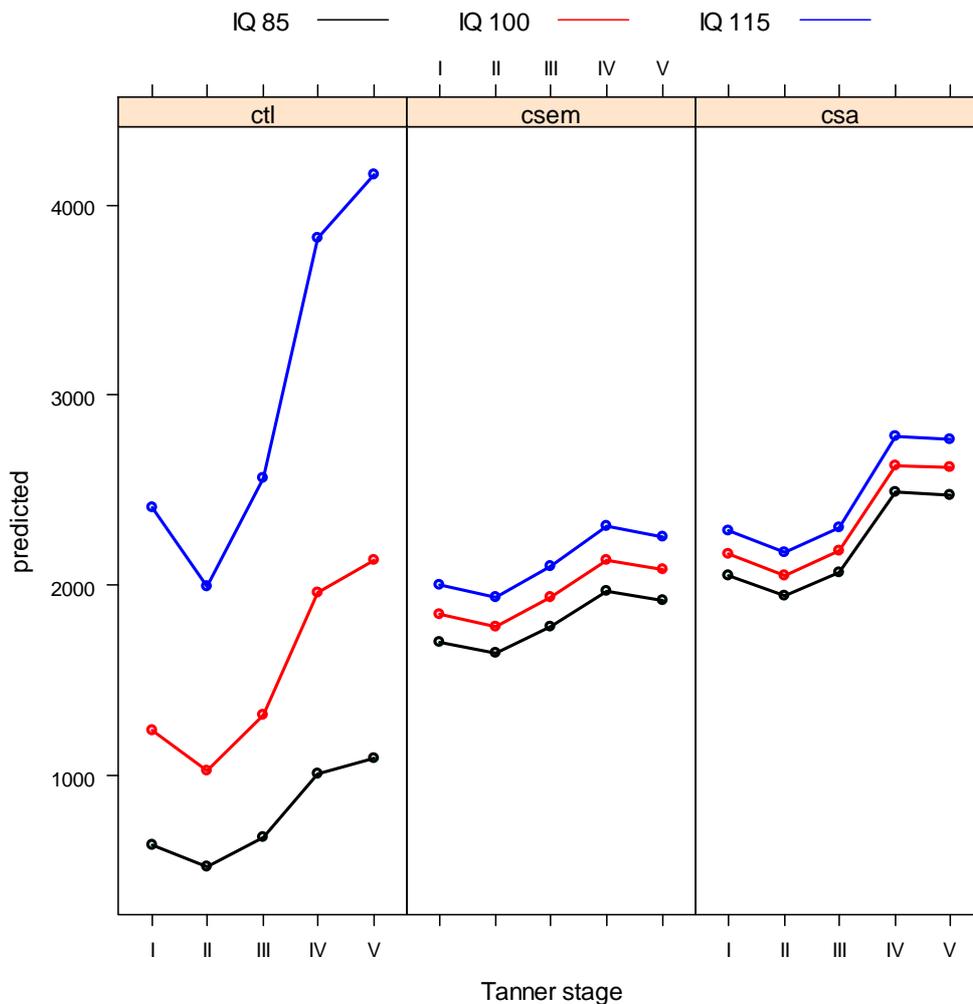


Abbildung 17: Prädiktionsmodell des VT-Verfahrens. Die vorhergesagte mittlere Betrachtungszeit (ms) der Tannerstadien (I bis V) in den drei Gruppen CTL, CSEM und CSA unter der Annahme von drei Abstufungen des verbalen IQs (IQ 85 = schwarz, IQ 100 = rot und IQ 115 = blau).

Einzelfallvorhersage: Trotz des komplexen Modells und der unerwarteten Zusammenhänge beim VT-Verfahren, soll die Nützlichkeit zur Einzelfallbewertung demonstriert werden. In Abbildung 18 sind vier Beispiele dargestellt. Die individuellen VT-Profile der vier Probanden über die fünf Tannerstadien sind als schwarze Punkte dargestellt. Die rote Fläche umschreibt den Bereich, der typisch ist für das adultsexuelle Gruppenprofil (Gruppe CTL), die grüne Fläche für das pädosexuelle Gruppenprofil (CSO). Die Flächen werden jeweils für das vIQ-Niveau des Probanden berechnet. Für jedes Individualprofil wird nun ein Mass für die Distanz zu den beiden Gruppenprofilen berechnet, was im Diagramm als ‚adult prox‘ und ‚child prox‘ angegeben ist. Die Nähe zum jeweiligen Gruppenprofil ist ein wichtiges Klassifikationskriterium: Die kleinere Zahl verrät in der Regel, welcher Gruppe das Profil zuzuschlagen ist. Allerdings ist dieses Kriterium der Nähe nicht immer ausreichend, siehe Grenzfall in Abbildung 4, rechts oben. Aus diesem Grund wurde es durch das Verhältnis von ‚adult prox‘ zu ‚child prox.‘ ersetzt. Ein Grenzwert für diese Ratio wurde mit Hilfe einer ROC-Analyse be-

stimmt, er lag bei adult/child prox. = 1.142 (AUC = 76.5), d.h. Werte kleiner als 1.142 sprechen für eine adultsexuelle Präferenz.

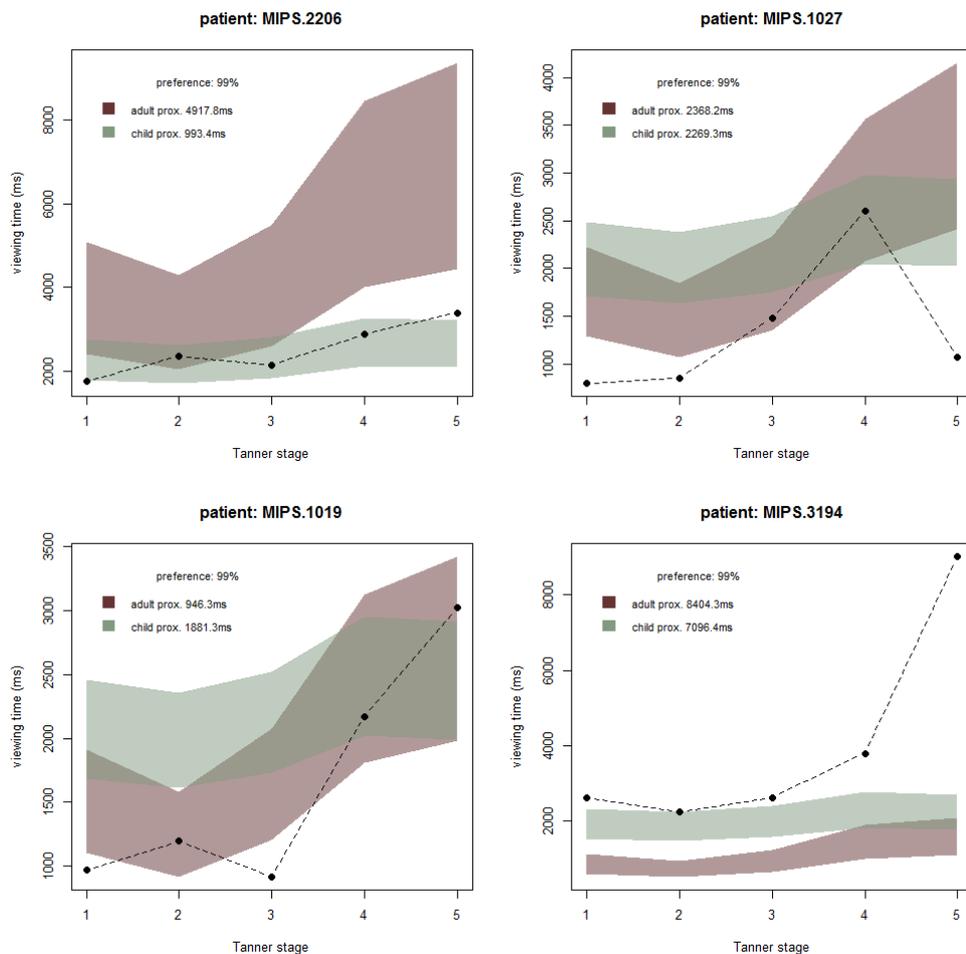


Abbildung 18: Vier korrekte Einzelfallklassifizierungen, basierend auf dem individuellen VT-Profil.

Punkte mit gestrichelter Linie: Individuelles VT-Profil über die fünf Tannerstadien.

Flächen: Nach vIQ des Probanden bestimmte Gruppenprofile. Rot: Adultsexuell; grün: Pädosexuell.

Wahrscheinliche Kinderpräferenz: Child prox. kleiner als adult prox.

Wahrscheinliche Erwachsenenpräferenz: Adult prox. kleiner als child prox.

Wichtigstes Kriterium: Ratio adult/child prox. < 1.142 höchstwahrscheinliche Erwachsenenpräferenz.

Links oben: Besonders gute und überzeugende Klassifizierung:

Child prox. kleiner & Ratio ~ 4.9 → Kinderpräferenz.

Recht oben: Grenzfall einer Klassifikation: Child prox. kleiner, aber Ratio 1.044 < 1.142 → Erwachsenenpräferenz.

Links unten: Korrekte Klassifikation: Adult prox. kleiner und Ratio ~ 0.5 → Erwachsenenpräferenz

Rechts unten: Viewing Time ist maximal bei Tanner 5 (spricht für Erwachsenenpräferenz) und dennoch korrekte wengleich knappe Klassifikation: Child prox. kleiner und Ratio 1.184 > 1.142 → Kinderpräferenz.

## Bildbewertung

Die Auswertung der Bewertung der Bildattraktivität ergab signifikante Interaktionen der drei Gruppen mit den wichtigsten beteiligten Faktoren: Geschlecht und Tannerstadium. Auch hier zeigte sich ein signifikanter Einfluss von verbalem IQ auf das Bewertungsverhalten ( $\rho = -.22$ ,  $p < .001$ ); intelligenter Probanden vergaben weniger hohe Bewertungen.

Der Befund ist wie schon bei der VT sehr komplex. Generell vergab die CTL-Gruppe fast überall die höchsten Bewertungen und hob sich besonders bei den Tannerstadien 1 bis 3 von den Indexgruppen ab (Interaktionseffekt Gruppe CTL vs. CSO x Tannergruppe 1-3 vs. 4-5:  $F_{1,1127} = 4.603, p = .03$ ), vgl. Abbildung 19. Weibliche Bilder der Tannerstadien 4 und 5 wurden gruppenübergreifend niedriger bewertet als die Stadien 1 bis 3 (Interaktionseffekt Geschlecht x Tannergruppe 1-3 vs. 4-5:  $F_{1,1127} = 99.2824, p < .001$ ). Bei den männlichen Bildern fielen die Unterschiede zwischen den Tannerstadien gruppenübergreifend wesentlich geringer aus.

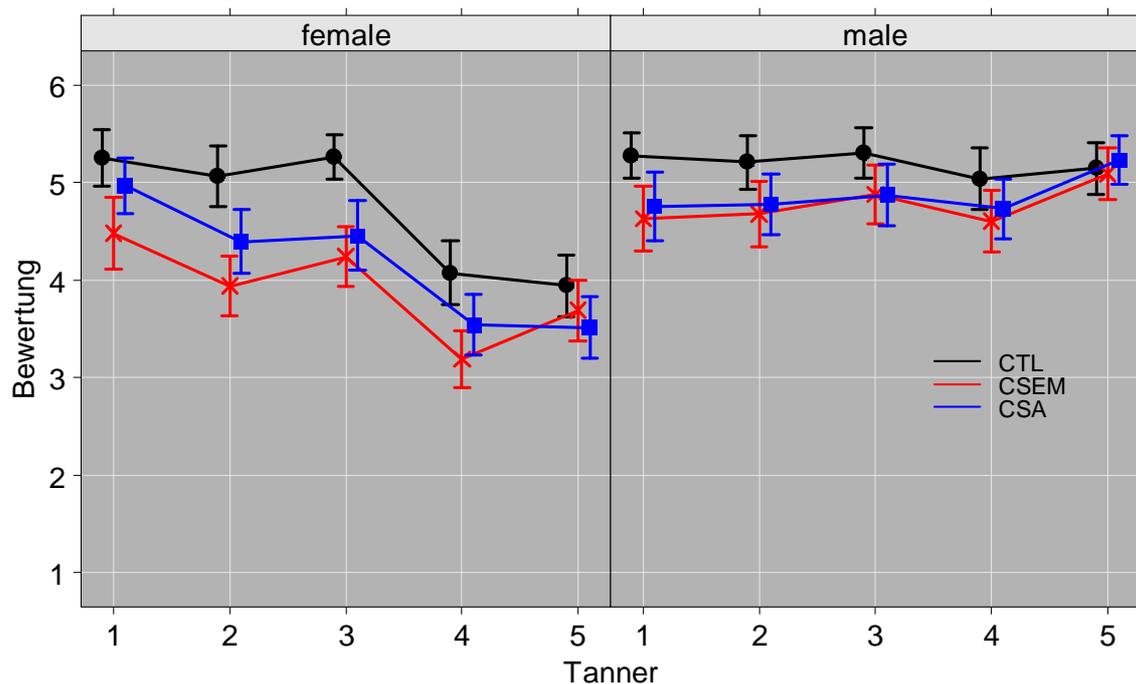


Abbildung 19: Rating-Mittelwerte (Likert-Skala von 1 = sehr unattraktiv bis 6 = sehr attraktiv) des NRP-Bildersatzes in Abhängigkeit vom Tannerstadium, getrennt nach Geschlecht (W, M) in den drei Gruppen CTL, CSEM und CSA. Die Fehlerbalken beschreiben das 95% Konfidenzintervall.

### 3.4.4 Choice Reaction Time-Aufgabe (CRT)

Die CRT-Aufgabe fügt dem VT-Paradigma eine Komponente der zeitlichen Taktung sowie eine kognitive Aufgabe hinzu. Sie erreicht damit einen höheren Grad an Standardisierung und wird zudem robuster gegen Vermeidungsstrategien. Es wurden auch hier Bilder aus dem NRP-Bildersatz verwendet, faktorisiert nach drei Tannerstadien (niedrig, mittel und hoch), bekleidet und unbekleidet und beiderlei Geschlechts. Die Bilder wurden nacheinander am Bildschirm präsentiert, nach kurzer Zeit erschien zufällig an einer von neun möglichen Positionen auf dem Bild eine Marke, woraufhin die Probanden eine Taste drücken sollten, welche mit der Position der Marke korrespondierte. Der Schluss auf die Präferenz für ein Bild

ist demnach abhängig von der Zeit, welche benötigt wird, um die Aufmerksamkeit von dem Bildreiz abzuwenden und die richtige Taste zu drücken.

Erwartet wurde auch hier ein zwischen den Gruppen unterschiedliches Profil der Reaktionszeiten, d.h. kürzere bei niedrigen Tannerstadien und längere bei höheren Tannerstadien in der Gruppe CTL, sowie ein umgekehrtes Muster in der CSO-Gruppe.

## **Ergebnis**

Ein Proband der CSEM-Gruppe ging nicht in die Auswertung ein. Es ergab sich ein Gruppeneffekt, die Reaktionszeiten waren in der CSO - Gruppe höher als in der CTL-Gruppe ( $F_{1,60} = 11.514$ ,  $p = .001$ ), allerdings über alle Tannerstufen hinweg und auch unabhängig vom Bekleidungsstatus. Diese generelle Verlangsamung der Indexgruppen war auch bei anderen Tests zu beobachten, z.B. Go/NoGo. Die CRT-Aufgabe ergab keine differenziellen Effekte, die auf die sexuelle Präferenz in den jeweiligen Gruppen schließen lassen. Als mögliche Ursache für dieses sehr unerwartete Ergebnis kommt die Qualität des Stimulusmaterials, d.h. des NRP-Bildersets in Frage, besonders Aspekte der Heterogenität der Darstellung und insgesamt die Attraktivität der dargestellten Personen. Zudem wurde die CRT-Aufgabe im Anschluss an die Aufdeckung der Coverstory durchgeführt. Es ist daher möglich, dass die individuelle Reaktion der Probanden darauf eine unkontrollierbare Varianzquelle für darauffolgende Tests darstellte, welche die Testeffekte überlagerte.

### **3.4.5 Virtuelle Exposition**

Um die Habituation von Sexual-Content-Related Arousal zu messen, wurde in der Virtuellen Realität eigens eine Simulation entwickelt, die einen Spaziergang am Strand einer Ferieninsel zum Gegenstand hat (siehe Abb. 20). Dabei wird der Proband in einem Abstand von 3.9m vor einer raumfüllenden Leinwand positioniert und erlebt den virtuellen Spaziergang in Echtzeit und 3D. Im Zuge dieses Spazierganges kommt der Untersuchungsteilnehmer an einzelnen Personen oder Gruppen von Personen vorbei (Szenerien), die zuweilen bekleidet aber auch unbekleidet sind. Der Proband beobachtet die Szenerien passiv. Er wird instruiert, die in der Szenerie enthaltenen Personen nach einem bestimmten Merkmal auf dem Körper (Hautpflaster) zu prüfen und, wenn dieses entdeckt wird, eine Taste zu drücken. Somit ist sichergestellt, dass die in der Szenerie enthaltenen Personen visuell „gescannt“ werden. Die Szenerien werden insgesamt viermal in gleicher Reihenfolge wiederholt. Ein Durchgang dauert etwa 12.5 Minuten. Währenddessen werden physiologische Parameter (Augenbewegungen, Fixationsverhalten, elektrodermale Aktivität, Elektroenzephalogramm) des Probanden abgeleitet, sodass Erregung, Müdigkeit und das Blickverhalten der Probanden objektiv erfasst werden.

## Hypothesen

Das Paradigma erlaubt somit die Aufzeichnung des physiologischen Erregungsniveaus in Abhängigkeit von der Szenerie. Arousal-Muster können damit der Szenerie zugeschrieben und über die wiederholten Darbietungen hinweg untersucht werden. Dabei erwarten wir in Abhängigkeit der jeweiligen sexuellen Präferenz der Probanden ein zwischen den Versuchsgruppen unterschiedliches Erregungsmuster mit im Vergleich zur Kontrollgruppe stärkerem Arousal in der CSA-Gruppe für Szenerien, die Kinder beinhalteten. Zudem wurde angenommen, dass die wiederholte Exposition im Sinne einer Habituation (Antwortabnahme über die Zeit) zwischen den Gruppen unterschiedlich ausfällt.

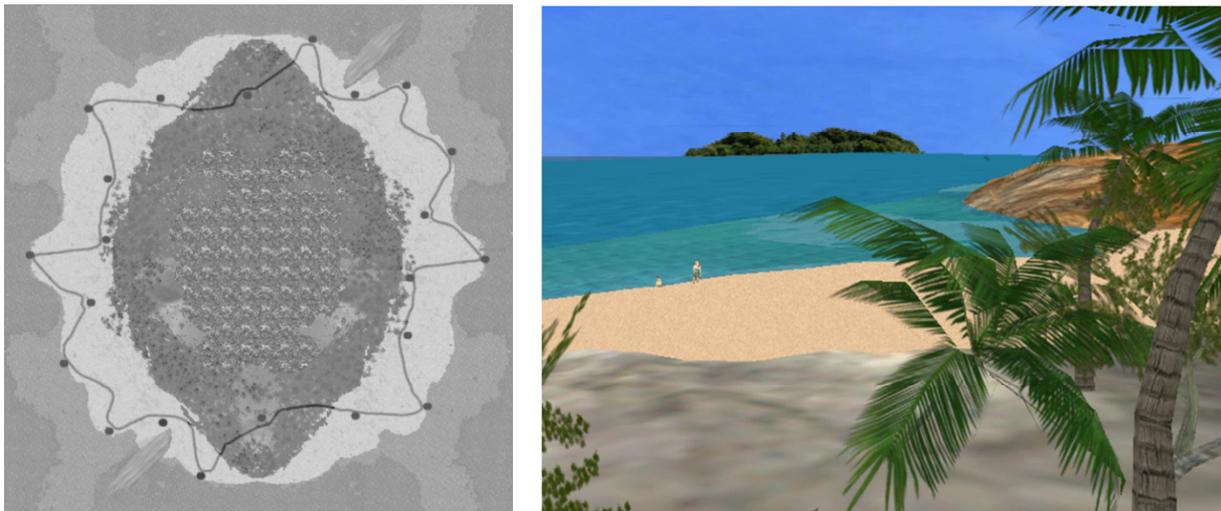


Abb. 20: Links: Schematischer Aufbau der virtuellen Insel. Die durchgezogene Linie repräsentiert den Spaziergang des Probanden entlang des Strandes. Überall dort, wo sich schwarze Punkte befinden, findet der Proband entweder Personen, alleine oder in Gruppen, vor (siehe Abb. 21). Diesen begegnet der Proband in vier Durchgängen wiederholt. Die Insel ist vollkommen punktsymmetrisch. Ein Durchgang entspricht damit einer halben Umrundung. Rechts: Eine Szenerie mit einem Jungen des Tannerstadiums 1 und einem erwachsenen Mann (Tannerstadium 5) aus der Ferne.



Abb. 21: Die elf Szenarien der virtuellen Insel, bestehend aus Personen, die dem Not-Real-People Bilderset entnommen wurden.

Die Resultate werden wie folgt dargestellt: Zunächst folgt die Analyse der Verhaltensdaten (Häufigkeiten, mit denen die Hautpflaster als Zielreize entdeckt wurden), um sicherzustellen, dass den Szenarien über die jeweiligen untersuchten Gruppen jeweils in vergleichbarem Masse Aufmerksamkeit entgegengebracht wurde. Der nächste Abschnitt wendet sich dann der elektrodermalen Aktivität (EDA) zu. Dabei wird deren Änderung über die Zeit und in Abhängigkeit der Szenarien betrachtet. Dann soll geprüft werden, inwiefern die Hypothesen über ein gruppenspezifisches Habituationmuster gültig sind. In einem weiteren Schritt wird der Einfluss der Müdigkeit und ihrer Änderung über die Zeit auf die elektrodermale Aktivität geprüft. Des Weiteren wird geprüft, inwiefern sich ein gruppencharakteristisches Blick- und Fixationsverhalten identifizieren lässt. Die Klassifikationsanalyse erstreckt sich auf den wichtigsten Ergebnisparameter der Virtuellen Exposition, nämlich auf EDA-Muster, die eine optimale Zuordnung der Gruppen ermöglichen. Abschliessend stellen wir eine alternative, rein datengestützte Auswertung vor, die zeigen soll, dass es eine Fülle von Möglichkeiten gibt Informationen zu erschliessen, die den Daten der Virtuellen Exposition zugrunde liegen, mit jeweils eigenen Implikationen für die klinische Bedeutsamkeit des Resultats und dem primären Endpunkt: der Vorhersage der sexuellen Präferenz und des Übergriffrisikos.

### 3.4.5.1 Verhaltensdaten

Die Aufgabe der Teilnehmer bestand darin, Pflaster auf den Körpern der Personen zu entdecken, und dies mit Druck auf eine Maustaste anzuzeigen. Pro Durchgang waren vier bis sechs Pflaster zu finden, welche zufällig über die Stimuli verteilt angebracht waren. Abbildung 22 stellt die Häufigkeiten dar, mit denen es den Probanden der drei Gruppen gelungen ist, die Pflaster zu entdecken.

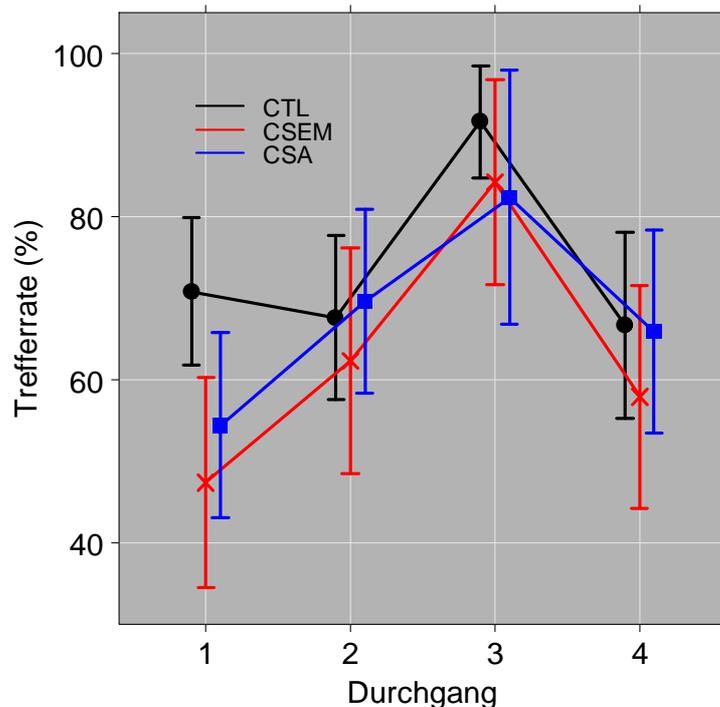


Abbildung 22: Trefferrate bei der Erkennung des Pflasters in den drei Gruppen CTL, CSEM und CSA über die vier Durchgänge hinweg. Die Fehlerbalken beschreiben das 95% Konfidenzintervall.

Pro Gruppe gibt es jeweils einen fehlenden Datensatz, bedingt durch technische Ausfälle, so dass insgesamt nur  $N = 21$  CTL-, 20 CSEM- und 20 CSA-Probanden in die Analysen eingehen können. Man erkennt einen weitgehend parallelen Verlauf der Trefferquoten bei den drei Gruppen. Die Variabilität zwischen den Durchgängen ist durch Übungseffekte und die unterschiedliche Schwierigkeit erklärbar die Pflaster zu entdecken. Daneben allerdings fällt auf, dass die Kontrollgruppe initial eine deutlich höhere Trefferquote (71%) gegenüber den beiden anderen Gruppen aufweist (CSEM: 47%,  $t = -2.9$ ,  $df = 51$ ,  $p = .005$ ; CSA: 54%,  $t = 2.0$ ,  $df = 51$ ,  $p = .052$ ), welche zumindest für die CSEM-Gruppe statistisch signifikant ist. Dies könnte darauf hindeuten, dass die Probanden der CSO-Gruppe zu Beginn der Untersuchung Schwierigkeiten mit der eher unüblichen Aufgabe hatten. Möglicherweise waren diese initial aber auch durch die Bildinhalte stärker abgelenkt als die Kontrollgruppe. Mit Ausnahme des ersten Durchgangs zeigt sich jedoch, dass die Aufmerksamkeit, die über die Durchgänge hinweg den Szenarien entgegengebracht wurde, über die untersuchten Gruppen vergleichbar war.

### 3.4.5.2 Elektrodermale Aktivität

Die Hautleitaktivität (EDA) wurde durch eine Dekompositionsmethode in drei Bestandteile zerlegt: (1) in einen periodischen - über die vier Durchgängen hinweg bestehenden Anteil, (2) einen globalen Trend, der die Veränderung über die vier Durchgänge hinweg beschreibt, und (3) in darüber hinaus verbleibende Restwerte (Residuen). Aufgrund fehlender EDA-Daten stehen für die Analyse jedoch nur  $N = 20$  CTL-,  $N = 19$  CSEM- und  $N = 18$  CSA-Probanden zur Verfügung. Die Abbildungen (vgl. Abb. 23) zeigen ausschliesslich den mittleren periodischen Anteil über 150 Zeitintervalle (zu je 5 Sek.) eines Durchganges. Die grau schattierten Bereiche kennzeichnen die elf definierten Personenszenarien. Deutlich sichtbar sind Aktivitätsspitzen, die jeweils gegen Ende einer Szenerie auftreten und bisweilen sogar ganz in die nächste Szenerie hineinreichen, da die volle Sichtbarkeit der in der Szenerie enthaltenen Personen erst gegen deren Ende erreicht wird. Mittels des Klassifikationsalgorithmus konnten nun genau drei Zeitpunkte identifiziert werden, deren korrespondierende EDA-Werte eine möglichst gute Trennung von der CTL- und der CSO-Gruppe mit einer Genauigkeit von 75.5% korrekter Zuordnungen ermöglichen, in Abb. 23 durch die gestrichelten Linien gekennzeichnet. Die Punkte liegen in der Szenerie 11 (unbekleidetes Mädchen Tannerstadium 1) und Szenerie 10. Dort jedoch im Rahmen einer Aktivitätsspitze, die in Szenerie 9 (zwei unbekleidete Frauen der Tannerstadien 4 und 5) ausgelöst wurde. Entgegen der Hypothese, wonach CSO-Probanden stärker auf Kinder und CTL-Probanden auf Erwachsene ansprechen sollten, zeigen die tatsächlichen Hautleitaktivitäten einen genau entgegengesetzten Effekt, der mittels der herkömmlichen Statistiken jedoch nicht abzusichern ist. Ein statistischer Trend lässt sich für die Aktivitätsspitze nachweisen die zwischen Szenerie 10 und 11 liegt ( $W = 469, p < .1$ ).

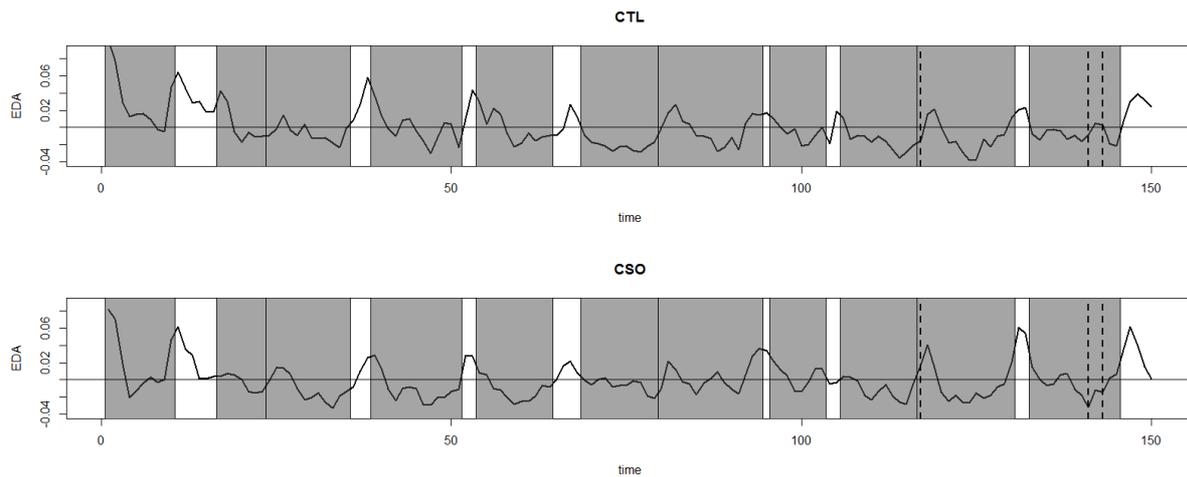


Abb. 23: Periodischer Anteil der Hautleitaktivität als Funktion der Zeit (Zeitintervalle) der Gruppen CTL und CSO. Grau schattierte Bereiche kennzeichnen die elf definierten Personenszenarien. Am Ende einer Szenerie kommt es häufig zu einer EDA-Aktivitätsspitze. Die durch die gestrichelten Linien gekennzeichneten Intervalle wurden durch den Klassifikationsalgorithmus ausgewählt und repräsentieren EDA-Aktivitätsmuster mit besonders hoher Diskriminabilität.

Der Vergleich zwischen der CSEM- und CSA-Gruppe führt zu ähnlichen Mustern. Die Klassifikationsanalyse indiziert nun jedoch für vier Zeitintervalle eine besondere Diskriminabilität. Erneut sind die neunte und elfte Szenerie betroffen. Nun gewinnen aber auch die dritte und vierte Szenerie an Bedeutung. Dabei handelt es sich um Personengruppen, die jeweils vier Stimuli beinhalten (vgl. Abb. 23), die ihrerseits die Tannerstadien eins, drei und fünf, beide Geschlechter und ausschliesslich bekleidete Figuren (Badebekleidung) umfassen. Die Klassifikationsanalyse lässt auf Grundlage dieser Muster eine Genauigkeit von ca. 80% korrekter Zuordnungen schliessen. Im Einzelvergleich erweist sich eine Aktivitätsspitze zwischen den Gruppen CSEM und CSA, die der Szenerie 8 zugehört, als signifikant unterschiedlich (CSEM > CSA  $W = 103$ ,  $p = 0.039$ ).

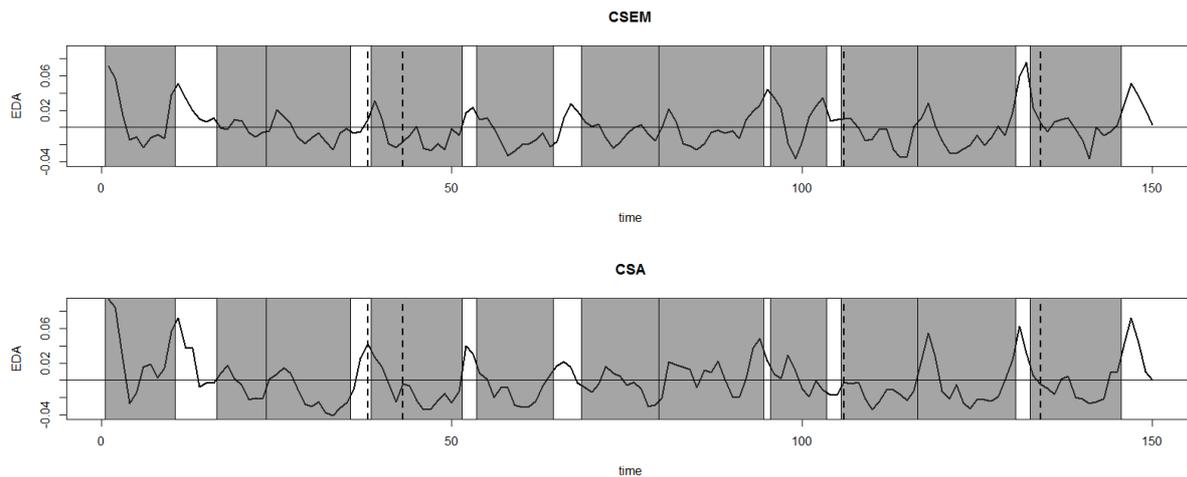


Abb. 24: Periodischer Anteil der Hautleitaktivität als Funktion der Zeit (Zeitintervalle) der Gruppen CSEM und CSA. Zur Bedeutung der Abbildung siehe Untertext Abb. 23.

### 3.4.5.3 Kontrolle der Müdigkeit (EEG)

Die Müdigkeit während der Untersuchung wurde durch ein EEG-Mass quantifiziert, welches als Indikator für „Drowsiness“ (Müdigkeit) gilt ( $\alpha + \theta/\beta$ ). Dieses Mass zeigt einen Abfall über die Zeit, was andeutet, dass die Versuchsteilnehmer im Laufe der Virtuellen Exposition eher wacher werden. Die Gruppen unterschieden sich dabei nicht voneinander.

Der mögliche Einfluss der Müdigkeit auf die Hautleitaktivität lässt sich mit Korrelationsanalysen des EEG-Müdigkeitsmasses mit den dekomponierten EDA-Komponenten klären. Dabei zeigt sich ein signifikanter Zusammenhang mit den EDA-Trendkomponenten der CSEM-Gruppe ( $r = -.20$ ,  $t = -5.1$ ,  $df = 598$ ,  $p < .001$ ) und der CTL-Gruppe ( $r = -.84$ ,  $t = -37.5$ ,  $df = 598$ ,  $p < .001$ ). Der korrelative Zusammenhang zwischen der periodischen und der Residualkomponente mit dem extrahierten EEG-Müdigkeitsmass ist dagegen vernachlässigbar gering ( $r \sim 0$ ). Daraus kann geschlossen werden, dass die weiter oben dargestellten EDA-Analysen frei von Effekten der Müdigkeit sind.

### 3.4.5.4 Habituation

Um unsere Hypothese zu prüfen, inwiefern das durch die visuellen Stimuli induzierte Erregungsniveau (gemessen an der Hautleitfähigkeit) nach wiederholter Stimulation habituiert, wurde für jede Aktivitätsspitze des periodischen EDA-Anteils und für jeden Untersuchungsteilnehmer gesondert ein Veränderungswert berechnet. Dieser Veränderungswert beruht im Wesentlichen auf den Residuen der Dekompositionsmethode und wurde durch Regression auf die vier Durchgänge bestimmt.

Tab. 9: Veränderungswerte der EDA-Aktivität (in %/100 des Skalenbereichs) über die Durchgänge gesondert nach der jeweiligen Szenerie. Negative Werte kennzeichnen eine abnehmende Tendenz. Im Einzelvergleich zeigt

sich eine signifikant stärkere Aktivitätsverminderung bei der CSO- gegenüber der CTL-Gruppe in Szene 11 (unbekleidetes Mädchen Tannerstadium 1).

Szenerie	CSO		CTL		Statistik <sup>1</sup>
	m	s	m	s	
scn 1	-0.004	0.089	-0.012	0.093	W=395.0,p=0.685
scn 2	-0.018	0.063	0.001	0.081	W=278.0,p=0.127
scn 3	-0.016	0.043	0.008	0.083	W=265.0,p=0.081
scn 4	-0.01	0.051	0.015	0.067	W=277.0,p=0.123
scn 5	-0.015	0.068	-0.001	0.062	W=354.0,p=0.797
scn 6	-0.009	0.055	-0.013	0.077	W=342.0,p=0.648
scn 7	-0.038	0.069	-0.022	0.068	W=312.0,p=0.339
scn 8	-0.029	0.08	-0.045	0.071	W=379.0,p=0.888
scn 9	-0.017	0.049	-0.016	0.085	W=299.0,p=0.241
scn 10	-0.005	0.065	-0.018	0.038	W=406.0,p=0.556
scn 11	-0.035	0.076	0.009	0.073	W=232.0,p=0.021

<sup>1</sup>Wilcoxon Rangsummen Test

In der Analyse der Veränderungswerte konnte jedoch weder ein allgemeiner Trend für die Gruppen ( $F_{1,55} = 2.38, p = .129$ ) noch für eine bestimmte Szenerie ( $F_{10,560} = 1.36, p = .194$ ) ermittelt werden. Eine differentielle Habituation die eine Diskriminierung der Gruppen ermöglichen könnte liegt demnach nicht vor. Gleiches gilt für einen Vergleich der Indexgruppen untereinander (CSEM vs. CSA). Die Daten sind aus Platzgründen hier nicht dargestellt.

### 3.4.5.5 Blickverhalten (Fixationen)

Fixationen sind durch eine Blickverweildauer von mindestens 100 ms gekennzeichnet. Mittels Eye-Tracking konnten solche Fixationen registriert werden, wenn sie in einer jeweiligen Szenerie einem bestimmten Stimulus anhafteten. Hierdurch ist es möglich, die totale Blickverweildauer über alle Durchgänge hinweg zu berechnen und sie mit den jeweiligen Eigen-

schaften der Stimuli (Tannerstadium, Geschlecht, Bekleidung/keine Bekleidung) in Verbindung zu bringen. Im Zuge dieser Berechnung findet eine Normalisierung der Fixationsdauer aufgrund der nicht gleichverteilten Stimulieigenschaften statt, so dass vereinzelt Fixationen unter 100ms beobachtet werden können.

Hypothetisch erwarteten wir einen Zusammenhang zwischen sexueller Präferenz mit der Verteilung der Fixationen auf die Merkmale der Stimuli, so dass ausschliesslich ein Vergleich zwischen den Gruppen CTL und CSO sinnvoll ist. Treffen die Hypothesen zu, dann sollten grössere Fixationszeiten mit einem höheren Interesse der Indexgruppen für unbedeckte weibliche Stimuli der Tannerstadien < 4 (vgl. Seite 19, SSPI-2) im Vergleich zu den CTL-Probanden vorherrschen. Für die CTL-Gruppe erwarteten wir dagegen eine Präferenz (grössere Fixationszeit) für unbedeckte weibliche Stimuli der Tannerstadien > 3.

### Fixationen in Abhängigkeit von Entwicklungsstadien (Tannerstadien)

Tab. 10: Mittlere Fixationszeiten (ms) in Abhängigkeit der Tannerstadien 1-5.

	CSO		CTL		Statistik <sup>1</sup>
	m	s	M	s	
tanner1	592.7	371.3	411.2	292.3	W = 357.0, p = 0.079
tanner2	925.5	444.4	841.1	457.2	W = 335.0, p = 0.194
tanner3	964	537.4	689.9	408.2	W = 358.0, p = 0.075
tanner4	670.1	471.9	391.1	359.9	W = 420.0, p = 0.002
tanner5	953.5	421.8	628.2	358	W = 386.0, p = 0.018

<sup>1</sup> Wilcoxon Rangsummen Test

Erneut führen fehlende Daten zu einer modifizierten Stichprobengrösse. Für die Analyse des Blickverhaltens standen Daten von N = 18 CTL-, N = 17 CSEM- und N = 19 CSA-Probanden zur Verfügung.

Mit Ausnahme von Stimuli des Tannerstadiums 2 betrachten CSO-Probanden im Einzelvergleich (siehe Tab. 10) die Stimuli signifikant länger als die CTL-Gruppe. Abbildung 25 zeigt das Blickverhalten gesondert nach CTL- und CSO-Gruppe und Tannerstadien über die vier Durchgänge der Virtuellen Exposition hinweg. Die Daten dieser Darstellung wurden mittels

eines gemischten Modelles modelliert, um spezifische Effekte herausarbeiten zu können. Es ist notwendig zu betonen, dass gemischte Modelle geeignet sind, Zufallskomponenten zu berücksichtigen, die auf die Selektion der vorliegenden Stichprobe zurückgeführt werden können. Damit wird das Signal/Rausch-Verhältnis optimiert und Effekte können damit verzerrungsfreier geschätzt werden. Grundsätzlich finden wir mit dieser Methode einen Effekt für die Gruppe ( $F_{1,52} = 11.9$ ,  $p = .001$ ), Tannerstadien ( $F_{4,908} = 29.3$ ,  $p < .001$ ), Verlauf ( $F_{1,908} = 45.8$ ,  $p < .001$ ) und eine Wechselwirkung Gruppe x Tannerstadium ( $F_{4,908} = 3.8$ ,  $p = .004$ ). Die längste Betrachtungszeit entfällt für die Kontrollpersonen auf Stimuli der Tannerstadien 2, 5 und 3, gefolgt von Stimuli der Tannerstadien 1 und 4. Eine ähnliche Reihenfolge lässt sich bei den Indexgruppen beobachten (5, 3, 2, 4, 1). Deren mittlere Fixationsdauer übertrifft jedoch leicht die der Kontrollprobanden. Die signifikante Interaktion Gruppe x Tannerstadium lässt darauf schliessen, dass signifikant kürzere Fixationsdauern der Indexgruppen hinsichtlich der Stimuli des Tannerstadiums 1 und 2 beobachtet werden können. Pro Durchgang nimmt die Fixationsdauer in beiden Gruppen ab.

Im Vergleich zu den Rohdaten sind die von der Modellberechnung vorhergesagten Daten damit deutlich kleiner und lassen daher auf ein sehr ungünstiges Signal/Rausch-Verhältnis schliessen. Das kann durch die vom Modell geschätzten Zufallseffekte objektiviert werden.

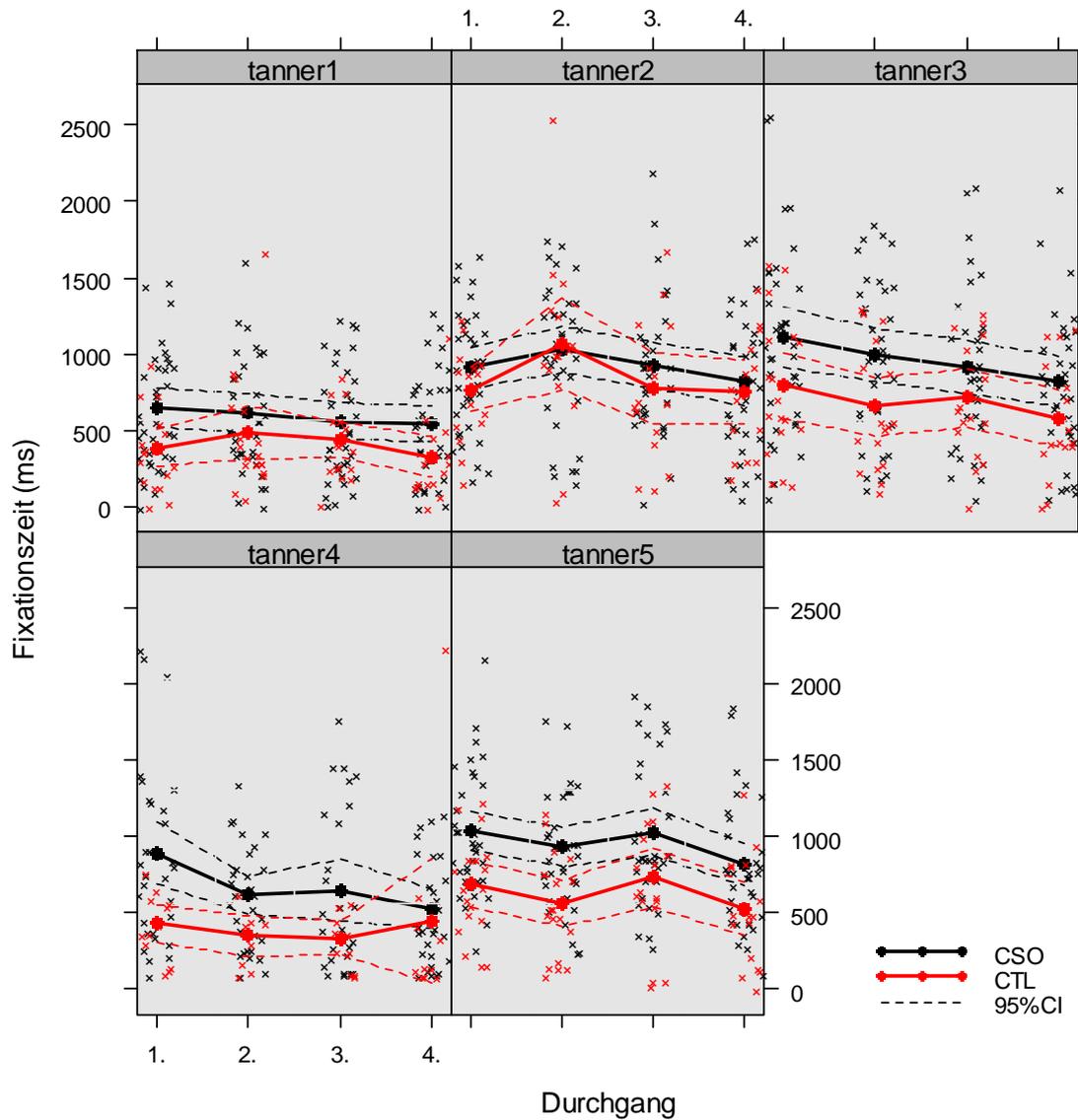


Abb. 25: Fixationszeit (ms) der beiden Gruppen CTL und CSO über die vier Durchgänge der Virtuellen Exposition. Die Unterschiede in den Fixationszeiten zwischen der CSO- und Kontrollgruppe sind hinsichtlich der Tannerstadien 1 und 2 signifikant geringer als gegenüber Stimuli der anderen Tannerstadien.

Das Ergebnis der Analyse deutet darauf hin, dass die CSO-Gruppe das Blickverhalten gegenüber Stimuli eines geringen körperlichen Entwicklungsstandes im Sinne eines Vermeidungsverhaltens unterdrückt.

## Fixationen in Abhängigkeit vom Geschlecht der Stimuli

Tab. 11: Mittlere Fixationszeiten (ms) in Abhängigkeit des Geschlechts.

	CSO		CTL		Statistik <sup>1</sup>
Geschlecht	m	s	m	s	
weiblich	854.2	435.9	580.3	326.3	$W = 435.0, p = 0.025$
männlich	728.7	383.2	497.2	308.5	$W = 457.0, p = 0.015$

<sup>1</sup> Wilcoxon Rangsummen Test

Auch hinsichtlich des Geschlechts der Stimuli ist die Blickverweildauer der CSO-Gruppe gegenüber der CTL-Gruppe im Einzelvergleich verlängert. Die Modellberechnung ergibt hier einen signifikanten Effekt für Gruppe ( $F_{1,52} = 8.7, p = .005$ ), Geschlecht der Stimuli ( $F_{1,360} = 43.6, p < .001$ ) und einen quadratischen Effekt für die Durchgänge ( $F_{1,360} = 38.6, p < .001$ ).

Die Stimuli des männlichen Geschlechts werden seitens der Kontrollprobanden und der Indexgruppen durchschnittlich etwas kürzer betrachtet als weibliche Stimuli. Insgesamt jedoch fixieren die Indexgruppen die Stimuli im Durchschnitt länger als die CTL-Gruppe. In beiden Gruppen verringert sich die Fixationsdauer über die Durchgänge hinweg. Auch hier besteht wieder ein ungünstiges Signal/Rausch-Verhältnis, wodurch die vom Modell vorhergesagten Effekte geringer sind, als sie in den Rohdaten erscheinen.

## Fixationen in Abhängigkeit von der Bekleidetheit der Stimuli

Tab. 12: Mittlere Fixationszeiten (ms) in Abhängigkeit der Bekleidetheit.

	CSO		CTL		Statistik <sup>1</sup>
Bekleidung	m	s	M	s	
bekleidet	662.1	327.2	456.4	274.7	$W = 454.0, p = 0.016$
unbekleidet	1153.6	654.4	756.2	492.5	$W = 450.0, p = 0.020$

<sup>1</sup> Wilcoxon Rangsummen Test

Erneut ist die Blickverweildauer der CSO- gegenüber der CTL-Gruppe im Einzelvergleich verlängert. Dabei werden unbekleidete Stimuli etwas länger betrachtet als bekleidete. Die

Modellberechnung zeigt erneut einen Effekt für Gruppe ( $F_{1,52} = 8.5$ ,  $p = .005$ ), für die Durchgänge ( $F_{1,359} = 15.9$ ,  $p < .001$ ) und eine Wechselwirkung Durchgang x Bekleidetheit (bekleidet vs. unbekleidet) ( $F_{1,359} = 183.2$ ,  $p < .001$ ). Die Reduktion der Fixationsdauer über die Durchgänge hinweg entwickelt sich differentiell für bekleidete und unbekleidete Stimuli. Hinsichtlich der bekleideten Stimuli vermindert sich die Fixationsdauer über die Durchgänge hinweg, für unbekleidete hingegen nimmt die Fixationsdauer pro Durchgang leicht zu. Dies verhält sich für beide Gruppen gleich.

Insgesamt zeigt das Ergebnis der Analyse, dass eine Präferenzbestimmung hinsichtlich Geschlecht und Bekleidetheit prinzipiell möglich erscheint, die Präferenz gegenüber einem Entwicklungsstadium jedoch eher mit Schwierigkeiten behaftet ist. Wenn man hingegen von der Annahme ausgeht, dass die Unterdrückung des Blickverhaltens gegenüber dem präferierten geringen Entwicklungsstadium die eigentliche Zielgröße und damit als implizites (unwillkürliches) Verhalten von Interesse ist, dann kann dies für eine Einschätzung einer pädophilen Präferenz jedoch durchaus nutzbar gemacht werden.

#### **3.4.5.6 Alternative Auswertungsmethode (Musterextraktion)**

Wie einleitend angekündigt, wurde auch ein alternativer Weg beschritten, um die für eine möglichst optimale Vorhersage der pädosexuellen Präferenz und des Übergriffsrisikos wesentliche Information aus den Daten der virtuellen Exposition zu gewinnen. Dazu wurde die Klassifikationsanalyse direkt auf die in „roher“ Form vorliegenden Daten unter Einschluss der Herzätigkeit angewandt und bestimmt, welche Variablen die höchste diskriminative Stärke besitzen.

#### **Diskrimination von CTL und CSO auf Basis sämtlicher Daten der Virtuellen Exposition**

Mittels der Klassifikationsanalyse konnten vier Variablen identifiziert werden, welche die CTL-Gruppe maximal von der CSO-Gruppe zu trennen vermochten: Dabei handelte es sich um die Kodierung für die Durchgänge, die unterschiedlichen Szenarien und um sowohl den tonischen als auch phasischen Anteil der elektrodermalen Aktivität. Da sämtliche Datenpunkte für die Klassifikation eines jeweiligen Probanden Berücksichtigung fanden (insg. 4 Durchgänge x 150 Zeitintervalle zu je 5 sek), wurden entsprechend der Vorhersage aus dem Klassifikationsalgorithmus für jeden Probanden 600 Datenpunkte danach beurteilt, ob sie eher mit der Klasse der CTL- oder eher mit der Klasse der CSO-Gruppe vereinbar waren. Dies ermöglichte es aus den tonischen und den phasischen Anteilen der elektrodermalen Aktivität ein Muster zu extrahieren, das für die Klassifikation der jeweiligen Gruppen geeignet ist. Abbildung 26 zeigt die so gewonnenen Muster. Initial (1. Szenerie und 1. Durchgang) ist über beide Gruppen eine maximale elektrodermale Aktivität zu beobachten (sowohl phasisch als

auch tonisch), die im weiteren zeitlichen Verlauf deutlich abnimmt. Ab Beginn des Durchgangs drei intensiviert sich jedoch die Aktivität bei den Kontrollprobanden (stärker hinsichtlich des phasischen als des tonischen Anteils), während ähnliches bei den Indexgruppen nicht zu beobachten ist. Erneut können keine eindeutigen Habituationismuster entlang der elf Szenerien beobachtet werden. Die Muster besitzen jedoch höchste Diskriminabilität ( $AUC = .99$ ). Sowohl 95% der Probanden der Indexgruppen als auch der Kontrollprobanden werden korrekt klassifiziert. Somit haben diese elektrodermalen Aktivitätsmuster eine grosse Bedeutung für die pädosexuelle Präferenz.

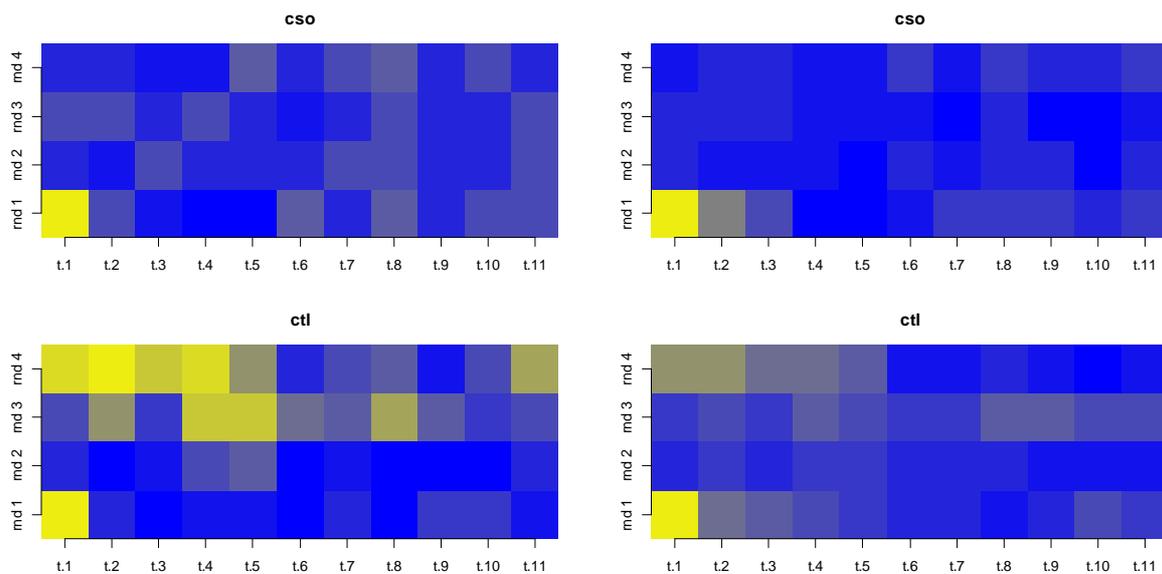


Abb. 26: Phasische (links) und tonische (rechts) elektrodermale Aktivität entsprechend der Vorhersage des Klassifikationsalgorithmus über die Abfolge der elf Szenerien (t.1 – t.11) und geschichtet nach den Durchgängen 1 – 4 (rnd 1 – 4) im Rahmen der virtuellen Exposition. Die farbliche Graduierung von tiefem Blau bis hellem Gelb kodiert die Intensität der elektrodermalen Aktivität, wobei blau für geringe und gelb für hohe Aktivität steht. Die zeitliche Abfolge verläuft von links unten nach rechts oben.

## Diskrimination von CSEM und CSA auf Basis sämtlicher Daten der Virtuellen Exposition

In analoger Weise wurden elektrodermale Aktivitätsmuster für die Binnenklassifikation innerhalb der Indexgruppen gewonnen. Die Klassifikationsanalyse identifizierte jedoch zusätzlich noch die Herzaktivität als wichtigen Indikator. Daraus ergaben sich die Muster aus Abbildung 27. Erneut finden sich initial hohe Aktivitätsmuster, die sich danach vermindern. Entgegen der CSA-Probanden zeigt die Gruppe der CSEM intermittierende Aktivitätsspitzen im Bereich der elektrodermalen Aktivität (phasisch stärker als tonisch). Im Bereich der Herztätigkeit lässt sich weniger gut eine Systematik erkennen. Jedoch im letzten Durchgang scheint nochmals eine szenenspezifische Aktivierung bei den jeweiligen Gruppen einzutreten.

Auch hier zeigen die extrahierten Muster eine maximale Diskriminabilität (AUC = 1.0). Alle Probanden mit Ausnahme eines CSEM-Probanden können korrekt klassifiziert werden.

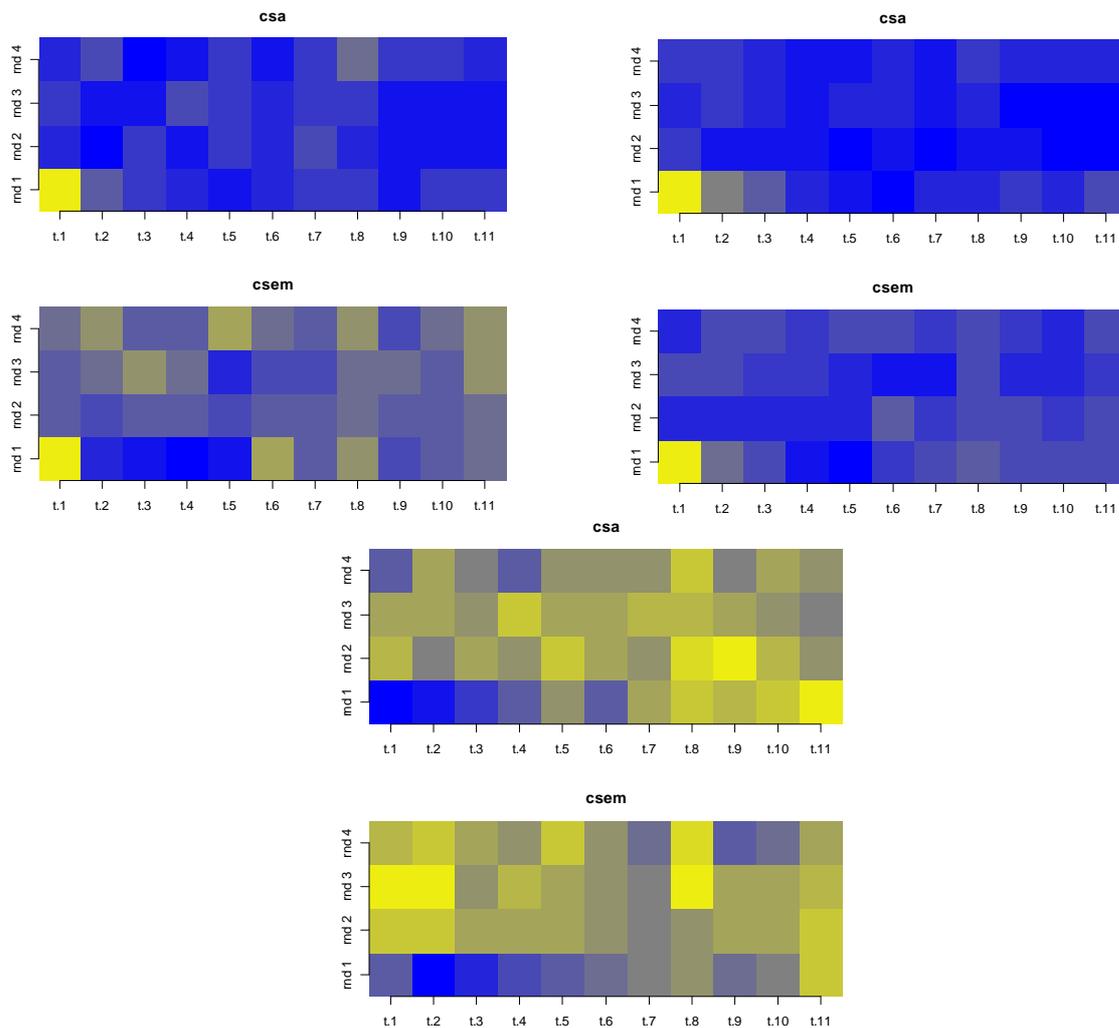


Abb. 27: Phasische (oben links) tonische (oben rechts) elektrodermale Aktivität und die Herzrätigkeit (Inter-Beat-Intervall, Herzschlagspanne) entsprechend der Vorhersage des Klassifikationsalgorithmus über die Abfolge der elf Szenarien (t.1 – t.11) und geschichtet nach den Durchgängen 1 – 4 (rnd 1 – 4) im Rahmen der Virtuellen Exposition. Entgegen der elektrodermalen Aktivität kodiert gelb Herzrätigkeit geringer und blau hoher Intensität.

Diese alternativen Auswertemethoden zeigt, dass die Daten der Virtuellen Exposition über ein erhebliches Potential für die Vorhersage der pädosexuellen Präferenz aber auch des Übergriffsrisikos verfügen. Dieses Potential konnte aus den zuvor dargestellten Analysen nicht abgeleitet werden. Dem entgegen liegt jedoch eine schlüssige inhaltliche Interpretation der Muster nicht auf der Hand. Es ist keineswegs klar, inwiefern ihnen etwas Spezifisches für pädosexuelle Präferenz oder gegenüber sexuellen Übergriffen anhaftet. Diese Frage konnte auch in der für die Datenanalyse vorgesehen Zeit nicht mehr sinnvoll vertieft werden. Sie wird aber auf jeden Fall Gegenstand weiterer Forschungsbemühungen bleiben.

### 3.5 Gesamtklassifikation

Unter den jeweiligen Prädiktoren der unterschiedlichen Datenbereiche (Neuropsychologie, Neurophysiologie, Implizite Verfahren) konnten im Zuge der Gesamtklassifikation sechs Variablen identifiziert werden, die geeignet erscheinen die Indexgruppen von der Kontrollgruppe zu trennen um damit Aussagen über die sexuelle Präferenz treffen zu können. Weitere drei Variable können mit hoher Genauigkeit die Indexgruppen untereinander in die CSEM- und CSA-Gruppe klassifizieren, wodurch wir Rückschlüsse über das Risikoverhalten erhoffen.

#### 3.5.1 Vorhersage der sexuellen Präferenz (Klassifikation von CTL und CSO)

Die aus der Klassifikationsanalyse hervorgegangen Variablen erstrecken sich auf zwei Verfahren der Neuropsychologie und auf vier spezifische EKPs des Semantic Misattribution-Paradigmas (SMP).

Unter den neuropsychologischen Verfahren ist es der kurzfristige freie Abruf einer zuvor gelernten Liste mit 16 Wörtern und Impulsivität im Zusammenhang mit Risikoentscheidungen, die zur Diskrimination der Indexgruppen beitragen. Im Zusammenhang mit dem SMP konnten vier EKPs identifiziert werden, die für frühe Informationsverarbeitungs-komponenten kodieren und an hochzentralen und okzipitalen (Hinterhaupt) Elektroden abgeleitet wurden. Dabei handelte es sich bei allen vier EKPs um Potentiale, die unter subliminaler Darbietung des visuellen Reizes (NRP-Bild) abgeleitet wurden. Die visuellen Reize ihrerseits erstreckten sich in drei Fällen auf die Darstellung von weiblichen Abbildungen der Tannerstadien 3 und 5 und in einem Fall auf eine männliche Abbildung der Tannerstadiums 1.

Im jeweiligen Einzelfall kommt den sechs Variablen allenfalls eine mässige Klassifikations-tüchtigkeit zu. Abbildung 28 zeigt eine sogenannte "Receiver Operating Characteristic" (ROC). Dort wird die Sensitivität und Spezifität abgebildet mit der unterschiedliche Werte einer Variablen eine Klassifikation ermöglichen. Dabei ist die Sensitivität der prozentuale Anteil korrekt klassifizierter Probanden der Indexgruppen und die Spezifität der entsprechende Anteil korrekter Klassifikationen der Kontrollgruppe.

Je weiter sich die ROC-Kurve einer Variablen in die linke obere Ecke erstreckt, desto besser ist dabei ihre Klassifikationsfähigkeit. Je näher sie an der mittleren Diagonale liegt, je schlechter ist sie. Dies ist gleichbedeutend mit der Grösse der Fläche, die sich unter der Kurve (AUC) berechnen lässt. Ist eine Variable zu 100% trennscharf, dann liegt dieser Wert bei  $AUC = 1.0$  und in umgekehrtem Falle bei  $AUC = 0.5$ . Ein  $AUC > 0.8$  gilt als guter Wert.

Wie der Abbildung zu entnehmen ist erreicht keine der Variablen einen solchen Wert. Kombiniert in der Gesamtklassifikation (CLASS) erreicht die Fläche unter der ROC-Kurve jedoch einen Wert von  $AUC = .87$ .

Tab. 13: Klassifikationstabelle CSO vs. CTL.

		beobachtet	
		CSO	CTL
vorher-gesagt	CSO	41	7
	CTL	2	14

Durch die Gesamtklassifikation können 41 von 43 CSO-Probanden und 14 von 21 CTL-Probanden korrekt klassifiziert werden. Das entspricht einer Sensitivität von 95% und einer Spezifität von 67%. 86% der Probanden wurden korrekt klassifiziert.

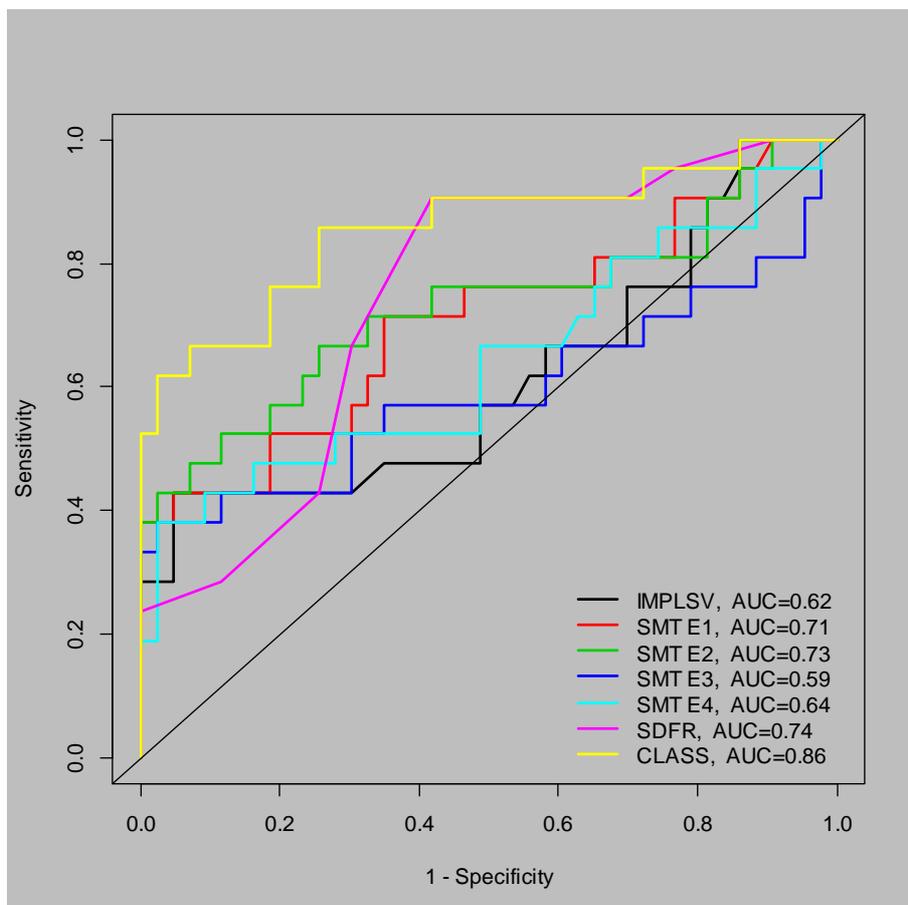


Abb. 28: Die „Receiver Operating Characteristic“ bildet die Diskriminabilität der gewählten Variablen ab. Die Kombination der Variablen (CLASS) ist erfolgreicher in der Klassifikation von CTL- und CSO-Probanden, wie jede

Variable alleine. IMPLSV: Impulsivität hinsichtlich einer Risikoentscheidung, SDFR: kurzfristiger freier Abruf nach Erlernen einer Wortliste, Semantic Misattribution-Paradigma EKPs - SMP E1: (Cz) männlich, Tannerstadium 1, SMP E2: (Cz) weiblich, Tannerstadium 5, SMP E3: (O1) weiblich, Tannerstadium 3, SMP E4: (Oz) weiblich Tannerstadium 3.

Insgesamt werden zufriedenstellende Klassifikationsergebnisse erzielt, die jedoch nur wenig über das Diskriminationsvermögen hinausgehen, welches schon durch die Betrachtung ausschliesslich neuropsychologischer Variablen erreicht werden kann. Die Auswahl ausschliesslich neuropsychologischer und neurophysiologischer Variablen und die Zurückweisung der Impliziten Verfahren (IAT, SMP) zur optimalen Klassifikation von Kontrollprobanden einerseits und Indexgruppen andererseits zeigt, dass das zugrundeliegende Selektionskriterium nicht spezifisch für die sexuelle Präferenz ist, sondern vielmehr einem Muster folgt, das durch hirnganische Defizite gekennzeichnet ist.

Wird das Klassifikationsverfahren unter ausschliesslicher Berücksichtigung der für die sexuelle Präferenz spezifischen Variablen (implizite Verfahren & Virtuelle Exposition EDA) wiederholt, so führt dies zur Auswahl des IAT und eines EDA-Musters der Virtuellen Exposition, die jedoch gemeinsam keine bessere Diskriminabilität ( $AUC = .76$ ) erreichen, als der IAT alleine. Daraus kann geschlossen werden, dass keine Kombination spezifischer Verfahren die Diskriminabilität der Einzelverfahren übertrifft, um damit sexuelle Präferenz besser bestimmen zu können.

### **3.5.2 Vorhersage des Risikoverhaltens (Klassifikation von CSEM und CSA)**

Hinsichtlich der Gesamtklassifikation der Indexgruppen untereinander wurden abschliessend drei Variablen selektiert, von denen wiederum eine aus einem neuropsychologischen Verfahren abgeleitet werden konnte und für die Schwierigkeit im Unterdrücken unangemessener Verhaltenstendenzen kodiert. Die anderen beiden jedoch sind EDA-Muster, die im Rahmen der Virtuellen Exposition im Zusammenhang mit einem bekleideten Mädchen (Tannerstadium 1) und einer Gruppe von vier badenden und bekleideten Kindern (Jungen und Mädchen, Tannerstadium 1) abgeleitet werden konnten. Die ROC-Analyse zeigt erneut, dass die Variablen jeweils für sich genommen nicht die klassifikatorische Tüchtigkeit haben, wie in einer kombinierten Vorhersage (CLASS,  $AUC = .87$ ).

Tab. 14: Klassifikationstabelle CSA vs. CSEM

		beobachtet	
		CSA	CSEM
Vorher-gesagt	CSA	19	5
	CSEM	3	16

Durch die Gesamtklassifikation (vgl. Tab. 14) können 16 von 21 CSEM- und 19 von 22 CSA-Probanden korrekt klassifiziert werden. Dies entspricht einer Sensitivität von 86% und einer Spezifität von 76%. 81% der Probanden wurden korrekt klassifiziert.

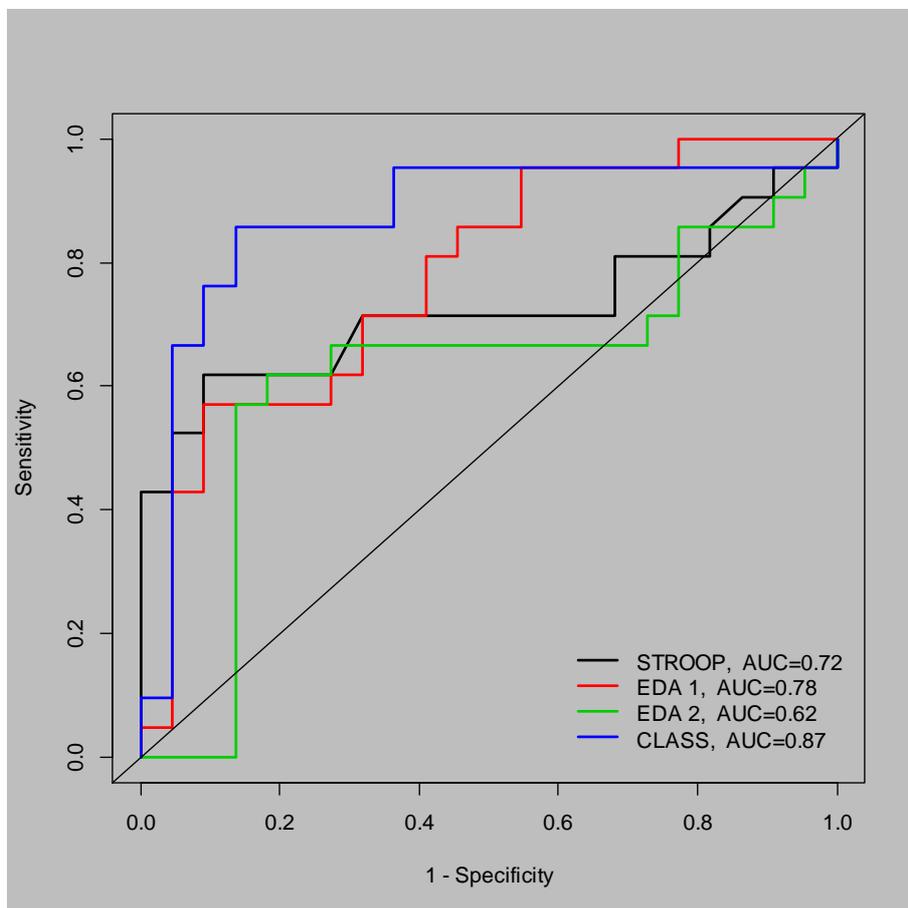


Abb. 29: Die „Receiver Operating Characteristic“ bildet die Diskriminabilität der gewählten Variablen ab. Die Kombination der Variablen (CLASS) ist erneut erfolgreicher bei der Klassifikation von CSEM- und CSA-Probanden, wie jede Variable alleine. STROOP: Unterdrückung unangemessener Verhaltenstendenzen, EDA 1: Hautleitaktivität im Zusammenhang mit einer Gruppe von vier badenden und bekleideten Kindern (Jungen, Mädchen, Tannerstadium 1), EDA 2: Hautleitaktivität im Zusammenhang mit einem einzelnen bekleideten Mädchen (Tannerstadium 1).

Entgegen der Vorhersage der sexuellen Präferenz verhält es sich bei der Klassifikation innerhalb der Indexgruppen in eine CSA- und eine CSEM-Gruppe anders. Hier werden neben einer neuropsychologischen Variablen auch EDA-Muster selektiert, die auf spezifische visuelle Stimuli zurückgeführt werden können. Die beiden Muster sind im Vergleich der beiden Indexgruppen komplementär, CSA-Probanden reagieren stärker auf eine Gruppe badender Kinder (Mädchen und Jungen, Tannerstadium 1), während CSEM-Probanden stärker auf ein allein dargestelltes Mädchen (Tannerstadium 1) reagieren. Die Auswahl einer neuropsychologischen Variablen zeigt aber auch hinsichtlich dieser Klassifikationsanalyse, dass auch hier die Hirnorganik eine Rolle bei der Zuordnung spielt.

## **4. Diskussion**

### **4.1 Schwierigkeiten bei der Versuchsdurchführung**

Zwei Ereignisse erschwerten den Vertrauensgewinn der Indexgruppen und somit den Rekrutierungsverlauf neben den bereits beschriebenen Verzögerungen bei der Terminierung zusätzlich: Die Annahme der Schweizerischen Volksinitiative „Pädophile sollen nicht mehr mit Kindern arbeiten dürfen“ vom 18. Mai 2014 und deren politische Kampagnen im Vorfeld, sowie die voreilige Berichterstattung über unser Projekt unter dem sinnwidrigen Schlagwort „Pädophilendetektor“ im April 2015.

Die unmittelbaren Folgen der Berichterstattung erstreckten sich zunächst vor allem auf das Verhältnis zuweisender Institutionen und Fachpersonen. Es bedurfte persönlicher Gespräche, um Bedenken hinsichtlich der weiteren Kooperation zu zerstreuen. In einem Fall kam es zum Kontaktabbruch eines potentiellen Probanden. Darüber hinaus bestand aufgrund des kompetitiven Drucks in der medialen Berichterstattung auch das Risiko einer inkognito Teilnahme von Journalisten an der MIPS-Studie. Diesem Risiko wurde mit einer juristisch verbindlichen Verschwiegenheitsklausel entgegengewirkt. Die öffentliche Berichterstattung machte die weitere Rekrutierung mittels der bisherigen Coverstory obsolet. Die Studienziele wurden demzufolge ab diesem Zeitpunkt (April 2015) zu Beginn einer Untersuchung offen gelegt und die Informationsmaterialien entsprechend angepasst. Insgesamt zogen die politischen „Anti-Pädophilie“-Kampagnen und die Medienberichterstattungen gewisse Rekrutierungsprobleme nach sich, jedoch weniger als befürchtet. Zudem wurden bei einzelnen Probanden veränderte Erwartungen sichtbar (z.B. die Erwartung, eine pädophile Neigung nachweisen oder ausschliessen zu können).

Die erschwerte Rekrutierung, bedingt durch die politischen Kampagnen, die Berichterstattung, sowie durch einen unfallbedingten Ausfall einer Mitarbeiterin führte dazu, dass wir im

Oktober 2015 eine Verlängerung der Projektförderung bis Ende April 2016 beantragen mussten, welche im Dezember 2015 genehmigt wurde.

Zuvor stellte die Aufdeckung der Coverstory einen heiklen Punkt in der Untersuchung dar. Die Offenlegung des Studienziels und die Erklärung der einzelnen Tests führten dabei zu sehr unterschiedlichen Reaktionen. Einige der Probanden interessierten sich wenig für die Studie und gaben an, lediglich aufgrund der Bezahlung teilzunehmen. Drei Probanden aus den Indexgruppen waren dagegen initial über den Gebrauch einer Coverstory empört und fühlten sich dadurch hinters Licht geführt. Nach einem erklärenden Gespräch stimmten diese Probanden aber einer Verwendung der erhobenen Daten zu. Gesamthaft waren die Reaktionen jedoch sehr wohlwollend und die grosse Mehrheit der Teilnehmer gab an, den wissenschaftlichen Ansatz der Erforschung des Themenbereichs der sexuellen Präferenzstörung, insbesondere der Pädophilie unterstützen zu wollen.

## **4.2 Diskussion der Resultate**

Zusammenfassend lässt sich feststellen, dass die in der MIPS-Studie („Prognostisch und Therapeutisch Relevante Messbare Indikatoren bei Pädosexuellen Straftätern“) untersuchten N = 22 Sexualstraftäter mit zurückliegendem Kindesmissbrauch (CSA) zunächst durch ein gehäuftes Auftreten von vorrangig antisozialen Persönlichkeitsstörungen gekennzeichnet waren und überdurchschnittlichen häufig durch emotionalen Missbrauch, physische Vernachlässigung und Missbrauch in der Kindheit traumatisiert worden waren. Dasselbe überproportionale Traumatisierungsmuster konnte zwar auch bei den N = 21 Konsumenten von Internetpornographie (CSEM) gefunden werden, jedoch waren etwas weniger Probanden von einer Persönlichkeitsstörung betroffen. Probanden beider Indexgruppen berichteten auch deutlich vermehrt als Kontrollprobanden (CTL) unter psychiatrischen Problemen insbesondere depressiven Symptomen und interpersonellen Problemen zu leiden. CSEM-Probanden berichteten darüber hinaus auch phobische und paranoide Symptome, während CSA-Probanden Angstsymptome und einen vermehrten Psychotizismus angaben. Das Vorliegen von Persönlichkeitsstörungen innerhalb der Indexgruppen dürfte sich auch auf die dimensionale Charakterisierung durch die sogenannten „Big-5“-Persönlichkeitsfaktoren ausgewirkt haben, die zwischen den Probanden der Indexgruppen relativ gleichförmig ausgeprägt sind und sich von den Kontrollprobanden in einer verringerten Gewissenhaftigkeit und in einem erhöhten Neurotizismus unterscheiden.

In keiner der beiden Indexgruppen liessen sich Hinweise für das aktuell erhöhte Auftreten psychopathischer Charakteristika finden. Die Psychopathie-Checkliste (PCL:SV) lieferte lediglich Hinweise auf ein überdurchschnittliches antisoziales Verhalten in der Adoleszenz und

im Erwachsenenalter auf Seiten der Sexualstraftäter insgesamt (CSO-Gruppe) (etwas akzentuierter bei den CSA-Probanden, was mit dem gehäuftem Auftreten der antisozialen Persönlichkeitsstörung in der Gruppe korrespondiert). Dies kann vor dem Hintergrund einer expliziten Low-Risk-Strategie bei der Rekrutierung der CSA-Gruppe nicht sonderlich überraschen.

Etwa 75% der Indexgruppen berichteten von einer heterosexuellen Ansprechbarkeit. Dies korrespondierte mit dem Geschlecht der Mehrzahl der Opfer seitens der CSA-Gruppe, die im Übrigen auch zumeist jünger als 11 Jahre waren. Knapp die Hälfte der CSA-Probanden hatten mehrere Opfer missbraucht. Die Probanden beider Indexgruppen berichteten, sich gleich häufig in sexuellen Fantasien mit Kindern zu ergehen. Insgesamt können die Berichte als zuverlässig betrachtet werden, da mehrere, in unterschiedlichen Fragebögen enthaltene Lügenskalen, keine Indikatoren für das Gegenteil lieferten. Daraus kann darüber hinaus auch auf eine hohe Compliance geschlossen werden.

Die neurophysiologischen und neuropsychologischen Daten legen hinsichtlich der Indexgruppen das Vorhandensein moderater aber signifikanter kognitiver Defizite nahe. Insgesamt scheinen aber die Defizite beider Gruppen von einander dissoziiert zu sein: Während die CSEM-Gruppe vorrangig durch verminderte Gedächtnisleistungen charakterisiert sind, findet sich hinsichtlich der CSA-Gruppe eine erhöhte Impulsivität und ein vermindertes Fehlerbewusstsein.

Bevor die Resultate, einschliesslich der Klassifikationsanalysen der MIPS-Studie, interpretiert werden können, ist es notwendig über Studien immanente Einschränkungen zu berichten, deren Vernachlässigung ansonsten zu Fehlinterpretationen führen könnte. Im Hinblick auf die demographischen Charakteristika ergeben sich zwar über die meisten Bereiche wie z.B. sozioökonomischer Status und Intelligenz keine Unterschiede zwischen den Gruppen. Gleichwohl besteht ein Altersunterschied von zehn Jahren zwischen den Indexgruppen und den jüngeren Kontrollprobanden, der auf eine sequentielle Rekrutierungsstrategie (erste Studien-Phase CTL, anschliessend CSO) bei gleichzeitig geringen Spielräumen in der Rekrutierung der Indexgruppen zurückzuführen ist. Darüber hinaus liess sich auch beobachten, dass in einigen Fällen (über alle Gruppen) verteilt positive Cannabisbefunde festgestellt werden können. Desgleichen findet sich innerhalb der Indexgruppen eine höhere Rate in der Einnahme psychoaktiver Substanzen (z.B. Benzodiazepine, Neuroleptika, Antidepressiva). Diese Faktoren sind geeignet vor allem die neuropsychologischen und neurophysiologischen Daten zu verzerren, wenngleich der Einfluss der psychoaktiven Substanzen aufgrund der relativ wenigen Fälle eher gering bleiben sollte. Das Alter der Probanden jedoch zeigt kleine

bis mittlere Effekte hinsichtlich des P100 ereigniskorrelierten Potentials (EKPs) im Zusammenhang mit dem „Semantic Misattribution-Paradigma“ (SMP), kaum jedoch hinsichtlich der Neuropsychologie. Die relative Homogenität der Gruppen hinsichtlich der demographischen Charakteristika könnten durch höhere Stichprobengrößen erschüttert werden, da dann die statistische Messkraft stärker wird und damit kleinere Unterschiede zwischen den Gruppen abgesichert werden könnten. Grundsätzlich gilt dies für alle Effekte, die in der MIPS-Studie berichtet werden können. Gegen die relativ kleinen Stichprobenumfänge von  $N \sim 20$  pro Gruppe könnte man daher kritische Einwände erheben. Da die MIPS-Studie jedoch das Ziel verfolgt, klinisch-forensisch relevante Aspekte zu eruieren, können darunter naturgemäss nur grosse Gruppenunterschiede fallen, zumal die Effekte ja auch im Einzelfall sichtbar sein sollen, wodurch die Stichprobengröße von  $N \sim 20$  gerechtfertigt ist. Weit kritischere Erwägungen betrifft die Definition der Gruppen. Für die CSEM-Gruppe gilt das Ausschlusskriterium des Kindesmissbrauchs, hingegen ist der Konsum illegaler Internetpornographie kein Ausschlussgrund für die CSA-Gruppe. Tatsächlich berichten auch ca. 50% der CSA-Probanden den Konsum illegaler Internetpornographie. Umgekehrt bleiben wir bei den CSEM-Probanden jedoch ausschliesslich auf die Zuverlässigkeit und Ehrlichkeit der Aussagen angewiesen, dass ein sexueller Übergriff auf Kinder in der Vergangenheit tatsächlich unterblieben ist. Unaufgeklärte Übergriffe seitens der CSEM-Probanden könnten daher zu einer Nivellierung der Gruppen führen. Eine gewisse Heterogenität innerhalb der Gruppe der CSA ist auch aus dem unterschiedlich fortgeschrittenen Therapieprozess zu erwarten, in den die Probanden innerhalb ihrer Massnahmen jeweils eingebunden sind. Dieser Einfluss könnte vor allem bei Effekten zum Tragen kommen, die gegenüber dem therapeutischen Prozess veränderungssensitiv sind. Schliesslich muss noch auf die Qualität des visuellen Stimulusmaterials eingegangen werden: Medien mit erotischen Inhalten in höchster Qualität sind heute mittels des Internets frei zugänglich, wogegen die NRP („Not-Real People“)-Bilder qualitativ deutlich abfallen. Dieser Kompromiss wurde eingegangen, um die Vergleichbarkeit mit existierenden internationalen Studien nicht zu kompromittieren. Dieser Qualitätsaspekt und die wiederholte Darbietung des Stimulusmaterials könnte zu einer Verminderung ihrer Wirksamkeit in den jeweils eingesetzten Verfahren (Semantic Misattribution-Paradigma (SMP), Impliziter Assoziationstest (IAT), Virtuelle Exposition (VE), Choice Reaction Time (CRT)-Aufgabe, Betrachtungszeit (VT)) geführt haben. Offenbar ist dieser Umstand besonders kritisch für die SMP, wo die abgeleiteten EKPs hauptsächlich die physikalischen Eigenschaften des Bildmaterials kodieren und auch die Verhaltensdaten keinen Priming-Effekt zeigen konnten. Ebenso davon betroffen könnte auch die CRT-Aufgabe sein, welche gleichermassen nur ein Null-Ergebnis lieferte. Abschliessend muss noch erwähnt werden, dass es u.U. einen Einfluss der Presseveröffentlichung auf die Selektion der Probanden gegeben haben könnte: Wir konnten nämlich beobachten, dass die Häufigkeiten, mit welcher ein chinesisches

Schriftzeichen als „sexuell“ konnotiert wurde, vor dem Pressebericht differentiell durch die beiden Indexgruppen beurteilt wurde ( $N = 12$  CSEM  $>$   $N = 16$  CSA) und sich bei den danach rekrutierten  $N = 14$  Probanden nivelliert hat, so dass vorher zwar ein Unterschied, nachher aber kein Unterschied mehr zwischen den Gruppen bestand.

Einer der Hauptbefunde der MIPS-Studie ist die Beobachtung nicht unerheblicher und differentieller kognitiver Defizite bei den Probanden der Indexgruppen: Erhöhte Impulsivität nicht nur aber auch im Zusammenhang mit risikobehafteten Entscheidung und die Bereitschaft zur Übernahme riskanter Entscheidungen bei gleichzeitigem vermindertem Fehlerbewusstsein (Unrechtsbewusstsein?) korrespondieren mit einer erhöhten Rate antisozialer Persönlichkeitsstörungen innerhalb der CSA-Gruppe. Dies entspricht einem Muster, welches im Zusammenhang mit psychopathischen Persönlichkeitszügen diskutiert wird, deren Vorliegen wir aufgrund der Ausprägung der PCL-Werte hinsichtlich CSA-Gruppe jedoch ausschliessen können. Kognitive Defizite bei pädophilen Sexualstraftäter werden auch in der Fachliteratur beschrieben (Cantor et al., 2004; Kruger & Schiffer, 2011b; Schiffer & Vonlaufen, 2011), auch im Zusammenhang mit den exekutiven Funktionen (Adjorlolo & Egbenya, 2016), deren Integrität über das Vermögen zu zielorientiertem und rationalem Verhalten entscheidet. Darüber hinaus zeigen bildgebende Studien verminderte graue und weisse Substanz in Gehirnen von Probanden mit pädophiler Neigung (Cantor et al., 2008; Poepl et al., 2013; Schiffer et al., 2008; Schiltz, Witzel, Northoff, & et al., 2007). Insgesamt zeichnen diese Befunde ein Bild von pädophilen Sexualstraftätern, welches auf eine zugrundeliegende hirnrorganische Entwicklungsstörung hinweist. Dies ist auch in Übereinstimmung mit den von uns berichteten Befunden und wird komplettiert durch die von den CSA-Probanden berichteten vermehrten psychiatrischen Problemen.

Die bisweilen phänomenologische Ähnlichkeit mit psychopathischen Persönlichkeitszügen legt eine Störung des ventromedialen-orbitofrontalen Kortex nahe, die mit dem Bild der „erworbenen Soziopathie“ korrespondiert (Anderson, Bechara, Damasio, Tranel, & Damasio, 1999; Saver & Damasio, 1991), aber entgegen der Psychopathie darüber hinaus auch mit interpersonellen Problemen in Verbindung steht. Störungen des ventromedialen-orbitofrontalen Kortex (Danek & Göhringer, 2005) sind vor allem mit emotionaler und motivationaler Dysregulation assoziiert. So ist beispielsweise die Kompetenz zur „Theory of Mind“ respektive die kognitive Empathie beeinträchtigt.

Die kognitiven Defizite der CSEM-Gruppe haben dagegen ihren Schwerpunkt im Bereich der Gedächtnisleistung. Auch neurophysiologisch sind nicht dieselben Defizite zu beobachten. So kann ein reduziertes Fehlerbewusstsein aufgrund der Pe-Komponente des Go/NoGo-

Paradigmas nicht nachgewiesen werden. Insgesamt ist aber die Beobachtung einer verminderten kognitiven Leistungsfähigkeit dennoch überraschend, zumal diese Gruppe in Abgrenzung zu der CSA-Gruppe in der Fachliteratur eher mit höherem sozioökonomischem Status, höherer Ausbildung und höherer Intelligenz in Zusammenhang gebracht wird (K. M. Babchishin, Hanson, & Hermann, 2011; Kelly M. Babchishin, Hanson, & VanZuylen, 2015; Meridian, Wilson, & Boer, 2009). Dies ist nur schwer mit kognitiven Störungen vereinbar. Darüber hinaus lassen auch die demographischen Charakteristika der CSEM-Probanden nicht auf eine besonders hohe Funktionalität dieser Gruppe schließen. Die Diskrepanz zwischen den Beobachtungen in der MIPS-Studie und der Fachliteratur könnte Ausdruck eines Alterskohorteneffektes sein, der mit der Zugänglichkeit zur Internet-Technologie in Verbindung steht. Noch vor wenigen Jahren war der Zugang zu Inhalten des World Wide Web insgesamt weniger und auch eher technisch affinen Menschen vorbehalten, wodurch ein Selektionseffekt wirksam wurde, der einen sozioökonomisch eher privilegierten Personenkreis betroffen hat. Durch die inzwischen hohe Verfügbarkeit und den immer leichter werdenden Zugang zur Internet-Technologie, entfällt nun dieser Effekt. Entgegen der Gruppe der CSA bleibt die Gesamtcharakterisierung der CSEM-Gruppe eher unscharf.

Die Klassifikationsanalysen wurden mit dem Ziel durchgeführt, eine Kombination von Verfahren zu identifizieren, deren Information geeignet ist (1) pädosexuelle Präferenz und ggf. ihre Ausprägung festzustellen und (2) Risikofaktoren zu identifizieren, die eine Transition von der CSEM-Gruppe zur CSA-Gruppe vorhersagen könnten.

Zur Identifikation der sexuellen Präferenz wurden sogenannte implizite Verfahren implementiert, deren Wirksamkeit aus der Fachliteratur bekannt ist. So z.B. der IAT, die CRT-Aufgabe, der VT-Test und eine modifizierte Variante des SMPs. Darüber hinaus entwickelten wir einen vollkommen neuen Ansatz für ein implizites Testverfahren, das durch den Einsatz der Technologie der Virtuellen Realität auf einer besonders hohen ökologischen Validität beruht, d.h. besonders lebensnah sein sollte (VE).

Wie schon berichtet, lieferten das SMP und die CRT-Aufgabe lediglich Null-Resultate. Eini-germassen diskriminierend, im Sinne unserer Hypothesenbildung, waren hingegen der IAT, das VT-Paradigma und die VE, deren Diskriminabilität jeweils isoliert betrachtet in einem leicht suboptimalen Bereich ( $AUC < .80$ ) lag, aber auch von vergleichbaren Verfahren wie in der Fachliteratur berichtet nicht wesentlich übertroffen wurde (Akerman & Beech, 2012; Mokros et al., 2013). Entgegen der Verfahrensweise die Bild-Stimuli im Rahmen des VT-Paradigmas nach ihrer sexuellen Attraktivität zu bewerten (Banse, Schmidt, & Clarbour, 2010; Mokros et al., 2013), wurden die Teilnehmer der MIPS-Studie instruiert die „Ästhetik“

des Bildes zu qualifizieren, was auf Seiten der Kontrollprobanden einen Intelligenzeffekt bezüglich Reaktionszeiten induziert haben könnte. Entgegen den von Mokros und Kollegen (Mokros et al., 2013) berichteten Daten, wonach das Betrachtungszeitmuster von pädophilen Sexualstraftäter maximal für Abbildungen der Tannerstadien  $< 5$  und das der Kontrollprobanden maximal für Abbildungen der Tannerstadien  $5$  ist, finden wir hypothesenkonforme Muster lediglich bei den Kontrollprobanden der MIPS-Studie. Durch die Induktion eines Intelligenzeffektes konnte aber somit allein aufgrund der Kenntnis der Intelligenz des jeweiligen Teilnehmers ein Betrachtungsmuster in Abhängigkeit der Tannerstadien vorausgesagt werden, welches wiederum die Klassifikationsgenauigkeit des Verfahrens steigerte. Mokros und Kollegen berichten für ein multivariates Verfahren der Klassifikation, das sowohl auf den Betrachtungszeiten, als auch auf der Beurteilung der Bilder beruht, eine Diskriminabilität von  $AUC = .80$  (Mokros et al., 2013). Wir erreichen hingegen, unter Berücksichtigung der Intelligenz, eine  $AUC = .77$ .

Die Diskriminabilität des IATs ist in etwa vergleichbar mit der des VT- Paradigmas. Jedoch beobachteten wir einen Einfluss der kognitiven Umstellfähigkeit auf das Ergebnis, dessen Berücksichtigung bei der Klassifikationsanalyse zu einer leichten Verminderung der Diskriminabilität führt. Unser Effekt ist jedoch etwa um  $1/3$  geringer ( $d \sim .41$ ) als der, in einer Metaanalyse berichtete IAT-Effekt ( $d = .63$ ) (Kelly M. Babchishin, Nunes, & Hermann, 2012).

Auch mit der VE wurde eine ähnliche Diskriminabilität erreicht. Dort wurden drei Datenmuster der elektrodermalen Aktivität (EDA) extrahiert, die mit zwei Szenarien korrespondieren, die zum einen zwei unbedeckte erwachsene Frauen und zum anderen ein unbedecktes Mädchen des Tannerstadiums 1 zeigen. Entgegen unserer Erwartungen ist die Hautleitaktivität der Kontrollprobanden im Zusammenhang mit dem unbedeckten Mädchen höher als bei den Indexgruppen. Genau umgekehrt verhält sich dies im Zusammenhang mit den erwachsenen Frauen. Eine schlüssige Interpretation liegt hier nicht auf der Hand. Auffällig ist jedoch, dass es sich bei den Abbildungen jeweils um unbedeckte Figuren handelte.

Im Rahmen der Gesamtklassifikation, bei der es galt sämtliche Verfahren zu kombinieren, die über eine prädiktive Valenz für die Indexgruppen verfügen (und damit sexuelle Präferenz erfassen sollten), zeigte sich jedoch, dass die physiologischen und neuropsychologischen Verfahren die impliziten Verfahren verdrängen. Die hypothesenkonformen vermuteten hirneigenen Entwicklungsdefizite sind das dominierende Muster für eine Klassifikation der Kontrollprobanden und der Indexgruppen. Zwar wird hierdurch eine sehr hohe Klassifikationsgenauigkeit erzielt ( $AUC = .86$ ), jedoch geht dies zulasten der Spezifität hinsichtlich der sexuellen Präferenz. Schliesst man hingegen die unspezifischen neuropsychologischen und neurophysiologischen Verfahren aus, dann endet man wieder bei einem einzelnen Verfah-

ren, dem IAT. Insgesamt ist also eine Kombination von Verfahren nicht besser zur Bestimmung der sexuellen Präferenz geeignet, als die jeweiligen Einzelverfahren.

Anders verhält es sich bei der Klassifikation der CSEM- und der CSA-Gruppe, im Hinblick auf eine Risikoprädiktion. Neben einem neuropsychologischen Verfahren, dessen Auswahl erneut die Bedeutung der differentiellen kognitiven Störungsmuster der Indexgruppen betont, fanden sich EDA-Muster für zwei spezifische Szenen, die einerseits eine gemischtgeschlechtliche Gruppe mit vier Kindern (Tannerstadium 1) enthielt und andererseits ein einzelnes Mädchen (Tannerstadium 1) abbildete. Alle Kinder waren in Badebekleidung und ohne Begleitung von Erwachsenen dargestellt. Die CSA-Probanden reagierten stärker auf die Gruppenszene während die CSEM-Probanden stärker auf das alleine dargestellte Mädchen ansprachen. Die Diskriminabilität war durch einen hohen Wert gekennzeichnet ( $AUC = .87$ ), der minimal die besten aktuarischen Prognoseinstrumente übertrifft (Eher et al., 2013). Die durch die Hautleitaktivität indizierten Szenerien könnten durchaus unspezifisch und lediglich selektiv für unsere Stichprobe sein. Sie könnten jedoch auch mit dem „Grooming“-Verhalten (Whittle, Hamilton-Giachritsis, Beech, & Collings, 2013) der CSA-Gruppe in Zusammenhang stehen und daher eine grundsätzliche Bedeutung haben. Als typische Kontaktorte für übergreifswillige Sexualstraftäter gelten u.a. Kaufhäuser, Schwimmbäder, Seen, Haltestellen, Schulen und Treffpunkte für Kinder (Kröber, 2013). All dies sind soziale Situationen, in denen eine Vielzahl von Kindern zugleich anwesend ist. Darüber hinaus handelt es sich bei der Gruppendarstellung um die einzige Szenerie, die potentielle Interaktionsmöglichkeiten anbietet (Spielzeug), somit eine Kontaktaufnahme unter natürlichen Rahmenbedingungen erleichtert. Hierin liegt ein Potential für die zukünftige Entwicklung ökologisch valider Szenarien in einer virtuellen Umgebung. Die Gruppensituation könnte daher tatsächlich die Induktion eines CSA-Gruppe-spezifischen Erregungsniveaus zufolge gehabt haben.

Dass die Darstellung von nackten oder leicht bekleideten Kindern für die CSEM-Gruppe eine grössere Bedeutung hat, als die Darstellung expliziter sexueller Handlungen, zeigen bereits Untersuchungen aus Österreich, in denen die Festplatten von  $N = 36$  Kinderpornografie-Konsumenten ausgewertet wurden (Fotr, 2013). Die Kategorien „Indikative Darstellungen“ (Copine 1), Posingbilder (Copine 4) erotisches Posing (Copine 5) und explizit erotisches Posing (Copine 6), waren auf den untersuchten Datenträgern weitaus am häufigsten vertreten. Die NRP-Abbildungen lassen sich ebenfalls dem Grad 1-3 der Copine-Skala (Taylor & Quayle, 2003) zuordnen, so dass auch der, durch die Hautleitaktivität indizierte Szenerie des alleine dargestellten Mädchens, u.U. eine gewisse Spezifität für die CSEM-Gruppe zukommen könnte.

### 4.3 Hypothesenprüfung

In unserem Forschungsplan vom 25. Februar 2010 hatten wir die folgenden Fragen und Hypothesen formuliert:

*Fragen:*

1. *Mit welcher Trennschärfe können in einem experimentellen Setting Probanden aus den drei Gruppen „Kontrolle“, „Konsumenten von Kinderpornographie“ und „Kindsmisssbraucher“ mit aktuellen Messmethoden unterschieden werden?*

Kontrollprobanden (CTL) können mit dem Gesamt-Klassifikationsmodell mit einer Trennschärfe von 0.86 AUC von den Straftätern mit pädosexueller Präferenz (CSO = CSEM und CSO) richtig klassifiziert werden und die gleiche Unterscheidung ermöglicht die Neuropsychologie mit 83% richtig klassifizierten Fällen in diesen beiden Gruppen. Ebenso unterschieden der IAT mit einer AUC von 0.74 bis 0.77 sowie die VT mit einer AUC von 0.77 diese beiden Gruppen recht gut. Die Trennschärfe hingegen zwischen den beiden Gruppen CSEM und CSA ist am besten in der VE mit EDA mit einem Wert von 0.87 AUC.

2. *Welche Methode zeigt allein die beste Trennschärfe?*

1. VE: CSEM vs. CSA, AUC = 0.87
2. Neuropsychologie: CTL vs. CSO, in 83 % korrekte Klassifikation
3. VT: CTL vs. CSO, AUC = 0.77
4. IAT: CTL vs. CSO, AUC = 0.74 bis 0.77

3. *Welche Kombination von Methoden zeigt die beste Trennschärfe?*

Das Gesamtklassifikationsmodell

*Hypothesen:*

1. *Hypothese 1 ( $H_1$ ): Kontrollpersonen können mit den beschriebenen Methoden von Konsumenten von Kinderpornographie und diese wiederum von Kindsmisssbrauchern unterschieden werden. Während bei Kindsmisssbrauchern und Konsumenten von Kinderpornographie im Gegensatz zu den Kontrollen eine pädosexuelle Präferenz nachweisbar ist, unterscheiden sich Kindsmisssbraucher durch mehr neurokognitive Defizite von den Konsumenten von Kinderpornographie. Damit sind alle drei Gruppen identifizierbar.*

2. Nullhypothese ( $H_0$ ): Die drei Gruppen können mit den beschriebenen Methoden nicht unterschieden werden, es besteht kein (gerichteter) Zusammenhang.

Abkürzungen:  $\mu$  = Mittelwerte der jeweiligen Messwerte für die entsprechende Gruppe

*adult* = adulte sexuelle Präferenz

*neurokog* = Neurokognitive Defizite

Die beiden Hypothesen werden wie folgt gegeneinander geprüft:

H1:  $\mu$  (CSO) CSA =  $\mu$  (CSO) CSEM >  $\mu$  (CSO) CTL

und

$\mu$  (adult) CSA =  $\mu$  (adult) CSEM <  $\mu$  (adult) CTL

und

$\mu$  (neurokog) CSA >  $\mu$  (neurokog) CSEM  $\approx$   $\mu$  (dis) CTL

Die Hypothese 1 kann wie folgt insgesamt mit einem bestimmten Vorbehalt validiert werden: Zwar lassen sich CTL-Probanden tatsächlich bezüglich pädosexueller Präferenz mit dem IAT und der VT von den CSO-Probanden unterscheiden (1. Zeile). Die CSEM-Gruppe ist aber bezüglich adulter Präferenz gemäss IAT und VT besser mit der CSA-Gruppe als mit der CTL-Gruppe vergleichbar (somit 2. Zeile bedingt valide) und die neurokognitiven Defizite verhalten sich gemäss Neuropsychologie ähnlich, d.h. Defizite in beiden Gruppen CSEM und CSA > CTL (Zeile 3).

H0: keine (gerichteten) Korrelationen wie in H1 beschrieben.

Es finden sich mehrere eindeutig gerichtete Korrelationen entsprechend Hypothese 1, die Nullhypothese kann also verworfen werden.

#### 4.4 Praktische Relevanz der Resultate

Die Frage, inwiefern die Ergebnisse der MIPS-Studie generalisierbar sind und sich auf alle Konsumenten von illegaler Internetpornographie oder pädophilen Kindsmisbrauchern übertragen lassen, hängt stark von der Repräsentativität der Stichprobe für die Gesamtpopulation ab. Im Wesentlichen lässt sich dabei die Frage der Repräsentativität auf die Frage nach Selektionseffekten reduzieren. Sofern es sich um zufällige Selektionseffekte handelt, spielt dabei auch die Stichprobengrösse eine entscheidende Rolle. Je grösser eine Stichprobe, desto kleiner ist der Einfluss von zufälligen Selektionseffekten. Bei einer Stichprobengrösse von

N = 20 pro Gruppe ist daher der Einfluss zufällig auftretender Selektionseffekte (z.B. überproportionaler Anteil von Cannabiskonsumenten, Berufsgruppen, älterer oder jüngerer Teilnehmer usw.) nicht ausgeschlossen. Systematische Selektionseffekte hängen von äusseren Faktoren, wie dem Interesse ab, welches z.B. die Kontrollprobanden der Studie entgegenbringen, der Zufriedenheit mit der Teilnahmeentschädigung und dem persönlichen Zeitbudget. Diese systematischen Selektionseffekte dürften aber auf Seiten der Sexualstraftäter deutlich schwerer wiegen als bei den Kontrollprobanden, zumal bei den Sexualstraftätern ein erster Selektionsprozess durch die Strafverfolgungs- und die Justizbehörden erfolgt ist, dann durch die jeweiligen Institutionen, die den Massnahmevollzug durchführen und schliesslich durch den Teilnehmer selbst, indem er sich freiwillig für die Teilnahme entscheidet.

Die Selektion durch die Strafverfolgungs- und Justizbehörden („Trichtermodell der Strafverfolgung“) könnte auch erklären, weshalb die Indexgruppen durch kognitive und andere Defizite gekennzeichnet sind, da es sich um eine Untergruppe von Sexualstraftätern handeln könnte, die aufgrund ihrer kognitiven Defizite besonders leicht für die Behörden identifizierbar ist und daher in den Massnahmezentren kumuliert. Studien mit behördlich unidentifizierten und vergleichbaren CSO-Probanden (Dunkelfeld) legen einen solchen Effekt nahe (Neutze, Grundmann, Scherner, & Beier, 2012). Darüber hinaus wird eine solche Hypothese auch durch die publizistische Berichterstattung über öffentlich bekannte Fälle nahe gelegt, die ganz offensichtlich auch einen Personenkreis betreffen, der nur mit geringer Wahrscheinlichkeit von erheblichen kognitiven und anderen psychiatrischen Defiziten betroffen ist (z.B. katholische Priester). Auf der anderen Seite jedoch erscheint das Defizitmuster insbesondere der CSA-Gruppe sehr kohärent, was man bei einer zufälligen Anhäufung eines allgemein defizitären Personenkreises nicht erwarten würde. Eine Alternativhypothese dazu ist, dass kognitive Defizite eines bestimmten Typus und als Folge einer hirnrorganischen Entwicklungsstörung, die Entwicklung einer pädophilen Neigung begünstigen und damit als Risikofaktor zu werten sind (Ward & Beech, 2006; Ward, Polaschek, & Beech, 2006).

Beide hier untersuchten Stichproben (CSA und CSEM) wiesen eine nicht unerhebliche Heterogenität auf, auch aufgrund der dargelegten Schwierigkeiten bei der Rekrutierung. So bestand eine deutliche Varianz im Bereich der sexuellen Orientierung (hetero- vs. homosexuell), der bestehenden Medikation (psychopharmakologische Medikation vs. keine Medikation), des Suchtmittelkonsums, der Verurteilung (bereits verurteilt vs. noch hängiges Verfahren) oder im Umfang und der Art der therapeutischen Vorbehandlung. Die Auswirkung dieser hier aufgelisteten Faktoren auf die gemessenen Ausprägungen in den einzelnen Testverfahren ist nicht genau abschätzbar und würde für jeden einzelnen dieser Faktoren weiterer Forschung bedürfen. Um einen beliebigen Einflussfaktor herauszugreifen: Es ist als wahrschein-

lich anzusehen, dass sich pädosexuelle Straftäter mit antidepressiver, medikamentöser Therapie in bestimmten Tests systematisch anders verhalten als pädosexuelle Straftäter ohne eine solche Therapie oder ohne die Notwendigkeit einer solchen Therapie.

Die meisten der eingesetzten Testverfahren erfassten sogenannte Traits (also zeitlich relativ überdauernde Eigenschaften), wie etwa die kognitive Leistungsfähigkeit. Es ist jedoch wichtig festzuhalten, dass sich States (also zeitlich eher variable Zustände) ebenfalls auf die Resultate auswirken können und so die Schätzung eines Traits verzerren können. So kann, um ein bekanntes Beispiel aufzugreifen, eine bestehende, aktuelle depressive Symptomatik (als State) zu einer Unterschätzung von kognitiven Leistungen führen. Die hier dargestellten Befunde betreffen zudem in erster Linie Gruppenunterschiede. Die Klassifikation von Individuen, wie weiter oben beschrieben (S.66 ff.), liefert lediglich probabilistische Aussagen über eine Gruppenzugehörigkeit aufgrund des Ausprägungsmusters in bestimmten Testverfahren. Es ist hierbei grundsätzlich zu bedenken, dass solche von Kontrollpersonen abweichende Ausprägungsmuster auch durch andere Faktoren als Pädosexualität zustande kommen können. Ein starker Hang zur Impulsivität, in einem neuropsychologischen Gamblingverfahren festgestellt, ist hier beispielsweise als ein wesentlicher Indikator für eine Klassifikation von pädosexuellen Straftätern gegenüber Kontrollen beschrieben (S. 66). Der Befund einer erhöhten Impulsivität kann allerdings auch völlig losgelöst von Pädosexualität auftreten, etwa bei Individuen mit Spielsucht. Selbst das Vorliegen eines für Pädosexualität typischen Profils (d.h. eines Ausprägungsmusters innerhalb *verschiedener* Testverfahren), bedarf der kritischen Hinterfragung, ob eine Kombination anderer Faktoren dieses Muster nicht ebenfalls erklären könnte.

Wie erwähnt, wurde das Forschungsvorhaben medial unter dem Schlagwort „Pädophilen-Detektor“ kolportiert. Dieses Schlagwort erachten wir als extrem unglücklich. Das Wort Detektor suggeriert, dass etwas mechanistisch und mit Bestimmtheit nachgewiesen werden kann, wie etwa bei einem Metalldetektor am Airport. Dem gegenüber lassen sich mit bestehenden Verfahren, lediglich Hinweise für das Vorliegen einer Pädosexualität erbringen. Eine wesentliche Motivation unserer Studie war, hierzu geeignete objektive Verfahren zu identifizieren. Indikatoren, die mittels solcher Verfahren erhoben werden, sollten aber als Ergänzung zu der fachlichen Beurteilung durch einen entsprechend ausgebildeten Psychiater oder Psychologen gesehen werden und keinesfalls als ein die fachliche Beurteilung ersetzendes Verfahren.

Aus den Resultaten der MIPS-Studie ergeben sich mehrere Konsequenzen für die Diagnostik und Behandlung von Sexualstraftätern der jeweiligen Subtypen. Leitend in der Beurteilung ist für den Kliniker zunächst die Frage nach der sexuellen Präferenz. Nach wie vor bereitet

diese erhebliche Schwierigkeiten, da sich der weitaus grösste Teil der Sexualität in neuronalen Erregungsmustern und hormonellen Veränderungen manifestiert, welche sich noch nicht einmal dem betroffenen Individuum vollständig durch Bewusstwerden erschliessen, geschweige denn von aussen auf der Verhaltensebene sichtbar werden. In einem strafrechtlichen Kontext kommen dann noch zusätzliche Verzerrungen in Form von gerichtetem Aussageverhalten hinzu. Aktuell begrenzt sich deshalb die Validität der diagnostischen Beurteilung der Störung der Sexualpräferenz auf einerseits die Angaben des betroffenen Patienten / Exploranden, die Tatanalyse sowie auf indirekte Rückschlüsse aus der Verhaltensbeobachtung.

Diesbezüglich ergeben sich aus der MIPS-Studie wie auch aus Resultaten von anderen Forschungsgruppen neue klinisch verwendbare Ansätze, wie die Blickdauer oder das IAT-Paradigma. Diese Verfahren sind derzeit für sich alleine genommen nicht geeignet, eine klare Zuordnung zu machen oder gar den Grad der pädophilen Präferenz quantitativ zu messen. Sie stellen aber schon jetzt ein Arbeitsinstrument dar, um im Rahmen einer Therapie mit den Befunden einen Einstieg in die oft unklare sexuelle Präferenz anbieten und in der gutachterlichen Beurteilung eine validere Prognose hinsichtlich Rückfälligkeit abgeben zu können. So haben zum Beispiel Banse et al. gute bis sehr gute Korrelationen zwischen dem EISIP mit aktuarischen Risikoinstrumenten wie Static-99 und SVR-20 aufzeigen können: Das EISIP (Explicit and Implicit Sexual Interest Profile) ist eine Kombination des EISIQ (Explicit and Implicit Sexual Interest Questionnaire) mit erstens Viewing Time und zweitens Implizitem Assoziationstest (Banse, Schmidt, & Clabour, 2010). Die Diskussionen anlässlich der IATSO (International Association for the Treatment of Sexual Offenders) Konferenz 2016 in Kopenhagen haben gezeigt, dass sich das EISIP als erster Standard in der wissenschaftlichen Untersuchung von fraglich pädophilen Sexualstraftätern entwickeln könnte mit einem guten Potential für die klinische Anwendung.

Forensische Klassifikationen unterscheiden meist „echte Pädophile“ („Kernpädophile“), die eine ausschliessliche, dominante sexuelle Präferenz für Kinder haben, von Personen mit einer „pädophilen Nebenströmung“ oder opportunistischen Tätern, welche kein oder nur ein sehr geringes sexuelles Interesse an Kindern haben und Delikte aus anderen Gründen begehen (etwa Gelegenheit, Intoxikation, finanzielles Interesse) (Berner, 2013).

In der vorliegenden Studie fallen die CSA-Probanden mit Hinweisen auf hirnorganische Defizite auf: Wie im klinischen Alltag finden sich auch bei den Indexgruppen, sowohl in der Anamnese, als auch in der neuropsychologischen Testung, Hinweise auf Hirnschädigungen. Während die erfragten Schädelhirntraumata im Gruppenvergleich keine Unterschiede zeigen

konnten, finden sich signifikante Testbefunde, die für einen impulsiven Tätertyp relevant sein dürften. Dieser Tätertyp zeichnet sich dadurch aus, dass nicht unbedingt eine stark ausgeprägte pädosexuelle Präferenz vorliegen muss, sondern die erhöhte Impulsivität und das Ausnutzen von „günstigen Gelegenheiten“ zur Delinquenz führen. Diese zeigen eher frontale Auffälligkeiten (Cohen & Galynker, 2009). Dies spricht deutlich in Richtung unserer im Forschungsplan formulierten und in wissenschaftlichen sowie klinischen Fachkreisen diskutierte Hypothese, dass Sexualdelinquenz eben nicht nur Ausdruck einer devianten Präferenz ist, sondern vielmehr das Resultat von Defiziten in verschiedenen Dimensionen. Diese manifestieren sich dann in unterschiedlichen, individuell stabilen Tatmustern, welche die schon seit langem bekannten Tätertypologien abbilden (Knight & Prentky, 1990). Unsere ersten Resultate lassen vermuten, dass sich mit den verwendeten Paradigmen die prognostisch so wichtige korrekte Zuordnung von Tätern zu den erwähnten Typologien verbessern lassen.

Berücksichtigt man z.B. das Einschlusskriterium der aktuellen „low-risk“-Legalprognose, so ist zu erwarten, dass bei diesem Tätertyp in einer gemischten Risikopopulation noch deutlichere Effekte in den neuropsychologischen Befunden erkennbar sein werden. Damit lassen sich hirngorganische Korrelate, die einen direkten Einfluss auf die Deliktdynamik haben, quantitativ bestimmen. Bei dieser Tätergruppe kann aus den Ergebnissen der Studie eine neuropsychologische Testung im Rahmen der Begutachtung empfohlen werden. Darüber hinaus müssen die kognitiven Defizite bzgl. der Art der zu ergreifenden psychotherapeutischen Massnahme Berücksichtigung finden, damit der erwünschte Lerneffekt in Richtung auf das Therapieziel auch sichergestellt werden kann. Auch hinsichtlich der Verlaufskontrollen im Rahmen einer Massnahme können neuropsychologische Befunde von Bedeutung sein. Bei diesem Tätertyp sollte sich idealerweise eine Verbesserung der Zielvariablen der Impulsivität und des Risikoverhaltens messen lassen. Standardisierte Messmethoden müssten aber noch festgelegt werden.

Da bis zum aktuellen Zeitpunkt in Anbetracht der Komplexität sowie der Menge der erhobenen Daten nur erste Auswertungen möglich waren, welche aber bereits jetzt hypothesenkonforme Resultate zeigen, erwarten wir von weiteren Analysen und dem Abgleich mit Resultaten von anderen Forschungsgruppen weitere aufschlussreiche Erkenntnisse. Damit können voraussichtlich die verwendeten Paradigmen technisch verbessert, in ihrer Anwendung auf portable Systeme reduziert und bezüglich ihrer klinisch-therapeutischen, als auch gutachterlichen Indikation und Validität verbessert werden.

Wir meinen dass es sich bei der vorliegenden Studie um ein gutes Beispiel translationaler Forschung handelt, wo auf Basis durch klinische Fragestellungen formulierter Hypothesen

Grundlagen und angewandte Forschung verschränkt werden und in beiden Bereichen neue Erkenntnisse resultieren. In der Grundlagenforschung z.B. die weitere Untersuchung neuronaler Prozesse bei der Verarbeitung visueller erotischer Stimuli und in der angewandten Forschung ganz wichtig die Verlaufsbeurteilung: Von immer grösserer forensisch-psychiatrischer klinischer aber auch juristischer und damit letztlich gesellschaftlicher Relevanz ist die Verlaufsbeurteilung von Sexualstraftätern unter Therapien in Hinblick auf Vollzugsentscheide wie Lockerungen oder Verlängerungen von Massnahmen respektive die Umwandlung in eine Verwahrung. Auch in Anbetracht der Heterogenität der Sexualstraftäter sind Längsschnittuntersuchungen aus methodischen Überlegungen mit intraindividuellen zeitlichen Vergleichen vielversprechend. Mögliche Anwendungsfelder erscheinen im Moment die Verlaufsuntersuchung von Sexualstraftätern vor antiandrogener medikamentöser Behandlung und während einer solchen, allenfalls im Vergleich zu psychotherapeutisch behandelten Tätern oder die Evaluation von unterschiedlichen Programmen zur Sexualstraftäterbehandlung zu sein. Unterschiedliche zeitliche Verlaufsmuster bei unterschiedlichen Subtypen von Sexualstraftätern sowie Studien zur differentiellen Wirksamkeit von Therapien (Lösel & Schmucker, 2005) legen nahe, dass sich eben auch zu Grunde liegende neuronale sowie hormonelle Muster intraindividuell zeitlich verändern und somit messbar sein sollten.

Schliesslich haben die international renommierten Sexualmediziner und Wissenschaftler Peer Briken, J. Paul Fedoroff und John W. Bradford mittlerweile einen wichtigen Kommentar zum neuen Diagnostischen Statistischen Manual DSM-5 der American Psychiatric Association publiziert: *Why can't pedophilic disorders remit?* (Briken, Fedoroff, & Bradford, 2014). Sie postulieren, dass mittlerweile genügend wissenschaftliche Evidenz besteht, dass paraphile Störungen wie die Pädophilie im Gegensatz zu den sexuellen Orientierungen (Hetero-, Homo- und Bisexualität), bei vielen Betroffenen nicht lebenszeitlich überdauernd sind. Sie fordern die Einführung der Spezifikation „in remission“, wie dies bei anderen paraphilen Störungen und bei Störungen aus dem schizophrenen Formenkreis oder affektiven Störungen schon lange etabliert ist. Dies zeigt einmal mehr die Notwendigkeit entsprechender Instrumente für die Messung des Verlaufs von paraphilen Präferenzen auf.

Auch anlässlich des Arbeitstreffens zur Weiterentwicklung der Mindestanforderungen für Prognosegutachten (Boetticher et al., 2007) am 17. Dezember 2016 in Heidelberg wurde auf die spezifische Notwendigkeit einer sexualmedizinischen Untersuchung von Sexualstraftätern hingewiesen und wird mit hoher Wahrscheinlichkeit in die Neuauflage der Mindestanforderungen eingehen.

Zusammengefasst heisst das für die Praxis, dass in absehbarer Zeit im westeuropäischen Raum und somit auch in der Schweiz Instrumente, welche neuropsychologische und neurophysiologische Verfahren kombinieren, als Ergänzung zur bisherigen etablierten aktuari-schen und klinischen Risiko- als auch Verlaufsbeurteilung bei Straftätern mit potentiell pädosexuellen Präferenzen sehr wahrscheinlich Eingang in die Praxis finden werden. Damit dies mit der gebotenen Sorgfalt und Fachlichkeit erfolgt, sind notwendig:

1. Die Schulung von Mitarbeiterinnen und Mitarbeitern, die solche Verfahren anwenden.
2. Eine fortlaufende wissenschaftliche Evaluation dieser zusätzlichen Verfahren.
3. Weitere wissenschaftliche Untersuchungen: Grundlagenforschung zum besseren Verständnis der einem sexuellen Missbrauch zu Grunde liegenden neuropsychologischen und neurophysiologischen Mechanismen wie auch klinische Forschung prognostischer und veränderungssensitiver Faktoren.

Dass in der Behandlung von Sexualstraftätern in der Schweiz noch erheblicher Verbesserungsbedarf besteht (nota bene auch in der Prävention mit wenigen löblichen Ausnahmen) ist so zu sagen eine Nebenerkenntnis unserer Studie, welche wir aber auch erwähnt haben möchten. Die Studie wäre beinahe am Problem der Rekrutierung von Probanden gescheitert. In der Vorbereitung der Studie wurde selbstverständlich in vielen forensisch-psychiatrischen Diensten die Verfügbarkeit von Probanden vorsondiert und von den Verantwortlichen bestätigt. In der Realität zeigte sich dann aber rasch, dass viele solche Therapieangebote lediglich beabsichtigt waren durchgeführt zu werden, die notwendigen Therapeuten nicht vorhanden waren oder man im Moment einfach keine Patienten hatte. Dies deutet auf eine klare Unter-versorgung bezüglich Prävention und Behandlung von Sexualstraftätern in der Deutschschweiz hin. Ausnahmen betreffen sowohl gut strukturierte Angebot in Institutionen als auch in Praxen niedergelassener Psychiater und Psychologen. Eine diesbezügliche Planung und Umsetzung in den Konkordaten erscheint in aller gebotenen Zurückhaltung notwendig.

## Literatur

- Adjorlolo, S., & Egbenya, D. L. (2016). Executive functioning profiles of adult and juvenile male sexual offenders: A systematic review. *The Journal of Forensic Psychiatry & Psychology*, 1–27.
- Akerman, G., & Beech, A. R. (2012). A Systematic Review of Measures of Deviant Sexual Interest and Arousal. *Psychiatry, Psychology and Law*, 19(1), 118–143. <https://doi.org/10.1080/13218719.2010.547161>
- Anderson, S. W., Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, 2(11), 1032–1037. <https://doi.org/10.1038/14833>
- Babchishin, K. M., Hanson, K. R., & Hermann, C. A. (2011). The Characteristics of Online Sex Offenders: A Meta-Analysis. *Sexual Abuse: A Journal of Research and Treatment*, 23(1), 92–123. <https://doi.org/10.1177/1079063210370708>
- Babchishin, K. M., Hanson, R. K., & VanZuylen, H. (2015). Online Child Pornography Offenders are Different: A Meta-analysis of the Characteristics of Online and Offline Sex Offenders Against Children. *Archives of Sexual Behavior*, 44(1), 45–66. <https://doi.org/10.1007/s10508-014-0270-x>
- Babchishin, K. M., Nunes, K. L., & Hermann, C. A. (2012). The Validity of Implicit Association Test (IAT) Measures of Sexual Attraction to Children: A Meta-Analysis. *Archives of Sexual Behavior*, 42(3), 487–499. <https://doi.org/10.1007/s10508-012-0022-8>
- Banse, R., Schmidt, A. F., & Clabour, J. (2010a). Indirect Measures of Sexual Interest in Child Sex Offenders A Multimethod Approach. *Criminal Justice and Behavior*, 37(3), 319–335. <https://doi.org/10.1177/0093854809357598>
- Banse, R., Schmidt, A. F., & Clabour, J. (2010b). Indirect Measures of Sexual Interest in Child Sex Offenders A Multimethod Approach. *Criminal justice and Behavior*, 37(3), 319–335.
- Bäumler, G. (1984). *Farbe-Wort-Interferenztest (FWIT) nach JR Stroop: Handanweisung*. Verlag für Psychologie Hogrefe.
- Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50(1), 7–15.
- Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *The Journal of general psychology*, 39(1), 15–22.
- Berner, W. (2013). Sexueller Missbrauch – Epidemiologie und Phänomenologie. In H. Schanda, T. Stompe, & W. Laubichler (Hrsg.), *Sexueller Kindesmissbrauch und Pädophilie* (S. 1–14). Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft (MWV).
- Blanchard, R., Christensen, B. K., Strong, S. M., Cantor, J. M., Kuban, M. E., Klassen, P., ... Blak, T. (2002). Retrospective Self-Reports of Childhood Accidents Causing Unconsciousness in Phal-

- lometrically Diagnosed Pedophiles. *Archives of Sexual Behavior*, 31(6), 511–526. <https://doi.org/10.1023/A:1020659331965>
- Blanchard, R., Kuban, M. E., Klassen, P., Dickey, R., Christensen, B. K., Cantor, J. M., & Blak, T. (2003). Self-Reported Head Injuries Before and After Age 13 in Pedophilic and Nonpedophilic Men Referred for Clinical Assessment. *Archives of Sexual Behavior*, 32(6), 573–581. <https://doi.org/10.1023/A:1026093612434>
- Boetticher, A., Krober, H., Muller-Isberner, R., Bohm, K. M., Muller-Metz, R., & Wolf, T. (2007). Mindestanforderungen für Prognosegutachten. *SEXUOLOGIE-STUTTGART*-, 14(1/2), 36.
- Breiman, L. (2001). Random Forests. *Mach. Learn.*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. New York: Chapman & Hall/CRC press.
- Briken, P., Fedoroff, J. P., & Bradford, J. W. (2014). Why Can't Pedophilic Disorder Remit? *Archives of Sexual Behavior*, 43(7), 1237–1239. <https://doi.org/10.1007/s10508-014-0323-1>
- Cantor, J. M., Blanchard, R., Christensen, B. K., Dickey, R., Klassen, P. E., Beckstead, A. L., ... Kuban, M. E. (2004). Intelligence, Memory, and Handedness in Pedophilia. *Neuropsychology*, 18(1), 3–14. <https://doi.org/10.1037/0894-4105.18.1.3>
- Cantor, J. M., Kabani, N., Christensen, B. K., Zipursky, R. B., Barbaree, H. E., Dickey, R., ... Blanchard, R. (2008). Cerebral white matter deficiencies in pedophilic men. *Journal of Psychiatric Research*, 42(3), 167–183. <https://doi.org/10.1016/j.jpsychires.2007.10.013>
- Clark, L., Cools, R., & Robbins, T. W. (2004). The neuropsychology of ventral prefrontal cortex: Decision-making and reversal learning. *Brain and Cognition*, 55(1), 41–53. [https://doi.org/10.1016/S0278-2626\(03\)00284-7](https://doi.org/10.1016/S0278-2626(03)00284-7)
- Cohen, L. J., & Galynker, I. (2009). Psychopathology and personality traits of pedophiles: Issues for diagnosis and treatment. *Psychiatric Times*, 26(6), 25–25.
- Cohen, L., McGeoch, P., Watras-Gans, S., Acker, S., Poznansky, M. S. W., Cullen, K., & Itskovich, Y. (2002). Personality impairment in male pedophiles. *The Journal of Clinical Psychiatry*, 63(10), 912–919.
- Covell, C. N., & Scalora, M. J. (2002). Empathic deficits in sexual offenders: An integration of affective, social, and cognitive constructs. *Aggression and Violent Behavior*, 7(3), 251–270. [https://doi.org/10.1016/S1359-1789\(01\)00046-5](https://doi.org/10.1016/S1359-1789(01)00046-5)
- Danek, A., & Göhringer, T. (2005). Kognitive Neurologie und Neuropsychologie. In P. D. H. Förstl (Hrsg.), *Frontalhirn* (S. 41–82). Springer Berlin Heidelberg. Abgerufen von [http://link.springer.com/chapter/10.1007/3-540-26841-3\\_3](http://link.springer.com/chapter/10.1007/3-540-26841-3_3)
- Deegener, G. (1996). *Multiphasic sex inventory (MSI): Fragebogen zur Erfassung psychosexueller Merkmale bei Sexualtätern: Handbuch*. Hogrefe, Verlag für Psychologie.

- Eher, R., Rettenberger, M., Gaunersdorfer, K., Haubner-MacLean, T., Matthes, A., Schilling, F., & Mokros, A. (2013). Über die Treffsicherheit der standardisierten Risikoeinschätzungsverfahren Static-99 und Stable-2007 bei aus einer Sicherungsmaßnahme entlassenen Sexualstraftätern. *Forensische Psychiatrie, Psychologie, Kriminologie*, 7(4), 264–272. <https://doi.org/10.1007/s11757-013-0212-9>
- Fagan, P., Wise, T., Schmidt, & Berlin, F. (2002). Pedophilia. *JAMA*, 288(19), 2458–2465. <https://doi.org/10.1001/jama.288.19.2458>
- Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *Neuroimage*, 26(2), 471–479.
- Featherman, D. L., Sobel, M. E., & Dickens, D. (1975). *A manual for coding occupations and industries into detailed 1970 categories and a listing of 1970-basis Duncan socioeconomic and NORC prestige scores*. Center for Demography and Ecology, University of Wisconsin-Madison. Abgerufen von <http://ssc.wisc.edu/cde/cdewp/75-1.pdf>
- Finkelhor, D., & Araji, S. (1986). Explanations of Pedophilia: A Four Factor Model. *The Journal of Sex Research*, 22(2), 145–161.
- Fisher, C., Cross, J., & Zuch, J. (1965). Cycle of penile erection synchronous with dreaming (REM) sleep: Preliminary report. *Archives of General Psychiatry*, 12(1), 29–45. <https://doi.org/10.1001/archpsyc.1965.01720310031005>
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101(2), 171–191. <https://doi.org/10.1037/0033-2909.101.2.171>
- Fotr, F. (2013). Computerforensische Untersuchungen im Umfeld von Kindesmissbrauchsfällen. In H. Schanda, T. Stompe, & W. Laubichler (Hrsg.), *Sexueller Kindesmissbrauch und Pädophilie* (S. 319–326). Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft (MWV).
- Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14), 2225–2236. <https://doi.org/10.1016/j.patrec.2010.03.014>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the Implicit Association Test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480.
- Hanson, R. K., & Morton-Bourgon, K. E. (2005). The Characteristics of Persistent Sexual Offenders: A Meta-Analysis of Recidivism Studies. *Journal of Consulting*, 73(6), 1154–1163.
- Harris, A., Phenix, A., Thornton, D., & Hanson, R. K. (2003). *Static 99: Coding rules revised 2003*. Solicitor General Canada Ottawa, ON. Abgerufen von [http://www.defenseforsvp.com/Resources/Hanson\\_Static-99/Static99codingRules\\_e.pdf](http://www.defenseforsvp.com/Resources/Hanson_Static-99/Static99codingRules_e.pdf)
- Holmes, T. H., & Rahe, R. H. (1967). The social readjustment rating scale. *Journal of psychosomatic research*, 11(2), 213–218.

- Imhoff, R., Schmidt, A. F., Bernhardt, J., Dierksmeier, A., & Banse, R. (2011). An inkblot for sexual preference: A semantic variant of the Affect Misattribution Procedure. *Cognition & Emotion*, 25(4), 676–690. <https://doi.org/10.1080/02699931.2010.508260>
- Jespersen, A. F., Lalumière, M. L., & Seto, M. C. (2009). Sexual abuse history among adult sex offenders and non-sex offenders: A meta-analysis. *Child Abuse & Neglect*, 33(3), 179–192. <https://doi.org/10.1016/j.chiabu.2008.07.004>
- Joyal, C. C., Black, D. N., & Dassylva, B. (2007). The Neuropsychology and Neurology of Sexual Deviance: A Review and Pilot Study. *Sexual Abuse: A Journal of Research and Treatment*, 19(2), 155–173. <https://doi.org/10.1007/s11194-007-9045-4>
- Kinsey, A. C., von Eckardt-Jaffé, M., von Eckardt-Jaffé, M., Martin, C. E., & Pomeroy, W. B. (1966). *Das sexuelle Verhalten des Mannes*. S. Fischer.
- Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of experimental psychology*, 55(4), 352.
- Knight, R. A., & Prentky, R. A. (1990). Classifying sexual offenders: The Development and Corroboration of Taxonomic Models, Chapter 3. In Marshall, Laws, & Barbaree (Hrsg.), *Handbook of Sexual Assault: Issues, Theories, and Treatment of the Offender* (S. 23–52). New York: Springer.
- Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology*, 17(2), 248–294. [https://doi.org/10.1016/0010-0285\(85\)90009-X](https://doi.org/10.1016/0010-0285(85)90009-X)
- Kröber, Hans-L. (2013). Sexuelle Übergriffe auf Kinder ausserhalb der Familie – auch in Schulen, Heimen, Vereinen und Kirche. In W. Laubichler, H. Schanda, & T. Stompe (Hrsg.), *Sexueller Kindesmissbrauch und Pädophilie* (Erste Auflage, S. 87–104). MWV Medizinisch Wissenschaftliche Verlagsgesellschaft.
- Kruger, T. H. C., & Schiffer, B. (2011a). Neurocognitive and Personality Factors in Homo- and Heterosexual Pedophiles and Controls. *The Journal of Sexual Medicine*, 8(6), 1650–1659. <https://doi.org/10.1111/j.1743-6109.2009.01564.x>
- Kruger, T. H. C., & Schiffer, B. (2011b). Neurocognitive and Personality Factors in Homo- and Heterosexual Pedophiles and Controls: Personality and Neurocognition in Pedophiles. *The Journal of Sexual Medicine*, 8(6), 1650–1659. <https://doi.org/10.1111/j.1743-6109.2009.01564.x>
- Langevin, R., Paitich, D., Freeman, R., Mann, K., & Handy, L. (1978). Personality characteristics and sexual anomalies in males. *Journal of Behavioural Science*, 10(3), 222–238.
- Larson, M. J., & Perlstein, W. M. (2009). Awareness of deficits and error processing after traumatic brain injury: *NeuroReport*, 20(16), 1486–1490. <https://doi.org/10.1097/WNR.0b013e32833283fe>
- Linnoila, M., Virkkunen, M., Scheinin, M., Nuutila, A., Rimon, R., & Goodwin, F. K. (1983). Low cerebrospinal fluid 5-hydroxyindoleacetic acid concentration differentiates impulsive from nonim-

- pulsive violent behavior. *Life Sciences*, 33(26), 2609–2614. [https://doi.org/10.1016/0024-3205\(83\)90344-2](https://doi.org/10.1016/0024-3205(83)90344-2)
- Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology*, 1(1), 117–146. <https://doi.org/10.1007/s11292-004-6466-7>
- Lykken, D. T. (1984). Polygraphic interrogation. *Nature*, 307, 681–684.
- Merdian, H. L., Wilson, N., & Boer, D. P. (2009). Characteristics of internet sexual offenders: a review. *Sexual Abuse in Australia and New Zealand*, 2(1), 34–47.
- Millon, T., Simonsen, E., Birket-Smith, M., & Davis, R. D. (2003). *Psychopathy: Antisocial, Criminal, and Violent Behavior*. Guilford Press.
- Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British medical bulletin*. Abgerufen von <http://psycnet.apa.org/psycinfo/1972-32010-001>
- Mokros, A., Gebhard, M., Heinz, V., Marschall, R. W., Nitschke, J., Glasgow, D. V., ... Laws, D. R. (2013). Computerized Assessment of Pedophilic Sexual Interest Through Self-Report and Viewing Time: Reliability, Validity, and Classification Accuracy of the Affinity Program. *Sexual Abuse: A Journal of Research and Treatment*, 25(3), 230–258. <https://doi.org/10.1177/1079063212454550>
- Mokros, Osterheider, M., & Nitschke, J. (2012). Pädophilie. *Der Nervenarzt*, 83(3), 355–358.
- Neutze, J., Grundmann, D., Scherner, G., & Beier, K. M. (2012). Undetected and detected child sexual abuse and child pornography offenders. *International Journal of Law and Psychiatry*, 35(3), 168–175. <https://doi.org/10.1016/j.ijlp.2012.02.004>
- Niemann, H., Sturm, W., Thöne-Otto, A., & Willmes-von-Hinkeldey, K. (2008). *California Verbal Learning Test-German adaptation*. Pearson As. Frankfurt.
- Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an anti-saccade task. *Psychophysiology*, 38(5), 752–760. <https://doi.org/10.1111/1469-8986.3850752>
- Oswald, W. D., & Fleischmann, U. M. (1997). *Nürnberger Alters-Inventar (NAI)*(4. Aufl.) Göttingen: Hogrefe.
- Pacific Behavioral Assessment Corporation. (2004). The NRP (Not Real People) Visual Stimulus Set.
- Partington, J. E., & Leiter, R. G. (1949). Halstead-Reitan Trail Making Test. *Arizona: Reitan Neuropsychology Laboratory*.
- Pereda, N., Guilera, G., Forns, M., & Gómez-Benito, J. (2009). The prevalence of child sexual abuse in community and student samples: A meta-analysis. *Clinical Psychology Review*, 29(4), 328–338. <https://doi.org/10.1016/j.cpr.2009.02.007>
- Poeppel, T. B., Nitschke, J., Santtila, P., Schecklmann, M., Langguth, B., Greenlee, M. W., ... Mokros, A. (2013). Association between brain structure and phenotypic characteristics in pedophilia.

- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>
- Raine A., Lencz T., Bihrlé S., LaCasse L., & Colletti P. (2000). Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. *Archives of General Psychiatry*, 57(2), 119–127. <https://doi.org/10.1001/archpsyc.57.2.119>
- Rogers, R. D., Everitt, B. J., Baldacchino, A., Blackshaw, A. J., Swainson, R., Wynne, K., ... others. (1999). Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. *Neuropsychopharmacology*, 20(4), 322–339.
- Rosenzweig, S. (1942). The Photoscope as an Objective Device for Evaluating Sexual Interest. *Psychosomatic Medicine*, 4(2), 150–158.
- Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R., Grant, M., & others. (1993). Development of the alcohol use disorders identification test (AUDIT). WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *ADDICTION-ABINGDON*, 88, 791–791.
- Saver, J. L., & Damasio, A. R. (1991). Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia*, 29(12), 1241–1249. [https://doi.org/10.1016/0028-3932\(91\)90037-9](https://doi.org/10.1016/0028-3932(91)90037-9)
- Schiffer, B., Paul, T., Gizewski, E., Forsting, M., Leygraf, N., Schedlowski, M., & Kruger, T. H. C. (2008). Functional brain correlates of heterosexual paedophilia. *NeuroImage*, 41(1), 80–91. <https://doi.org/10.1016/j.neuroimage.2008.02.008>
- Schiffer, B., & Vonlaufen, C. (2011). Executive Dysfunctions in Pedophilic and Nonpedophilic Child Molesters. *The Journal of Sexual Medicine*, 8(7), 1975–1984. <https://doi.org/10.1111/j.1743-6109.2010.02140.x>
- Schiltz K, Witzel J, Northoff G, & et al. (2007). Brain pathology in pedophilic offenders: Evidence of volume reduction in the right amygdala and related diencephalic structures. *Archives of General Psychiatry*, 64(6), 737–746. <https://doi.org/10.1001/archpsyc.64.6.737>
- Seto, M. C. (2008). *Pedophilia and sexual offending against children: Theory, assessment, and intervention* (Bd. xvi). Washington, DC, US: American Psychological Association.
- Seto, M. C., & Lalumière, M. L. (2001). A Brief Screening Scale to Identify Pedophilic Interests Among Child Molesters. *Sexual Abuse: A Journal of Research and Treatment*, 13(1), 15–25. <https://doi.org/10.1023/A:1009510328588>
- Seto, M. C., Stephens, S., Lalumière, M. L., & Cantor, J. M. (2015). The Revised Screening Scale for Pedophilic Interests (SSPI–2) Development and Criterion-Related Validation. *Sexual Abuse: A*

- Taylor, M., & Quayle, E. (2003). *Child Pornography: An Internet Crime*. Psychology Press.
- Volavka, J. (2008). *Neurobiology of Violence, Second Edition*. American Psychiatric Pub.
- Ward, T., & Beech, A. (2006a). An integrated theory of sexual offending. *Aggression and Violent Behavior*, 11(1), 44–63. <https://doi.org/10.1016/j.avb.2005.05.002>
- Ward, T., & Beech, A. (2006b). An integrated theory of sexual offending. *Aggression and Violent Behavior*, 11(1), 44–63. <https://doi.org/10.1016/j.avb.2005.05.002>
- Ward, T., Polaschek, D., & Beech, A. R. (2006). *Theories of sexual offending*. John Wiley & Sons.
- Wechsler, D. (1981). *WAIS-R manual: Wechsler adult intelligence scale-revised*. Psychological Corporation.
- Whittle, H., Hamilton-Giachritsis, C., Beech, A., & Collings, G. (2013). A review of online grooming: Characteristics and concerns. *Aggression and Violent Behavior*, 18(1), 62–70. <https://doi.org/10.1016/j.avb.2012.09.003>

# Anhang

## Abkürzungsverzeichnis

$\mu\text{V}$		$\mu\text{Volt}$
Abb.		Abbildung
Art.		Artikel
AUC	Area under the Curve	
CRT	Choice Reaction Time	
CSA	Child Sexual Abuse	Gruppe mit bekanntem Kindesmissbrauch
CSEM	Child Sexual Exploitation	Gruppe mit bekanntem Kinderinternetpornographiekonsum
CSO	Child Sexual Offenders	CSA und CSEM zusammengefasst
CTL	Controls	Kontrollgruppe
d.h.		das heisst
df		Freiheitsgrade
DSM	Diagnostic and Statistical Manual of Mental Disorders	diagnostischer und statistischer Leitfaden psychischer Störungen
e	error	
EDA		Elektrodermale Aktivität
EEG		Elektroenzephalogramm
EKG		Elektrokardiogramm
EKP		Ereigniskorreliertes Potential
IAT	Implicit Association Test	Impliziter Assoziationstest
ICD	International Statistical Classification of Diseases and Related Health Problems	Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme
IQ		Intelligenzquotient
M		männlich
m		Mittelwert
md		Median
MIPS		Messbare Indikatoren Pädosexueller Straftäter
ms	milliseconds	Millisekunden
NRP	Not Real People	

$p$	Probability	Wahrscheinlichkeit
prox.	proximity	Proximität/Nähe
$r$		Regressionskoeffizient
RLC	Reversal Learning Capacity	
Rnd.	Round	Durchgang
ROC	Receiver Operating Characteristic	
RT	Reaction Time	Reaktionszeit
SMP	Semantic Misattribution Paradigma	Semantisches Paradigma Missattributions-Paradigma
SOA	Stimulus Onset-Asynchrony	Zeitlicher Abstand zwischen dem Onset zweier sukzessiver Reize
SSPI	Screening Scale for Pedophilic Interest	
StGB		Strafgesetzbuch
T1-5		Tannerstadium 1-5
VE		Virtuelle Exposition
Vgl.		vergleiche
vIQ		Verbaler Intelligenzquotient
Vs.	versus	versus
VT	Viewing Time	Betrachtungszeit
W		weiblich

## Glossar

---

95%-Konfidenzintervall	Das 95%-Konfidenzintervall ist derjenige Bereich, in dem der wahre Messwert mit einer Wahrscheinlichkeit von 95% liegt.
Amplitude	Stärke eines ereigniskorrelierten Potentials.
antiandrogen	Wirkung der männlichen Sexualhormone (Androgene) hemmende Medikation.
Compliance	Oberbegriff für das Mass an kooperativem Verhalten eines Patienten/Probanden.
Diskriminabilität	Unterscheidbarkeit.
Elektrodermale Aktivität (EDA)	Aktive und passive elektrische Eigenschaften, welche sich auf Strukturen und Funktionen der Haut und der unter ihr liegenden Organe zurückführen lassen. Die Registrierung der EDA gilt als eine der am häufigsten verwendeten Methoden zur Erfassung von Emotionen und Stress.
Ereigniskorrelierte Potentiale (EKP)	Auf einem bestimmten Stimulus oder auf ein bestimmtes Verhalten hin gemittelte EEG-Aktivität. Durch die Mittelung werden unsystematische Anteile im EEG unterdrückt und systematische Anteile hervorgehoben.
explizit/implizit	Bewusst/unbewusst (im Detail unter „Implizite Verfahren“ beschrieben)
Dunkelfeld/Dunkelziffer	Unaufgeklärter Bereich einer Statistik. In der Kriminalistik die Differenz zwischen dem tatsächlich begangenen und den Strafverfolgungsbehörden bekannt gewordenen Straftaten und den vermutlichen begangenen Straftaten. Diese Delikte werden u. a. deshalb nicht bekannt, da sie von den Opfern nicht angezeigt werden.
„Grooming“-Verhalten	Prozess der Vorbereitung und Annäherung mit dem Ziel des Kindesmissbrauchs.

---

Habituation	Eine basale Form des Lernens, die gekennzeichnet ist durch eine Abschwächung der Reaktion auf wiederholte Stimulation.
Indikator	Merkmal, welches herangezogen wird, um eine probabilistische Aussage über die Ausprägung oder das Vorliegen einer Eigenschaft machen.
Klassifikationsalgorithmus	s. Random Forest-Analyse
Kognitive Verzerrung	Irrationale Gedankenmuster; in der klinischen Psychologie ein die klinische Störung aufrechterhaltender Faktor, der Verhaltenseinsicht behindert.
Korrelation	Stärke des Zusammenhangs zweier Variablen. Positive Korrelation: Je mehr A, desto mehr B (und umgekehrt). Negative Korrelation: Je mehr A, desto weniger B (und umgekehrt). Eine Korrelation beschreibt dabei jedoch lediglich einen statistischen Zusammenhang, keine Kausalität.
Kruskal-Wallis-Test	Ein statistisches Testverfahren, mit welchem sich mehrere Rangreihen miteinander vergleichen und auf signifikante Unterschiede prüfen lassen.
Latenz	Zeitpunkt in einem ereigniskorrelierten Potential.
Legalprognose	Risikobeurteilung einer straffälligen Person bezüglich Fähigkeit und Motivation zu einem späteren Zeitpunkt Gesetze einzuhalten.
Median	Bezeichnet den mittleren Wert einer Zahlenreihe und ist somit geeignet für Daten mit Rangreihenniveau und robust gegenüber statistischen Ausreißern.
Meta-Analyse	Statistisches Verfahren, um die Resultate aus verschiedenen, aber vergleichbaren Studien zu vereinen.
Minimum und Maximum	Zeigt den tiefsten bzw. höchsten erreichten Wert einer Person in der entsprechenden Gruppe an.

Operationalisierung	Konkretisierung theoretischer Begriffe und Hypothesen durch Angabe beobachtbarer und messbarer Ereignisse.
$p$ -Wert	s. Signifikanz.
Paraphilie	Oder Störung der Sexualpräferenz (ICD-10). Von empirischer Norm abweichendes sexuelles Verhalten, deviante Fantasien und dranghafte Bedürfnisse, welche überwiegend oder ausschließlich Inhalte des sexuellen Interesses sind (Fetischismus, Voyeurismus, <b>Pädophilie</b> , Sadomasochismus u. a.). Therapie indiziert bei klinisch bedeutsamen Leiden. Abzugrenzen von Sexualdelinquenz.
Phallometrie	Verfahren zur Messung der Penisreaktion bei sexualpsychologischen Untersuchungen.
Präfrontal	Lagebezeichnung im Gehirn aus der Neuroanatomie. Bezeichnet Areale, die im vorderen Teil des frontalen Kortex liegen.
Prävalenz	Rate der Erkrankten an einer bestimmten Krankheit zu einem bestimmten Zeitpunkt oder in einem bestimmten Zeitabschnitt (im Vergleich zur Zahl der Untersuchten).
Random Forest-Analyse	Klassifikationsmethode, welche aus unkorrelierten und randomisierten Entscheidungsbäumen besteht, welcher jeder eine Entscheidung zu einer Klassifikation trifft. Die endgültige Klassifikation wird durch die Entscheidung mit den meisten Stimmen bestimmt.
Randomisierung	Zufällige Gruppenzuteilung oder zufällige Reihenfolgen in der Stimulation; durch Randomisierung werden Reihenfolgeeffekte oder andere mögliche systematische Effekte eliminiert.
Regressionsanalyse	Statistisches Verfahren zur Analyse einer funktionalen Abhängigkeit einer Kriteriumsvariable

	von einer oder mehreren Prädiktorvariablen.
Residuum	Abweichung eines Wertes zur Regressionsgeraden nach Regression.
ROC-Analyse	<p>Graphische Darstellung der diagnostischen Sensitivität über der diagnostischen Spezifität, beides in Abhängigkeit eines repräsentativen cut-off-Punktes.</p> <p>Der cut-off-Punkt repräsentiert die Kriteriumsgrenze, z.B. zwischen „Gesund“ und „Krank“.</p>
Sensitivität	Fähigkeit eines Tests, bestehende Merkmale (z.B. Krankheit) zu erfassen. Angabe zur Wahrscheinlichkeit für einen positiven Befund bei Vorhandensein des Merkmals (vgl. Spezifität).
Sexuelle Orientierungsstörung	Störungen der Sexualpräferenz (s.a. Paraphilie).
Signifikanz	Ein gemessener Zusammenhang gilt als statistisch signifikant, wenn die Wahrscheinlichkeit, dass dieser zufällig zustande gekommen ist, unterhalb einer vordefinierten Irrtumswahrscheinlichkeit liegt. In der Forschung wird dieser Wert im Allgemeinen bei 5% ( $p < .05^*$ ) angesiedelt.
Spezifität	Eignung eines diagnostischen Tests, das fehlende Vorhandensein eines Merkmals (z.B. keine Krankheit) mit einem negativen Testergebnis zu beurteilen (vgl. Sensitivität).
supraliminal/subliminal	Für das Bewusstsein überschwellige/unterschwellige Darbietung von Reizen.

## Verwendete Tests

Alcohol Use Disorders Identification Test (**AUDIT**): Selbstbeurteilungsfragebogen zur Erfassung problematischen Alkoholkonsums. (Saunders et al., 1993)

Attention Network Test (**ANT**): Computergestützter Test zur Erfassung der Leistung in verschiedenen Aufmerksamkeitsdimensionen. (Fan, McCandliss, Fossella, Flombaum, & Posner, 2005)

California Verbal Learning Test (**CVLT-II**): Test zur Überprüfung der episodischen Gedächtnisleistung. (Niemann, Sturm, Thöne-Otto, & Willmes-von-Hinkeldey, 2008)

Cambridge Gambling Task (**CGT**): Test zur Erfassung von Risikoverhalten und Impulsivität. (Rogers et al., 1999)

Childhood Trauma Questionnaire (**CTQ**): Selbstbeurteilungsfragebogen zur retrospektiven Erfassung von erlebtem Missbrauch und erfahrener Vernachlässigung in der Kindheit. (Bernstein & Fink, 1998)

Computer-Maus Übung: Übung zum Umgang mit der Computermaus. (Eigenentwicklung)

Corsi Block Tapping Test (**CBTT**): Computergestützter Test zur Erfassung der räumlichen Merkfähigkeit des Arbeitsgedächtnisses passiv (Corsi vorwärts) und aktiv (Corsi rückwärts). (Milner, 1971)

Edinburgh Handedness Inventory (**EHI**): Selbstbeurteilungsfragebogen zur Erfassung der Handpräferenz. (Oldfield, 1971)

Gemeinsamkeiten Finden: Test zur Überprüfung der Abstraktionsfähigkeit. (Wechsler, 1981)

Inventars der Persönlichkeitsorganisation, Kurzversion (**IPO-16**): Selbstbeurteilungsfragebogen zur Messung des Schweregrads der strukturellen Beeinträchtigung in den Bereichen Identität, Abwehr und Realitätsprüfung. (Zimmermann et al., 2015)

Iowa Gambling Task (**IGT**): Test zur Untersuchung von probabilistischem Lernen, wobei der Vorteil einer langfristigen, im Mittel niedrigeren Belohnung zugunsten einer kurzfristig hohen Belohnung erkannt und angewendet werden muss. (Bechara, Damasio, Damasio, & Anderson, 1994)

Kinsey-Skala: Angabe der sexuellen Ausrichtung auf einer siebenstufigen Skala (homo, bi-, heterosexuell). (Kinsey, von Eckardt-Jaffé, von Eckardt-Jaffé, Martin, & Pomeroy, 1966)

Leistungsprüfsystem, Untertest 3 (**LPS 3**): Untertest zur Messung der nonverbalen, räumlich-figuralen Intelligenz durch das Erkennen von (Un-)regelmäßigkeiten in geometrischen Figuren. (Horn, 1983)

Manual der Arbeitsgemeinschaft für Methodik und Dokumentation in der Psychiatrie (**AMDP-System**): Manual zur Dokumentation psychiatrischer Befunde und anamnestischer Daten. (Fähndrich & Stieglitz, 2006)

Mehrfachwahl-Wortschatz-Test (**MWT-B**): Test zur Messung verbaler Intelligenz durch das Erkennen eines bekannten Wortes innerhalb einer Reihe von fiktiven Begriffen. (Lehrl, 1977)

Multiphasic Sex Inventory (**MSI**): Selbstbeurteilungsfragebogen zur Erfassung psychosexueller Merkmale und Verhaltensauffälligkeiten bei Sexualstraftätern. (Deegener, 1996)

**N-Back** Test: Computergestützter 2-Back-Test zur Erfassung der aktiven Komponente des Arbeitsgedächtnisses. (Kirchner, 1958)

NEO-Persönlichkeitsinventar-Revidiert (**NEO-PI-R**): Selbstbeurteilungsfragebogen zur Erfassung der Ausprägungen auf den fünf Persönlichkeitsfaktoren Offenheit, Gewissenhaftigkeit, Verträglichkeit, Extraversion und Neurotizismus. (Ostendorf & Angleitner, 2004)

Psychopathy Checklist: Screening Version (**PCL:SV**): Screening-Instrument zur Diagnose einer psychopathischen Persönlichkeitsstörung anhand zwölf Merkmalen. (Hart, Cox, & Hare, 1995)

Reversal Learning (**RL**): Test zur Untersuchung der Verhaltensanpassung an Veränderungen in Reiz-Reaktions-Kontingenzen (kognitive Flexibilität). (Clark, Cools, & Robbins, 2004)

Social Readjustment Rating Scale (**SRRS**): Fremdbeurteilungsskala zur Messung psychischer Stressoren anhand positiver und negativer Lebensereignisse. (Holmes & Rahe, 1967)

Speeded Motor **Tapping** Test: Computergestützte Messung der neuromotorischen Geschwindigkeit. (Eigenentwicklung)

**Stroop**-Test: Test zur Messung der individuellen Interferenzneigung. (Bäumler, 1984)

Strukturiertes Klinisches Interview für DSM-IV Achse II (**SKID-II**): Zur Diagnose von Persönlichkeitsstörungen nach Achse II, DSM-IV. 1. Selbstbeurteilungsfragebogen zur Erfassung von Merkmalen einer Persönlichkeitsstörung. 2. Strukturiertes Interview zur genaueren Überprüfung von Störungsbereichen, welche im Fragebogen einen cut-off-Wert überschritten haben. (Fydrich, Renneberg, Schmitz, & Wittchen, 1997)

Symptom Checkliste-90-R (**SCL-90-R**): Selbstbeurteilungsfragebogen zur Messung der subjektiven, durch körperliche und psychische Symptome, empfundene Beeinträchtigung, innerhalb der letzten sieben Tage. (Franke & Derogatis, 2002)

Tower-of-Hanoi (**TOH**): Modifizierte computergestützte Version zur Erfassung der Fähigkeit des planvollen Vorgehens zum Erreichen eines Zielzustandes. (Kotovsky, Hayes, & Simon, 1985)

Trail-Making-Test (**TMT A/B**): Test zur Messung der Verarbeitungsgeschwindigkeit (A) sowie der kognitiven Flexibilität (B). (Partington & Leiter, 1949)

Verbal Fluency (**VF**): Test zur Messung der phonematischen sowie semantischen Wortflüssigkeit bei zeitgleicher Anwendung vorgegebener Regeln. (Oswald & Fleischmann, 1997)

Wisconsin Card Sorting Test (**WCST**): Modifizierte computergestützte Version zur Erfassung der kognitiven Flexibilität. (Berg, 1948)

Zahlenspanne: Test zur Erfassung der Merkfähigkeitsspanne des Arbeitsgedächtnisses passiv (Zahlenspanne Vorwärts (**ZV**)) und aktiv (Zahlenspanne Rückwärts (**ZR**)). (Wechsler, 1981).

## Kongressbeiträge und Publikationen über die MIPS-Studie

- T. Spielmann, M. Graf, M. O. Pflueger, C. Boillat, G. Deuring, P. Lemoine, „Child sexual exploitation material use: Is there evidence for a new generation of offenders?“. 22nd European Congress of Psychiatrie (EPA) 01.-04.03.2014, Munich. Joint Symposium of the Section for Alcoholism and Drug Addiction and the Forensic Section: Internet, sex and addiction: An unhappy triad.
- C. Boillat, T. Spielmann, G. Deuring, P. Lemoine, M. Graf, M. O. Pflueger, „Is there a new generation of child sexual exploitation material users? First and preliminary results of the Basel MIPS-study (Commensurable Indicators of Paedophilic Offenders - Switzerland)“. 13. International Association for the Treatment of Sexual Offenders (IATSO) Conference, 2014.
- M. Delahaye, T. Spielmann, M. Graf, M. O. Pflueger, „Therapeutic climate in a longitudinal intervention study on sex offenders“. 13. International Association for the Treatment of Sexual Offenders (IATSO) Conference, 2014.
- G. Deuring, „Messbare Indikatoren Pädosexueller Straftäter“. Symposium Prävention sexuellen Missbrauchs Göttingen, 2015.
- C. Boillat, „MIPS - Messbare Indikatoren Pädosexueller Straftäter“. V. Symposium Empirische Forschung in der Forensischen Psychiatrie, Psychologie und Psychotherapie – EFPPP 2015.
- G. Deuring, „Assessment der sexuellen Präferenz bei pädosexuellen Straftätern in einer Virtual Reality“. V. Symposium Empirische Forschung in der Forensischen Psychiatrie, Psychologie und Psychotherapie – EFPPP 2015.
- M. O. Pflueger, T. Spielmann, „ MIPS Project/Basel Switzerland, what we are measuring and some neuropsychological preliminary findings“. School Seminar, School of Psychology, University of Lincoln, UK (18.11.2015).
- T. Spielmann, „Why measure physiological/psychological factors in sexual offenders? - MIPS project/Basel Switzerland, what we are measuring and our preliminary findings“. School Seminar, School of Psychology, University of Lincoln, UK (18.11.2015).
- M. O. Pflueger, G. Deuring, C. Boillat, T. Spielmann, P. Lemoine, M. Graf, „The high-functioning myth of CSEM offenders – A neuropsychological study“. *Abstract eingereicht bei* 14. International Association for the Treatment of Sexual Offenders (IATSO) Conference, 2016.
- T. Spielmann, G. Deuring, C. Boillat, P. Lemoine, M. Graf, M. O. Pflueger, „Differences between contact offenders and child exploitation material offenders in sexual preference, substance use, and personality traits – the Basel MIPS-study“. *Abstract eingereicht bei* 14. International Association for the Treatment of Sexual Offenders (IATSO) Conference, 2016.
- C. Boillat, M. Stutz, G. Deuring, T. Spielmann, P. Lemoine, T. Rosburg, M. Graf, M. O. Pflueger, „Correlations of childhood trauma experiences with personality traits and sex offence-related behaviour in adult contact sex offenders“. *Abstract eingereicht bei* 14. International Association for the Treatment of Sexual Offenders (IATSO) Conference, 2016.

G. Deuring, C. Boillat, T. Spielmann, P. Lemoine, T. Rosburg, M. Graf, M. O. Pflueger, „Implicit tests of sexual preference for the classification of paedophilic sex offenders“. *Abstract eingereicht bei* 14. International Association for the Treatment of Sexual Offenders (IATSO) Conference, 2016.

N. Schwab, C. Boillat, G. Deuring, P. Lemoine, T. Spielmann, H. Hachtel, M. Graf, M. O. Pflueger „Preliminary Data on Commensurable Indicators of Pedosexual Offenders“. *Submitted.*