

Aus der Universitätsklinik für Psychiatrie und Psychotherapie

Tübingen

Abteilung: Allgemeine Psychiatrie und Psychotherapie

mit Poliklinik

Ärztlicher Direktor: Professor Dr. A. J. Fallgatter

**Über die Möglichkeit
nicht-authentische
Beschwerden zu erkennen
Diagnostische Optimierung bei Simulation**

**Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin**

**der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen**

vorgelegt von

Benjamin Tscheuschner

aus

Warstein

2011

Dekan: Professor Dr. I. B. Autenrieth

1. Berichterstatter: Professor Dr. A. Stevens

2. Berichterstatter: Professor Dr. A. J. Fallgatter

Widmung

Meiner Frau Tatjana Tscheuschner.

Inhaltsverzeichnis

Widmung		2
Einleitung		6
Theorie		6
Aktueller Forschungsstand		8
Beschwerdenvalidierungstests		9
Ziel der Studie		10
1	Die verschiedenen Tests	12
1.1	Strukturierter Fragebogen simulierter Symptome (SFSS)	12
1.2	Word Memory Test (WMT)	13
1.3	Verhaltens- und Erlebensinventar (VEI)	15
1.3.1	VEI – Validitätsskalen	15
2	Patienten	17
2.1	Alter	19
2.2	Geschlecht	20
2.3	Bildung	20
2.4	Unfalldatum	20
2.5	Zerebrale Diagnosen	22
2.6	Beschwerdenschilderung	23

2.7	Testergebnisse der Patienten	24
2.7.1	SFSS, WMT oder VEI Validitätsskalen bestanden?	24
2.7.2	Mittelwerte und Standardabweichungen des WMT und VEI	25
2.8	Zusammenfassendes Patientenbild	27
3	Testauswertungen	28
3.1	Beschwerdemuster von Antwortverzerrern	28
3.1.1	Beschwerdeschilderung der Testergebnisgruppen	28
3.1.2	Passende Beschwerdeschilderung?	30
3.1.3	VEI-Beschwerden bei SFSS und WMT	33
3.2	Stärke der VEI-Validitätsskalen bei Antwortverzerrern	35
3.2.1	Korrelation zwischen VEI Validitätsskalen und SFSS / WMT	35
3.2.2	VEI-Validitätsskalen und SFSS bzw. WMT	37
4	Diskussion	40
4.1	Interpretation	40
4.1.1	Beschwerdemuster von Antwortverzerrern	40
4.1.2	Der VEI als Test zur Simulationsdiagnostik	41
4.2	Vergleich	41
4.2.1	Die Beschwerdemuster	41
4.2.2	Die VEI Validitätsskalen	42

4.3	Schwachpunkte	42
5	Zusammenfassung	43
5.1	Hintergrund	43
5.2	Methoden	43
5.3	Resultate	44
5.4	Konklusion	44
	Literaturverzeichnis	45
	Danksagung	50
	Lebenslauf	51

Einleitung

Theorie

Menschen kommunizieren immer wieder nicht ehrlich untereinander. Vor allem dann nicht, wenn sie sich dadurch einen Vorteil erhoffen. Wenn nun ein Patient zum Arzt oder Gutachter kommt und ihm seine Beschwerden schildert oder aber bestimmte Beschwerden nicht schildert; wenn eine Patient Antworten in ihrem Wahrheitsgehalt verzerrt oder aber Symptome vortäuscht; ist es dann nicht die Aufgabe der Ärzte und gerade der Gutachter herauszufinden, ob denn die Beschwerden korrekt angegeben, Antworten wahrheitsgemäß beantwortet wurden, um eine zielgerichtete Therapie oder Maßnahme einleiten zu können? Ist nicht gerade bei zivil- und strafrechtlichen Begutachtungen mit negativen Antwortverzerrungen oder vorgetäuschten Symptomen zu rechnen? Wenn man gerade in diesem Bereich die Augen verschließen würde, dann könnte das deutliche Konsequenzen für die Gesellschaft haben, wenn etwa ein Sexualstraftäter aufgrund seiner möglicherweise falsch positiv verzerrten Antworten vorzeitig entlassen wird.

Natürlich ist es nicht einfach zu erkennen, ob jemand Beschwerden nur vortäuscht oder Antworten verzerrt (Smith et al. 1997). Umso wichtiger ist es Methoden zu entwickeln, die diese Täuschungen so präzise aufdecken können, dass man möglichst einem aufrichtigen Menschen auch nicht unrecht tut und ihm möglicherweise fälschlich Unehrllichkeit unterstellt (Merten et al. 2009).

Wie erwähnt ist gerade im gutachterlichen Kontext (Benbadis 2005, Mittenberg et al. 2002) mit einer ausgesprochenen Häufigkeit von Antwort- als auch Beschwerdenverzerrungen zu rechnen. Logischerweise wirkt sich das auch auf die Testergebnisse aus. Der Gutachter handelt im Auftrag der Gesellschaft und ist geradezu dazu bestimmt die Aussagen des Begutachteten kritisch zu hinterfragen und zu prüfen. Um zu einem wahrheitsgemäßen Urteil zu kommen, braucht er eine qualitativ hochwertige Diagnostik der Validität der geschilderten Antworten und Beschwerden. Es ist angemessen dies als medizinisch not-

wendig zu beschreiben (Bush et al. 2005), gerade weil auch hier in Deutschland ein recht hoher Anteil von neurologisch-psychiatrisch Begutachteten bei Einsatz moderner Verfahren der Diagnostik als auffällig eingestuft werden können (Merten et al. 2006, Merten et al 2007).

Im forensischen Kontext scheint dieses Problem besonders ausgeprägt zu sein, da man schätzt, dass dort die Simulationsrate (absichtliches Erzeugen oder starkes Übertreiben körperlicher Symptome)(Resnick 1997) bei etwa 16% liegt (Rogers et al. 1994). In einer neueren Studie lag bei 89% der Probanden eine klare Diskrepanz zwischen angegebenen und beobachteten Symptomen vor (Edens 2007). Das verwundert auch nicht, denn als Straftäter hofft man durch z.B. Simulation auf Strafaufschubung, Strafmilderung oder bequemere Unterbringung in einer psychiatrischen Anstalt (Resnick 1997).

Um Simulation zu erkennen wurden im Allgemeinen 3 Methoden nachgegangen. Eine davon stützt sich auf das vorhandene Persönlichkeitsprofil (Schretlen 1998). Bei diesem Zugang geht man davon aus, dass Simulanten in einer diffizil abweichenden Art auf standardisierte Testanweisungen antworten. Um solche Antwortmuster zu entdecken, enthalten solche standardisierten Instrumente sogenannte Validitätsskalen. Das Minnesota Multiphasic Personality Inventory; MMPI-2 (Butcher et al. 1989) ist ein Beispiel für ein Selbsteinschätzungsinstrument mit Validitätsskalen, von denen eine die F Skala ist.

Einen weiteren Zugang zur Aufdeckung von Simulation stellen Interviewtechniken und Verhaltensbeobachtungen dar. Resnick (Resnick 1984) stellte z. B. fest, dass Simulanten bei üblichen Interviews einige bestimmte charakteristische Merkmale erkennen lassen. Das „Structured Interview of Reported Symptoms“ *SIRS* (Rogers et al. 1992) ist ein gutes Beispiel einer standardisierten Interviewtechnik, die diese charakteristischen Merkmale, wie z.B. ungläubhafte Symptome, widersprüchliche Darstellung, plötzliches Auftreten von Symptomen, Übertreiben von Symptomen und Ernsthaftigkeit der Symptome, hervorhebt. Allerdings werden solche Interviewtechniken praktisch dadurch limitiert, dass es viel Zeit und braucht und vor allem gut ausgebildete Interviewer.

Der 3. Zugang bezieht sich auf psychologische Tests, die speziell für die Erkennung von Simulation konzipiert wurden. Beispiele solcher Tests sind der „Coin-in-hand“-Test (Cochrane et al. 1998) und der M-Test (Grow et al. 1980), die besonders simulierte Amnesie beziehungsweise simulierte Schizophrenie erkennen können. Leider erfassen diese Testverfahren nur eine begrenzte Gruppe von Symptomen (Cima et al. 2003).

Aktueller Forschungsstand

Seit den 1990ern ist eine große Anzahl von empirischen und theoretischen Arbeiten erschienen, die sich damit befassen haben, authentische von nicht-authentischen kognitiven, psychoemotionalen und körperlichen Beschwerden zu unterscheiden. Bis vor wenigen Jahren fiel diese Art der Arbeit unter den Begriff „Simulationsdiagnostik“. Hier gab es einen Schwerpunkt in der klinischen und vor allem forensischen Neuropsychologie, da hier immer wieder an Methodiken gearbeitet worden ist. Es wurden zuverlässige Simulationsmarker gesucht, die nicht-authentische kognitive Störungen identifizieren und sogar Antwortverzerrungen entlarven können.

Seit 2002 werden die zuvor als „Simulationstests“ aufgefassten Verfahren und die „Simulationsdiagnostik“ generell konzeptionell eher sorgfältiger und politisch korrekter als Diagnostik der Beschwerdenuvalidität benannt.

Suboptimales Leistungsverhalten liegt vor, wenn ein Untersuchter wegen z.B. mangelnder Motivation in einem Test schlechter als sein tatsächliches Fähigkeitsniveau abschneidet. Einem solchen Verhalten, das auch weiter gefasst als negative Antwortverzerrung bezeichnet werden kann, können vorsätzliche Verzerrungsversuche, Verhaltensstörungen oder psychische Erkrankungen zugrunde liegen. Ein direkter Schluss von negativen Antwortverzerrungen auf eine Simulation ist sowohl sachlogisch nicht korrekt, als auch ethisch nicht vertretbar.

Es herrscht heute mittlerweile eine weitgehende Übereinstimmung darüber, dass eine Entscheidung über das Vorliegen negativer Antwortverzerrungen nie auf der Grundlage eines einzelnen Testes oder Indikators erfolgen sollte, sondern immer mehrere Verfahren oder Ansätze zur Diagnostik der Beschwerdvalidität zu berücksichtigen sind (Tan et al. 2002, Lynch 2004). Dabei sollten auch bekannte Anwendungsbeschränkungen der benutzten Verfahren wie mangelnde empirische Validierung beachtet werden (Martelli et al. 2002). Man ist sich einig, dass eine qualifizierte Beurteilung der Beschwerdvalidität ein komplexes multimodales Herangehen (Stieglitz 2005) braucht, das mit einem hohen Maß an Sachkenntnis und Verantwortungsbewusstsein gepaart ist. Nur so kann die Anzahl falscher Entscheidungen mit häufig weitreichenden Folgen klein gehalten werden (Blaskewitz et al. 2007).

Beschwerdvalidierungstests

Eine sehr umfangreiche und trotzdem noch wachsende internationale Literatur belegt, dass Beschwerdvalidierungstests durch Überprüfung der Leistungsmotivation sehr gut in der Lage sind, hirnormale Patienten mit nachweisbaren Erkrankungen wie schwerem Schädel-Hirn-Trauma oder Epilepsie von solchen zu trennen, die im Rahmen von Rechtsstreitigkeiten ohne nachweisbare, objektive Störungen subjektive Leistungsstörungen berichten (z.B. Binder & Rohling, 1996) – ja, dass Beschwerdvalidierungstests notwendig sind (van Gorp et al. 1999). Auch bei psychiatrischen Patienten trennen diese Verfahren zuverlässig (Goldberg, Back-Madruga & Boone, 2007).

Es gibt anhand solcher Verfahren eindeutige Kriterien oder Entscheidungsregeln für die Annahme von „Malingering“, d.h. die vorsätzliche Täuschung / Aggravation von Symptomen, um meist finanzielle Vorteile zu erhalten (Slick, Sherman & Iverson, 1999).

Es gibt unter Beschwerdvalidierungsinstrumenten erhebliche qualitative Unterschiede. Die hier angewandten beschwerdvalidierenden Instrumente sind die international best erforschten. Durch die Vielzahl der anhand großer

Patientenkollektive multizentrisch und international wiederholt kreuzvalidierten Ergebnisse und deren Veröffentlichungen sind inzwischen die sehr guten Testeigenschaften belegt worden: Die Sensitivität und die Spezifität der hier angewandten Tests sind bisher nicht überboten worden. Der erste hier eingesetzte Beschwerdenuvalidierungstest ist nach mehreren autorunabhängigen Studien in der Lage, 100% aller vorgetäuschten kognitiven Störungen zu erkennen, d.h. die Sensitivität für unzureichende Anstrengung lag bei 100% (Tan et al., 2002; Brockhaus & Merten, 2004; Merten et al., 2004). Keiner der kooperativen Untersuchten dagegen wurde fälschlicherweise als nicht-kooperativ eingestuft, d.h. auch die Spezifität lag bei 100%. Sensitivität gibt den Prozentsatz der untersuchten simulierenden Probanden an, die als solche durch einen Test korrekt identifiziert worden sind, und Spezifität bedeutet den Prozentsatz der mit einem Test untersuchten (u. U. hirneorganisch verletzten) optimal arbeitenden Probanden, die korrekt identifiziert worden sind.

Im Übrigen ist anhand dieser Instrumente vergleichsweise viel Erfahrung mit einzelnen klinischen Gruppen gesammelt worden. Diese klinische Erfahrung ist in der diagnostischen beschwerdenuvalidierenden Arbeit mit einzelnen Patienten mindestens so wichtig wie die Testeigenschaften der einzelnen Beschwerdenuvalidierungstests selber – und zwar deshalb, weil man wissen muss, wie bona fide kranke, sich bemühende Patienten einer klinischen Gruppe mit den Beschwerdenuvalidierungstests umgehen. Solche statistisch untermauerten gruppenspezifischen Erwartungen über das kognitive Verhalten im Umgang mit einem bestimmten Beschwerdenuvalidierungstest, die durchaus im frappanten Gegensatz zu gängigen klinischen Annahmen stehen können, nennt man „*positive*“ bzw. „*negative predictive power*“ des Instrumentes. Ins Deutsche übertragen heißt das die Vorhersagekraft eines Instrumentes, bei auffälligem (positivem) Befund einen tatsächlich suboptimal arbeitenden Probanden und bei unauffälligem (negativem) Befund einen tatsächlich optimal arbeitenden Probanden wahrscheinlich identifiziert zu haben.

Ziel der Studie

Aufgrund der enormen Wichtigkeit der Diagnostik der Beschwerdvalidität gerade im gutacherlichen Bereich, aber auch der komplexen multimodalen Herangehensweise, werden in dieser Arbeit Fragen zweier Themenbereiche zu beantworten versucht.

1. Welche Beschwerden werden durch Patienten, die Antworten verzerrt haben, angegeben, die sie tatsächlich gar nicht haben? Kann man eine bestimmte Tendenz erkennen? Gibt es ein bestimmtes Muster?

2. Wie stark ist der Test „Der Verhaltens- und Erlebensinventar“ (VEI), die deutsche Fassung des Personality Assessment Inventory, innerhalb der Diagnostik der Beschwerdvalidität? Kann er signifikant valide Aussagen zu Antwortverzerrungen machen? Gibt es Skalen, die ähnlich hohe Trefferquoten haben wie der WMT oder SFSS?

1 Die verschiedenen Tests

1.1 Strukturierter Fragebogen simulierter Symptome (SFSS)

Der Strukturierte Fragebogen simulierter Symptome (SFSS bzw. im englischsprachigen Raum SIMS („Structured Inventory of Malingered Symptomatology“ SIMS (Smith et al. 1997))) ist ein leicht durchführbarer Selbsteinschätzungsfragebogen, um Simulation und eine Vielzahl simulierter Symptome zu erkennen. Der SIMS umfasst 75 Items, bei denen man jeweils „ja“ oder „nein“ ankreuzen kann. Die Items wurden sowohl aus unterschiedlichen Quellen, z. B. MMPI und SIRS, zusammengesucht, da sich diese Tests in vergangenen Forschungen in bestimmter Hinsicht als brauchbar zur Erkennung von Simulation erwiesen haben (Smith 1997). Hierbei wurden die Items dann so umgeformt, dass sie sowohl eine bestimmte Aufdeckungsstrategie für Simulation enthalten als auch eine erhöhte Sensitivität für das Vortäuschen einer bestimmten Störung aufweisen. Zum anderen wurden auch zusätzliche Items entwickelt, welche die qualitativen Charakteristika von Simulanten, die in früheren Arbeiten festgestellt wurden (Resnick 1984, Rogers 1984, Seamons et al. 1981), widerspiegeln. Die Items wurden dann auf 5 Skalen zu jeweils 15 Items aufgeteilt. Eine Skala steht dann jeweils für eine der häufig simulierten Störungen: Neurologische Beeinträchtigung (N), Affektive Störung (AF), Psychose (P), Niedrige Intelligenz (NI) und Amnestische Störung (AM). Einige Items haben absichtlich einen ungewöhnlichen Inhalt (z.B. „Wenn ich Stimmen höre, fühlt es sich an, als würden meine Zähne aus dem Körper heraustreten“) oder beziehen sich auf unwirkliche Symptome (z.B. „Ich habe einen körperlichen Schmerz, so als ob Ameisen unter meiner Haut kribbelten“). Andere Items zielen auf eine ganz bestimmte Erkrankung (z.B. Depression) ab, wobei nur ein Fachmann erkennen kann, dass es sich hierbei um sehr untypische Symptombeschreibungen handelt (z.B. „Je depressiver ich bin, um so mehr möchte ich essen“). Ganz andere Items beinhalten die Wahl einer nahezu richtigen Antwort (z.B. „Die Woche hat 6 Tage“).

Die Antworten, die auf Simulation hinweisen, werden zum Gesamt-SIMS-Wert zusammengezählt. Zu jeder der 5 Subskalen kann ein einzelner Score ermittelt werden.

Eine erste Untersuchung des SIMS (Smith et al. 1997) wurde mit 476 Studenten durchgeführt. Die Studenten wurden in eine Kontrollgruppe, die ehrlich antworten sollte, oder in einzelne Simulantengruppen, die verschiedene Störungen simulierten, aufgeteilt. Der Gesamt-SIMS-Wert war dabei der effizienteste Indikator für Simulation. Er erkannte 95,6% der Simulanten (Sensitivität) und 87,9% der ehrlich antwortenden Studenten (Spezifität).

Rogers et al. (1996) untersuchten die Effizienz des SIMS bei jungen Erwachsenen. Insgesamt füllten 53 Teilnehmer den SIMS ehrlich oder gezielt unehrlich aus. Auch hier zeigte sich wiederum eine hohe Sensitivität (87%). Die Spezifität (62%) fiel in dieser Studie niedriger als in der ursprünglichen Smith-und-Burger-Studie (Smith et al. 1997) aus.

Es wurde bei einem Cutoff von 16 im Gesamt-SIMS-Wert eine PPP von 87% erzielt. Obwohl dieser Wert für eine Screeningmethode akzeptabel ist, impliziert er auch, dass 13% der Individuen mit SIMS-Werten oberhalb des Cutoffs nicht simulieren. Rogers et al. berichteten über eine NPP von 94% für diesen Cutoff-Wert. Im SIMS werden also relativ wenige falsch negative Ergebnisse produziert (Cima et al. 2003).

Aufgrund der einfachen Verständlichkeit, der hohen Sensitivität und der kurzen Form ist der SIMS als Screeningmethode zu empfehlen.

1.2 Word Memory Test (WMT)

Der Word Memory Test wurde um 1994 vom Paul Green vorgestellt. Er ist ein Verfahren, das speziell zur Erfassung von negativen Antwortverzerrungen und Unkooperativität in der neuropsychologischen Diagnostik entwickelt wurde. Er ist so erfolgreich, dass er inzwischen in verschiedenste Sprachen adaptiert wurde und auf der ganzen Welt von sehr vielen Neuropsychologen benutzt wird

(zum Beispiel in Kanada, den Vereinigten Staaten, Neuseeland, Australien, Großbritannien, den Niederlanden, Spanien, Deutschland, Argentinien oder Litauen). Im WMT werden 20 Wortpaare in zwei aufeinanderfolgenden Lerndurchgängen am Computerbildschirm zum Einprägen dargeboten. Jedes Paar regt zwei verschiedene, miteinander assoziativ verknüpfte alltagsnahe Vorstellungen an (z. B. Himmel–Sonne; Ente–Teich). Anschließend ist jedes einzelne dieser Wörter aus Alternativen (z. B. Himmel–Erde; Stern–Sonne) auszuwählen. Nach einem Intervall von 30 min werden die Wörter nun mit anderen Distraktoren (z. B. Himmel–Regen; Mond–Sonne) erneut abgefragt (Verzögerte Wiedererkennung). Als Grad der Übereinstimmung zwischen den Einzelantworten in diesen Durchgängen wird ein Konsistenzmaß bestimmt (Green 2005).

Personen, die zweifelsfrei an einer schweren Hirnverletzung bzw. einer neurologischen Erkrankung wie Hirntumor oder Aneurysmablutung litten, konnten ohne Mühe die beschriebene Wiedererkennungsaufgabe erfüllen und erreichten einen mittleren Testwert von rund 96 Prozent Richtigen (Green 2002). Die Leistungsmotivation eines Probanden muss dann ernsthaft infrage gestellt werden, wenn seine erreichten Punkte unterhalb eines Trennwertes (cutoff) fallen. Dieser Trennwert liegt etwa drei Standardabweichungen unterhalb des Mittelwertes jener Personen, die schwere Hirnverletzungen oder neurologische Erkrankungen aufwiesen. Wenn man den Intelligenzquotienten zum Vergleich heranzieht, würde dies einem IQ von weniger als 55 entsprechen, wobei allerdings zu berücksichtigen ist, dass die WMT-Werte nicht normalverteilt sind. Ungültige Testergebnisse sind besonders dann wahrscheinlich, wenn ein Proband, der den Symptomvalidierungstest nicht bestanden hat, eine Diagnose wie etwa leichtes Schädel-Hirn-Trauma, Depression, Fibromyalgie oder chronisches Müdigkeitssyndrom hat. Auch andere Diagnosen, zu denen üblicherweise keine ausgesprochen schweren kognitiven Störungen gehören, sondern bei denen im Allgemeinen überhaupt keine kognitiven Defizite in Leistungstests erscheinen, kommen hier infrage (Green et al. 2004).

1.3 Verhaltens- und Erlebensinventar (VEI)

Der Verhaltens- und Erlebensinventar (VEI), die deutsche Fassung des Personality Assessment Inventory (Morey 1991), ist ein mehrdimensionaler Persönlichkeitsinventar. Neben der ursprünglichen Form existieren mehrere Computerversionen sowie eine Kurzform. Er sollte in seiner Art eine überlegene Alternative zum MMPI-2 (Butcher et al. 1989) werden. Und White kommt auch zu folgender Schlussfolgerung: "When compared with the MMPI-2, the PAI has a number of significant advantages. The PAI is more logically structured, it is better designed, it is more relevant to contemporary issues and terminologies, it is easier for subjects to complete, it is easier for clinicians to score and it is more straight forward to interpret." (White 1993)

Der VEI besteht aus 344 Items und 22 Skalen. Unter diesen 22 Skalen befinden sich auch 4 Skalen für die Gültigkeit des Testes: Inkonsistenz (überprüft, ob der oder die Befragte konsistente Antworten geliefert hat), Seltenheit (stellt fest, ob nachlässig oder willkürlich geantwortet wurde), ungünstiger und günstiger Eindruck (wenn jemand einen übertrieben günstigen oder ungünstigen Eindruck erwecken will). Seit der Erstveröffentlichung wurden vier zusätzliche Validitätsindikatoren entwickelt, davon zwei zur Aufdeckung negativer Antwortverzerrung (s.u.). Daneben gibt es noch 11 klinische Skalen wie z.B. somatische Beschwerden, Depression oder Paranoia, 5 Behandlungsskalen wie z.B. Aggression, Stress oder Behandlungsablehnung und noch 2 Interpersonale Skalen (Morey 2007) Untersuchungen zu konkurrenenter und diskriminanter Validität bescheinigen dem VEI eine gute Eignung zur Messung von Antwortverzerrung (Duellmann et al. 2004). In einer Studie von Bagby et al. (2002) zeigte sich der VEI robust gegenüber Coaching. Bei einem Vergleich mit dem MMPI-2 erzielte der VEI etwas geringere Effektstärken (Blanchard et al 2003). Dennoch konnte der Test in allen durchgeführten Analysen zu einer Erhöhung der inkrementellen Validität beitragen und lässt eine Verwendung beider sinnvoll erscheinen (Blaskewitz 2005).

1.3.1 VEI – Validitätsskalen

Validitätsskalen geben an, ob der Proband den Fragebogen ordnungsgemäß ausgefüllt hat oder ob er in irgendeiner Hinsicht die Antworten im Fragebogen verfälscht hat. Hier werden die für unsere Studien interessanten Skalen vorgestellt.

1.3.1.1 Inkonsistenz

Die Skala überprüft, ob der oder die Befragte konsistente Antworten geliefert hat. Der Rohwert wird aus 10 Itempaaren berechnet, die hoch (positiv oder negativ) miteinander korrelieren.

1.3.1.2 Seltenheit

Die Seltenheits-Skala stellt fest, ob nachlässig oder willkürlich geantwortet wurde. Die Items haben eine extrem hohe oder extrem niedrige Zustimmungsrates, sind aber neutral hinsichtlich psychischer Auffälligkeiten.

1.3.1.3 Ungünstiger Eindruck

Die Skala spricht an, wenn jemand einen übertrieben ungünstigen Eindruck erwecken will oder auffällige Symptome simuliert. Bei psychisch Kranken haben die Items nur eine relativ niedrige Zustimmungsrates.

1.3.1.4 Rogers Discriminant Function (RDF)

Diese Skala ist ein Index, der aus 20 Skalen und Subskalen zusammengestellt wird, die im besten Falle die Personen abgrenzt, die bestimmte Beschwerden simuliert (Sellbom et al. 2008).

1.3.1.5 Malingering Index (MAL)

Der Malingering Index benutzt 8 bzw. 11 Skalen, um herauszustellen, ob sich jemand verstellt oder nicht (Sellbom et al. 2008)

2 Patienten

In unserer Testreihe wurden 101 Patienten im Medizinischen Gutachteninstitut Tübingen eingehend körperlich, neurologisch und psychiatrisch untersucht und einer psychologischen Testung unterzogen. Alle hatten schriftlich ihr Einverständnis zu einer anonymen Verwertung ihrer Daten zu wissenschaftlichen Zwecken gegeben. Alle Patienten hatten ein Motiv zu simulieren oder dissimulieren. Man konnte die Patienten in zwei Kategorien unterteilen: Motiv der positiven Antwortverzerrung und Motiv der negativen Antwortverzerrung.

In die Kategorie des Motivs der positiven Antwortverzerrung fielen Fahreignungsbegutachtete, Personen, die zur Frage der Eignung für die Verbeamtung untersucht wurden und Personen aus der Justizvollzugsanstalt, bei denen zur Kriminalprognose Stellung genommen werden sollte. Beide Personengruppen hätten eher das Anliegen sich in einem positiven Licht darzustellen um die Chancen auf den erneuten Erwerb der Fahrerlaubnis bzw. auf vorzeitige Entlassung zu verbessern.

In die Kategorie des Motivs der negativen Antwortverzerrung fielen Personen, die bei der Berufsgenossenschaftlichen Versicherung eine Erwerbsminderungsrente beantragten (59,4%), Personen, die bei der privaten Haftpflichtversicherung eine Entschädigung geltend machen wollten (13,9%), Beamte, die Dienstunfähigkeit beantragten (7,9%), Verunfallte, die von der privaten Unfallversicherung eine Entschädigung forderten (7,9%) oder andere (s. Tabelle 1). Allen Personen liegt eher etwas daran, die Antwort negativ zu verzerren, um ihrem Ziel näher zu kommen.

Tabelle 1

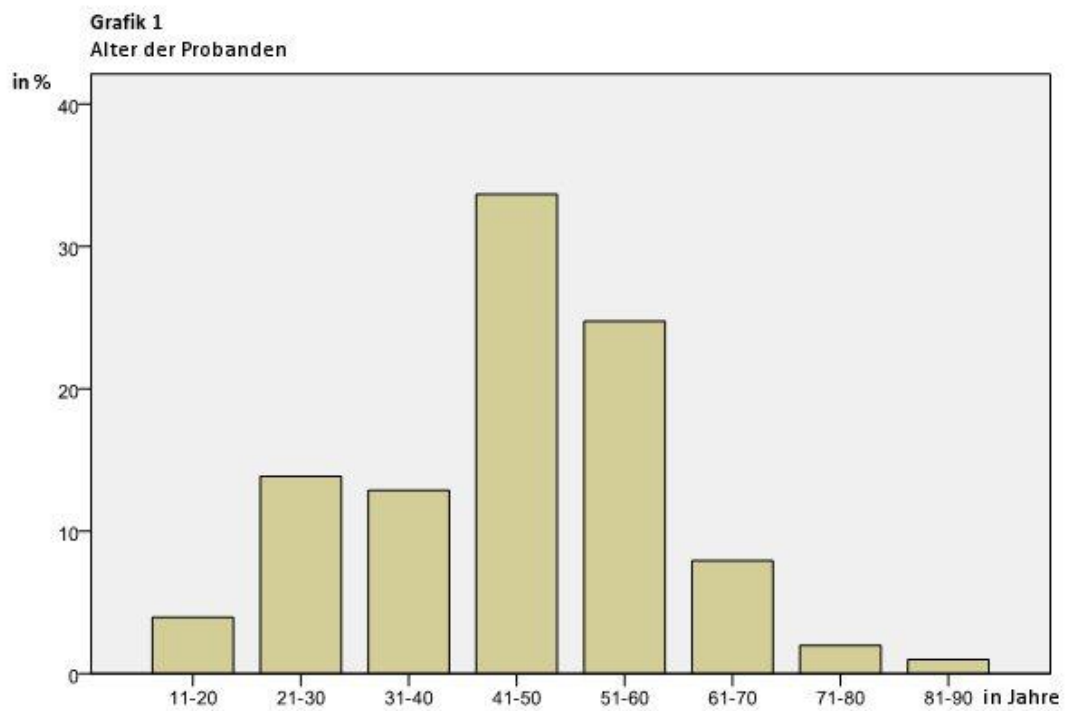
Wo wurde Entschädigung beantragt?

	Anzahl	Prozent
Berufsgenossenschaft	60	59,4
Priv. Haftpflichtversiche-	14	13,9
Dienstunfähigkeit	8	7,9
Private Unfallversicherung	8	7,9
Sonstige	11	10,9

Da in dieser Studie die negative Antwortverzerrung untersucht werden soll, wurde die erste Personengruppe, die eher das Motiv einer positiven Antwortverzerrung hat, von den nachfolgenden Analysen ausgeschlossen. In diese Personengruppe fielen die Fahreignungsbegutachteten, die Anwärter für die Verbeamtung und Personen aus der Justizvollzugsanstalt. Somit hat sich die Personen- bzw. Patientenanzahl auf 101 reduziert.

2.1 Alter

Das Alter der Patienten liegt hauptsächlich im mittleren Alter. 34,6% liegen zwischen 41-50 Jahren. Und 71,4% der Patienten liegen zwischen 31-60 Jahren (s. Grafik 1).



Das Durchschnittsalter liegt bei einer Standardabweichung von 13,8 Jahren bei 45,1 Jahren. Das Alter reicht von 19 Jahren bis hin zu 83 Jahren.

2.2 *Geschlecht*

Von den Probanden sind 66,3% männlich und 33,7% weiblich (s. Tabelle 2).

Tabelle 2

Geschlecht der Probanden

	Anzahl	Prozent
männlich	67	66,3
weiblich	34	33,7

2.3 *Bildung*

Die Bildungsebenen werden von geringeren zur höheren Schulausbildung aufgezählt. 1% der Probanden haben entweder gar keinen oder einen Sonder- schulabschluss, 54,5% haben die Hauptschule abgeschlossen, 27,7% haben die Realschule abgeschlossen und 16,8% haben das Abitur gemacht (s. Tabelle 3). Der Hauptteil der Probanden hat also nur die Hauptschule abgeschlossen.

Tabelle 3

Schulabschluss der Probanden

	Anzahl	Prozent
keiner/Sonderschule	2	1
Hauptschule	58	54,5
Realschule	28	27,7
Gymnasium	19	16,8

2.4 *Unfalldatum*

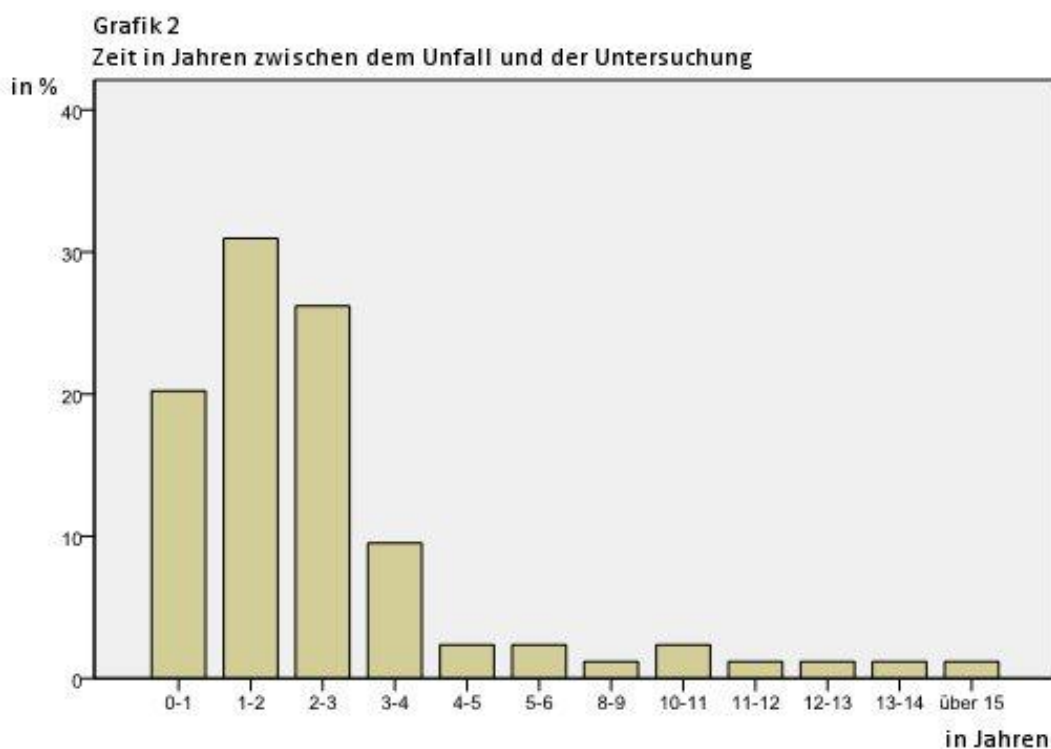
83,2% der Probanden hatten einen Unfall in der Vergangenheit. Von diesen 79,4% fallen 77,4% aller Unfälle in den Zeitraum der letzten 3 Jahren (s. Tabelle 4 – siehe Anhang).

Tabelle 4

Zeit in Jahren zwischen dem Unfall und der Untersuchung

	Anzahl	Prozent
0-1	17	16,8
1-2	26	25,7
2-3	22	21,8
3-4	8	7,9
4-5	2	2
5-6	2	2
8-9	1	1
10-11	2	2
11-12	1	1
12-13	1	1
13-14	1	1
Über 15	1	1

Der Unfall, der am weitesten zurückliegt, fand vor 27 Jahren statt. Der am wenigsten zurückliegende Unfall war vor gerade einmal einem Monat.



2.5 Zerebrale Diagnosen

Bei 40,4% der Patienten wurde eine contusio cerebri festgestellt, bei 14,1% eine commotio cerebri und bei 45,5% wurde gar keine Gehirnbeteiligung festgestellt (s. Tabelle 5).

Tabelle 5

Zerebrale Diagnosen der Probanden

	Anzahl	Prozent
Keine	45	45,5
Commotio Cerebri	14	14,1
Contusio Cerebri	40	40,4

Psychische Diagnosen wurden wie folgt gestellt: Bei 0% der Patienten wurde eine Angststörung festgestellt. Lediglich 1% leiden unter einer diagnostizierbaren Depression. Bei 32,3% wurden sonstige psychische Diagnosen nach DSM-IV gestellt (s. Tabelle 6).

Somatische Diagnosen wurden wie folgt gestellt: Bei 2% wurde ein HWS-Syndrom diagnostiziert. Niemand litt unter einer posttraumatischen Belastungsstörung. Bei 52,5% wurden weitere somatische Diagnosen gestellt (s. Tabelle 6).

Tabelle 6

Psychische und somatische Diagnosen der Probanden

	Anzahl	Prozent
Angststörung	0	0
Depression	1	1
Sonstige psych. Diagnose	32	32,3
HWS-Syndrom	2	2
PTBS	0	0
Weitere somat. Diagnosen	52	52,5

2.6 Beschwerdenschilderung

7,1% der Probanden gaben eine Angstsymptomatik an, 10,1% eine depressive Symptomatik, 9,1% Posttraumatische Belastungsstörungen und 66,7% sonstige psychische Beschwerden (s. Tabelle 7). Im somatischen Bereich gaben 79,8% medizinische Beschwerden und 12,1% HWS-Beschwerden an (s. Tabelle 7).

Tabelle 7

Beschwerdenschilderung der Probanden

	Anzahl	Prozent
Angstsymptomatik	7	7,1
Depressive Symptomatik	10	10,1
PTBS	9	9,1
Sonstige psych. Beschwerden	66	66,7
Medizinische Beschwerden	79	79,8
HWS Beschwerden	12	12,1

2.7 Testergebnisse der Patienten

Einen groben Überblick über die Patienten werden uns die Ergebnisse der Patienten in den jeweiligen Tests geben.

2.7.1 SFSS, WMT oder VEI Validitätsskalen bestanden?

Im SFSS hatten 67,3% einen unauffälligen Wert, wohingegen 32,7% auffällige Werte aufwiesen, die auf Simulation hindeuten (s. Tabelle 9). Den WMT haben 66,3% bestanden und 33,7% zeigten eine mangelnde Kooperation, d.h. sie zeigten eine unzureichende Anstrengungsbereitschaft (s. Tabelle 9).

Tabelle 8

Auffällige Werte in SFSS und WMT?

	Anzahl	Prozent
SFSS	33	32,7
WMT	34	33,7

Bei den VEI-Validitätsskalen hatten 5% eine Erhöhung in der Inkonsistenz, 16,9% eine Erhöhung in der Seltenheit und 10,9% eine Erhöhung beim ungünstigen Eindruck (Tabelle 22-24 – siehe Anhang).

Tabelle 9

Erhöhung der Validitätsskalen bei den Probanden

	Anzahl	Prozent
Inkonsistenz	5	5
Seltenheit	17	16,9
Ungünstiger Eindruck	11	10,9

2.7.2 Mittelwerte und Standardabweichungen des WMT und VEI

2.7.2.1 WMT

Tabelle 10

Mittelwerte und Standardabweichungen bei den Variablen des WMT

	Mittelwert	Standardabweichung
WMT: Immediate Recognition	90,9	11,3
WMT: Delayed Recognition	90,8	11
WMT: Consistency	87,3	13
WMT: Multiple Choice	77,1	19,4
WMT: Paired Associates	73,8	20,2
WMT: Free Recall (FR)	49	19,6
WMT: Lond delayed FR	50,8	20,7

Der Durchschnitt würde den WMT bestehen, denn bei allen drei ersten Werten liegen die Probanden über dem „Cut-off“.

2.7.2.2 VEI

In dieser Auswertung fallen die Skalen „somatische Beschwerden“, „Depression“ und „Manie“ auf. Bei den somatischen Beschwerden liegt der Mittelwert an der oberen Grenze der leichten Erhöhung. Es liegt schon fast im erhöhten Bereich, wo man schon von manifesten somatischen Beschwerden

sprechen kann. Der Mittelwert liegt bei dem Parameter Depression im T-Wert-Bereich zwischen 60-69; so schätzen sich die Teilnehmer im Schnitt als eher sensibel, pessimistisch und selbstzweifelnd ein. Bei der Manie liegt der Wert eindeutig im unteren Bereich. Aufgrund dieser Auffälligkeiten behalten wir die somatischen Beschwerden und die Depression im Hinterkopf.

Tabelle 11

Mittelwerte und Standardabweichungen bei den Variablen des VEI

	Mittelwert	Standardabweichung
Inkonsistenz	51,3	8,3
Seltenheit	52	8,7
Ungünstiger Eindruck	55,2	13,4
Günstiger Eindruck	53	11,8
MAL*	53,3	12
RDF*	55,9	11,7
Somatische Beschwerden	69,1	13,4
Angst	58,5	14,2
Depression	64,4	15,1
Manie	44,3	9,8
Paranoia	52,3	12
Schizophrenie	54,1	13,8
Borderline-Merkmale	51,7	11,5
Antisoz. Verhaltensmerk-	47,3	9,5
Alkoholprobleme	47,26	8,6
Drogenprobleme	50,7	8,7
Aggression	47	8,3
Suizidalität	55	15,9
Stress	51,8	9,6
Unterstützungsdefizit	52,7	12,9
Behandlungsbalehnung	52	10,8
Dominanz	44,3	9,7
Wärme	46,5	11

*T-Werte mit Grenzwerten nach Groves 2009

2.8 Zusammenfassendes Patientenbild

Die meisten Patienten liegen mit dem Alter zwischen 31 und 60 Jahren (71,4%), sind männlich (66,3%) und haben die Hauptschule abgeschlossen (54,5%). Im Weiteren hat der Großteil der Patienten einen Unfall in der Vergangenheit gehabt (83,2%) und medizinische (79,8%) oder sonstiges psychische Beschwerden (66,7%) angegeben. Zuletzt ist noch markant, dass 33,7% sich als unkooperativ im WMT zeigten und 32,7% unglaubwürdige Symptome im SFSS aufzeigten.

3 Testauswertungen

Nun wird der Frage nachgegangen, in welchem Zusammenhang die Tests stehen. Insbesondere werden die Zusammenhänge von dem VEI zum SFSS bzw. WMT untersucht und bewertet.

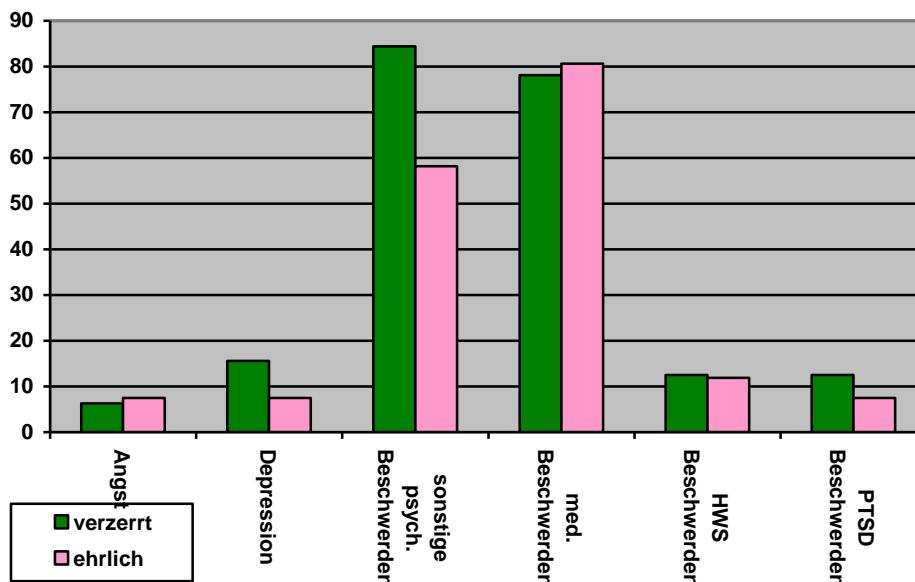
3.1 Beschwerdemuster von Antwortverzerrern

Welche Beschwerden werden durch Patienten, die Antworten verzerrt haben, angegeben, die sie tatsächlich gar nicht haben? Kann man eine bestimmte Tendenz erkennen? Gibt es ein bestimmtes Muster?

3.1.1 Beschwerdeschilderung der Testergebnisgruppen

Zu Beginn wird untersucht, welche Beschwerden diejenigen schildern, die die Testergebnisse des WMT und SFSS negativ verzerrten bzw. ehrlich geantwortet haben.

3.1.1.1 SFSS

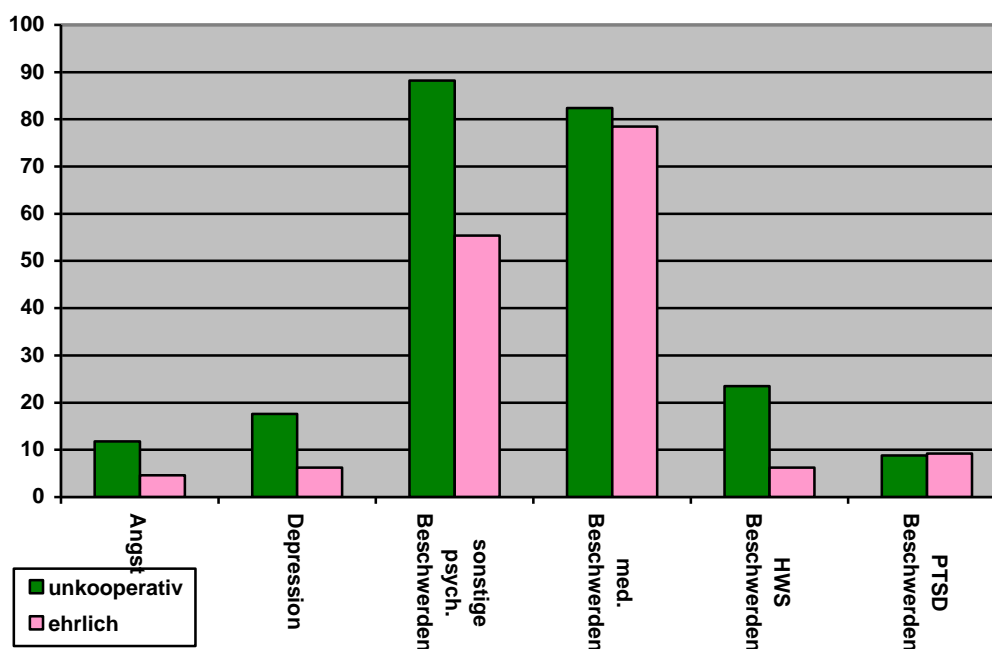


Auch hier sehen wir eine ähnliche Tendenz. Diejenigen, die eine Antwort negativ verzerrten, hier unglaubwürdige Beschwerden schildern, zeigen logischer-

weise eine höhere Tendenz Beschwerden anzugeben. Auch hier ist im Bereich der sonstigen psychischen Beschwerden zumindest eine starke Signifikanz zu erkennen ($p=0,01$).

3.1.1.2 WMT

Welche Beschwerden geben diejenigen an, die im WMT eine mangelnde Kooperation zeigten bzw. ehrlich geantwortet haben?



Prinzipiell geben diejenigen, die im WMT eine mangelnde Kooperation zeigten mehr Beschwerden an, als diejenigen, die ihren Test ehrlich bearbeitet haben. Ganz besonders wird dies im Bereich der sonstigen psychischen Beschwerden deutlich, bei denen wir einen hochsignifikanten Unterschied erkennen ($p=0,001$). Aber auch im Bereich der HWS-Beschwerden erkennt man mit $p=0,012$ einen signifikanten Unterschied. Die Depressionen gehen mit $p=0,072$ nur knapp an einer Signifikanz vorbei (s. Tabelle 12).

Tabelle 12

Signifikante Unterschiede

	Signifikanz p
Psychische Beschwerden	0,001
HWS-Beschwerden	0,012
Depression	0,072

3.1.2 *Passende Beschwerdenschilderung?*

Bei den bisherigen Ergebnissen taucht selbstverständlich die Frage auf, ob denn die geschilderten Beschwerden der Patienten zu den Diagnosen passen? Aus diesem Grund filtern wir im nächsten Schritt alle Personen raus, die eine bestimmte Beschwerde angeben und eruieren, wie viele von denen tatsächlich die entsprechende Störung diagnostiziert bekommen haben und wie sich das zu den entsprechenden Scores verhält.

3.1.2.1 *SFSS*

In den folgenden Tabellen werden diejenigen Personen aufgeführt, die die jeweilige Beschwerde angegeben haben (sonstige psychische Beschwerden, medizinische Beschwerden etc.) und mit den tatsächlichen Diagnosen verglichen (s. Tabelle 13).

74,1% der Personen, die vom SFSS als Simulanten angenommen wurden, haben entgegen ihrer Angaben gar keine diagnostizierbaren sonstigen psychischen Probleme. Während bei den „Ehrlichen“ nur 51,3% entgegen ihren Angaben keine diagnostizierbaren sonstigen psychischen Probleme haben (s. Tabelle 13).

44% der Personen, die vom SFSS als Simulanten angenommen wurden, haben entgegen ihrer Angaben gar keine diagnostizierbaren medizinischen Erkrankungen. Während bei den „Ehrlichen“ nur 33,3% entgegen ihren Angaben keine Erkrankungen medizinischen Probleme haben.

Tabelle 13

Zusammenhang zwischen Beschwerden und Diagnosen in %

	SFSS bestanden	SFSS nicht bestanden
Sonstige psych. Diagnose	51,3	74,1
Sonstige psych. Diagnose	48,7	25,9
Med. Diagnose pos.	33,3	44
Med. Diagnose neg.	66,7	56

3.1.2.2 WMT

In den folgenden Tabellen werden diejenigen Personen aufgeführt, die die jeweilige Beschwerde angegeben haben (sonstige psychische Beschwerden, medizinische Beschwerden etc.) und mit den tatsächlichen Diagnosen verglichen.

66,7% der Personen, die vom WMT als unkooperativ angenommen wurden, haben entgegen ihrer Angaben gar keine diagnostizierbaren sonstigen psychischen Probleme. Während bei den „Ehrlichen“ nur 55,6% entgegen ihren Angaben keine diagnostizierbaren sonstigen psychischen Probleme haben (s. Tabelle 14).

50% der Personen, die vom WMT als unkooperativ angenommen wurden, haben entgegen ihrer Angaben gar keine diagnostizierbaren medizinischen Probleme. Während bei den „Ehrlichen“ nur 29,4% entgegen ihren Angaben keine diagnostizierbaren medizinischen Probleme haben (s. Tabelle 14).

Tabelle 14

Zusammenhang zwischen Beschwerden und Diagnosen in %

	WMT bestanden	WMT nicht bestanden
Sonstige psych. Diagnose	55,6	66,7
Sonstige psych. Diagnose	44,4	33,3
Med. Diagnose pos.	29,4	50
Med. Diagnose neg.	70,4	50

3.1.2.3 VEI-Validitätsskala Ungünstiger Eindruck

In den folgenden Tabellen werden diejenigen Personen aufgeführt, die nach dem VEI einen ungünstigen Eindruck erwecken wollten und die jeweilige Beschwerde angegeben haben (sonstige psychische Beschwerden, medizinische Beschwerden etc.). Diese Angaben werden dann mit den tatsächlichen Diagnosen verglichen.

75% der Personen, die nach dem VEI einen ungünstigen Eindruck erwecken wollten, haben entgegen ihrer Angaben gar keine diagnostizierbaren sonstigen psychischen Probleme. Während bei den „Ehrlichen“ nur 58,6% entgegen ihren Angaben keine diagnostizierbaren sonstigen psychischen Probleme vorliegen (s. Tabelle 15).

57,1% der Personen, die nach dem VEI einen ungünstigen Eindruck erwecken wollten, haben entgegen ihrer Angaben gar keine medizinischen Probleme. Während bei den „Ehrlichen“ nur 34,7% entgegen ihren Angaben keine Probleme haben. Beim RDF sieht es ähnlich aus. 83,3% der Personen, die im RDF auffällig waren, haben entgegen ihrer Aussage gar keine medizinischen Probleme. Während bei den „Ehrlichen“ nur 58,3% entgegen ihren Angaben keine Probleme haben. Der MAL kann in diesem Bereich keine Aussage machen, da niemand auffällig wurde (s. Tabelle 15).

Tabelle 15

Zusammenhang zwischen Beschwerden und Diagnosen in %

	VEI-U ↑	VEI-U	MAL* ↑	MAL*	RDF* ↑	RDF*
psych. Diagnose pos.	25	41,4	---	39,4	16,7	41,7
psych. Diagnose neg.	75	58,6	---	60,6	83,3	58,3
Med. Diagnose pos.	42,9	65,3	---	63,3	28,6	66,7
Med. Diagnose neg.	57,1	34,7	---	36,7	71,4	33,3

*T-Werte mit Grenzwerten nach Groves 2009

3.1.2.4 Zusammenfassung

Zusammenfassend kann man sagen, dass bei allen drei Antwortverzerrungsparametern, diejenigen, die eine Beschwerde im medizinischen oder sonstigen psychischen Bereich angegeben haben und als negative Antwortverzerrer eingestuft worden sind, tendenziell häufiger entgegen ihren Angaben in den Bereichen gar keine Probleme haben bzw. diagnostisch das Gegenteil festgestellt wurde, im Vergleich zu denjenigen, die nicht als negative Antwortverzerrer eingestuft wurden.

3.1.3 VEI-Beschwerden bei SFSS und WMT

Nun wird geschaut, welche Beschwerden durch den VEI erkennbar bzw. vom Patienten so dargestellt werden. Es wird die Frage geklärt inwieweit die Beschwerden im VEI mit dem jeweiligen Score korrelieren.

3.1.3.1 SFSS

Das Ausmaß der Beschwerdenschilderung ist erwartungsgemäß bei denjenigen, die im SFSS als Personen, die ungläubwürdige Beschwerden schildern, identifiziert worden sind, noch viel größer als bei denjenigen, die sich im WMT als unkooperativ zeigten. Wir finden in den Skalen somatische Beschwerden, Angst, angstverbundene Störungen, Depression, paranoia, Schizophrenie, Borderline-Merkmale, Suizidalität, Stress und Unterstützungsdefizit eine höchst signifikante Korrelation zu Personen, die nach dem SFSS ungläubliche Beschwerden schilderten. In den Skalen Paranoia, Unterstützungsdefizit, Behandlungsablehnung und Wärme finden wir einen signifikanten Zusammenhang (s. Tabelle 16).

3.1.3.2 WMT

Die Korrelation zwischen Angaben zu Borderline-Merkmalen und Unterstützungsdefizit und ist signifikant zu mangelnder Kooperation im WMT. Dabei ist die Korrelation zu somatischen Beschwerden, Angst, Depression, Schizophrenie und Suizidalität sogar höchst signifikant (s. Tabelle 16).

Tabelle 16

Korrelationen zwischen den VEI-Skalen und den SFSS Ergebnissen

	SFSS	WMT
Somatische Beschwerden	0,442**	0,283**
Angst	0,417**	0,268**
Angstverbundene Störungen	0,343**	0,186
Depressionen	0,448**	0,333**
Manie	0,185	-0,51
Paranoia	0,299**	0,156
Schizophrenie	0,478**	0,279**
Borderline-Merkmale	0,368**	0,201*
Antisoz. verhaltensmerkmale	0,176	-0,14
Alkoholprobleme	-0,081	-0,070
Drogenprobleme	0,056	0,146
Aggression	0,116	0,086
Suizidalität	0,400**	0,282**
Stress	0,365**	0,147
Unterstützungsdefizit	0,390**	0,219*
Behandlungsablehnung	-0,335**	-0,187
Dominanz	-0,060	-0,160
Wärme	-0,282**	-0,179

** Korrelation ist hoch signifikant

* Korrelation ist signifikant

3.1.3.3 Zusammenfassung

Sowohl Patienten, die im SFSS als Personen eingestuft wurden, die unglaubwürdige Beschwerden simulieren, als auch Patienten, die im WMT als unkooperativ eingestuft wurden, zeigen eine bis zu höchst signifikante Korrelation zu erhöhten Beschwerdenschilderungen im Vergleich zu Patienten, die im SFSS oder WMT ehrlich geantwortet haben. Insbesondere findet man das in den Skalen somatische Beschwerden, Angst, Depression, Schizophrenie, Borderline-Merkmale und Suizidalität.

3.2 Stärke der VEI-Validitätsskalen bei Antwortverzerrern

Wie stark ist der Test „Der Verhaltens- und Erlebensinventar“ (VEI), die deutsche Fassung des Personality Assessment Inventory, innerhalb der Diagnostik der Beschwerdvalidität? Kann er signifikant valide Aussagen zu Antwortverzerrungen machen? Gibt es Skalen, die ähnlich hohe Trefferquoten aufweisen wie der WMT oder SFSS?

3.2.1 Korrelation zwischen VEI Validitätsskalen und SFSS / WMT

3.2.1.1 Korrelationen zwischen VEI Validitätsskalen und SFSS

Tabelle 17

Korrelationen zwischen den VEI-Validitätsskalen und den SFSS Ergebnissen

	SFSS
Inkonsistenz	0,109
Seltenheit	0,108
Ungünstiger Eindruck	0,668**
MAL	0,228*
RDF	0,124

** Korrelation ist hoch signifikant

* Korrelation ist signifikant

Wir haben hier eine hochsignifikante Korrelation. Je höher der Wert im SFSS ist, desto höher ist auch der Wert des ungünstigen Eindrucks. D.h. Personen, bei denen der SFSS ungläublichen Beschwerden angezeigt hat, bei denen zeigt der VEI auch an, dass sie einen ungünstigen Eindruck erwecken wollten.

Ebenso haben wir eine signifikante Korrelation zwischen dem SFSS und der VEI-Validitätsskala Malingering Index.

3.2.1.2 Korrelationen zwischen VEI Validitätsskalen und WMT

Tabelle 18

Korrelationen zwischen den VEI-Validitätsskalen und den WMT Ergebnissen

	WMT: IRV	WMT: DRV	WMT: CV
Inkonsistenz	-0,009	0,53	0,14
Seltenheit	-0,286**	-0,176	-0,262**
Ungünstiger Eindruck	-0,314**	-0,337**	-0,344**
MAL	-0,129	-0,207*	-0,134
RDF	-0,057	-0,040	-0,071

** Korrelation ist hoch signifikant

* Korrelation ist signifikant

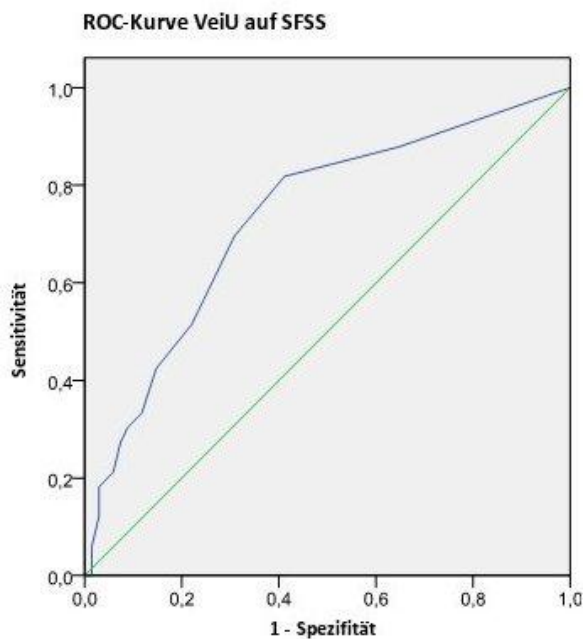
Auch hier finden wir wieder hochsignifikante Korrelationen. Je höher die WMT scores (IRV = Immediate Recognition Variable, DRV = Delayed Recognition Variable, CV = Consistency Variable) werden (mit der Höhe nimmt die Wahrscheinlichkeit zu, dass beim Test kooperativ geantwortet haben), einen desto weniger ungünstigen Eindruck wollen die Probanden machen. Je niedriger der Wert wird (die Wahrscheinlichkeit, dass sie den Test unkooperativ bearbeitet haben steigt), desto mehr versuchen sie einen ungünstigen Eindruck zu hinterlassen.

Zwei weitere Werte sind hier auffällig: Bei der Immediate Recognition Variable und Consistency Variable steigt die Seltenheit hoch-signifikant (d.h. die Testausfüllung beim VEI wird immer zufälliger) bzw. Delayed Recognition Variable steigt der Malingering Index signifikant, wenn die WMT scores sinken (d.h. wenn die Wahrscheinlichkeit steigt, dass der Test unkooperativ bearbeitet wurde) desto größer ist die Wahrscheinlichkeit, dass der Proband den Fragebogen zumindest teilw. zufällig bzw. negativ verzerrt ausfüllt.

3.2.2 VEI-Validitätsskalen und SFSS bzw. WMT

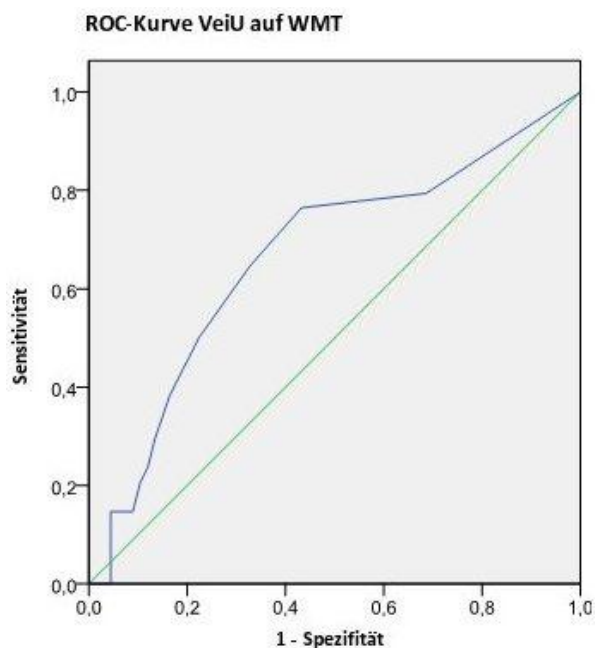
Es gibt signifikante bis hochsignifikante Zusammenhänge zwischen den VEI (Validitätsskalen) und den Scores des WMT und SFSS. Lassen sich diese Zusammenhänge auch anhand der ROC-Kurven bestätigen?

3.2.2.1 SFSS und VEI-Ungünstiger Eindruck



Da die AUC bei 0,733 liegt, gibt es einen starken Zusammenhang zwischen dem SFSS und dem VEI-Score Ungünstiger Eindruck. Das bestätigt auch die Signifikanz, die bei 0,000 liegt.

3.2.2.2 WMT und VEI-Ungünstiger Eindruck



Beim WMT ist die AUC mit 0,665 unter 0,7. Somit ist ein Zusammenhang da, der trotz der Signifikanz von 0,007 nicht so stark ausgeprägt ist wie beim SFSS.

3.2.2.3 Weitere Skalen des VEI und Korrelationen?

Die VEI-Skalen Inkonsistenz, Seltenheit, Malingering Index und Rogers Discriminant Function zeigen alle keine Korrelationen in der ROC-Kurve mit dem SFSS oder WMT. Ihre AUC liegt bei allen zwischen 0,5 und 0,6. Nur die Korrelation der Skala Seltenheit liegt mit dem WMT über AUC 0,6 und ihrer Signifikanz bei 0,015. Also ist auch hier ein Zusammenhang erkennbar (s. Tabellen 18 und 19).

Tabelle 18

ROC-Kurven Ergebnisse: VEI-Validitätsskalen auf SFSS

	AUC	Signifikanz
Veil	0,547	0,445
Veis	0,559	0,339
T_MAL	0,540	0,513
T_RDF	0,484	0,8

Tabelle 19

ROC-Kurven Ergebnisse: VEI-Validitätsskalen auf WMT

	AUC	Signifikanz
Veil	0,522	0,717
Veis	0,648	0,015
T_MAL	0,577	0,205
T_RDF	0,540	0,513

4 Diskussion

4.1 Interpretation

4.1.1 Beschwerdemuster von Antwortverzerrern

Patienten, die im SFSS als Antwortverzerrer eingestuft worden sind, geben hochsignifikant mehr somatische Beschwerden an als diejenigen, die im SFSS nicht als Antwortverzerrer eingestuft worden sind.

Patienten, die im WMT als Antwortverzerrer eingestuft worden sind, geben hochsignifikant mehr sonstige psychische Beschwerden und signifikant mehr HWS-Beschwerden an als diejenigen, die im WMT nicht als Antwortverzerrer eingestuft worden sind.

In allen drei Antwortverzerrungsparametern (SFSS, WMT, VEI-Ungünstiger Eindruck) haben diejenigen, die eine Beschwerde im medizinischen oder sonstigen psychischen Bereich angegeben haben und als negative Antwortverzerrer eingestuft worden sind, tendenziell häufiger entgegen ihren Angaben in den Bereichen gar keine Probleme bzw. ist diagnostisch auf Befundebene keine entsprechende Störung festzustellen.

Sowohl Patienten, die im SFSS als Personen eingestuft wurden, die unglaubwürdige Beschwerden simulieren, als auch Patienten, die im WMT als unkooperativ eingestuft wurden, zeigen eine bis zu höchst signifikante Korrelation zu erhöhten Beschwerdenschilderungen im Vergleich zu Patienten, die im SFSS oder WMT ehrlich geantwortet haben. Insbesondere finden man das in den Skalen somatische Beschwerden, Angst, Depression, Schizophrenie, Borderline-Merkmale und Suizidalität.

4.1.2 Der VEI als Test zur Simulationsdiagnostik

Die VEI-Validitätsskala „Ungünstiger Eindruck“ ist die einzige wirklich starke Skala im Bezug auf Antwortverzerrung. Sie zeigt eine hochsignifikante Korrelation zum SFSS und zum WMT. Und auch die AUC liegt beim SFSS oberhalb von 0,7 (0,733) und beim WMT nur knapp unterhalb von 0,7 (0,665).

4.2 Vergleich

Zu diesem Thema gibt es nicht allzu viele Studien. Aus diesem Grund fällt der Vergleich recht schwer. Trotzdem gibt es einige wenige Studien, zu denen man ein Vergleich ziehen kann.

4.2.1 Die Beschwerdemuster

Es ist nicht verwunderlich, dass Personen, die den SFSS nicht bestehen, mehr Beschwerden angeben, denn der Test will ja genau diese Personen herausfiltern. Es ist auch nicht unüblich, dass der Anteil solcher Personen bei diesem Hintergrund recht hoch sind (Merten et al. 2007).

Bei Personen, die durch den WMT als unkooperativ eingestuft wurden, wird ebenfalls simuliert, allerdings nicht so ausgeprägt wie bei den Personen, die den SFSS nicht bestanden haben.

Über das genaue Muster der Beschwerdeschilderung ist allerdings wenig bekannt. Primär werden laut dieser Studie eher psychische oder psychosomatische Beschwerden simuliert, als somatische Beschwerden. Vermutlich wird dies aus dem Motiv heraus getan, dass somatische Beschwerden wesentlich leichter als psychische oder psychosomatische Beschwerden überprüft werden können.

4.2.2 Die VEI Validitätsskalen

Laut Sellbom ist die VEI-Validitätsskala „ungünstiger Eindruck“ die wohl stärkste Skala in dem Erkennen von Antwortverzerrung. MAL und RDF haben zwar auch ihren Stellenwert, reichen an die VEI-Validitätsskala „ungünstiger Eindruck“ jedoch nicht heran (Sellbom & Bagby 2008)

Hawes sieht die Validitätsskala MAL ein wenig besser und schätzt sie ähnlich stark wie die die Validitätsskala „ungünstiger Eindruck“. Die Validitätsskala RDF nimmt er ebenfalls als weniger effektiv an (Hawes 2009).

Beide Studien gehen in eine ähnliche Richtung und unterstreichen die Ergebnisse dieser Studie.

4.3 Schwachpunkte

Die Stärke der Studie ist zugleich auch ein Schwachpunkt der Studie. Die Probanden sind keine Studenten oder sonstige Freiwillige aus bestimmten Bereichen, sondern Personen, die ein Motiv haben Antworten zu verzerren. Das bringt den klaren Vorteil, dass wir realistische Antwortverzerrung beobachten können. Es hat aber auch den Nachteil, dass wir nicht zu 100% wissen, ob tatsächlich die Antworten verzerrt worden sind, sondern nur auf Methoden zurückgreifen können, die mit einer hohen Wahrscheinlichkeit Antwortverzerrung entlarven (SFSS und WMT).

5 Zusammenfassung

5.1 Hintergrund

Der ambulante und stationäre Alltag konfrontiert den Arzt immer wieder mit der Frage, ob der gerade zu behandelnde oder zu begutachtende Patient eine behandlungswürdige Erkrankung hat oder nicht. Im Normalfall kommt der Patient, wenn er tatsächlich Beschwerden hat. In diesem Fall braucht der Arzt nur nach der Ursache zu suchen. Aber es gibt auch Situationen, in denen der Patient einen Vorteil davon hat, seine Symptome überzogener darzustellen als sie sind oder aber Beschwerden komplett zu simulieren. Unter dieser motivierenden Voraussetzung geht es für den Patienten entweder um Entschädigung, bezahlten Krankenurlaub oder ähnliches. Verlässliches ärztliches Handeln steht hier unter der erschwerenden Bedingung, einschätzen zu können, ob hier ein objektiver Befund vorliegt oder ob das Eigeninteresse einen erhöhten Simulationsgrad produziert.

Bei unserer Studie hatten wir eine große Gruppe von Patienten bei denen ein Anreiz für Simulation bestand, da sie genau die oben genannten Vorteile beanspruchen wollten.

5.2 Methoden

Wir haben die ausgewählten Personen verschiedene Tests machen lassen: SFSS, WMT und VEI. Mit Hilfe dieser Tests sollte herausgefunden werden, wer simuliert (SFSS) oder mangelnd kooperiert (WMT), ob der VEI das auch erkennt und welche Beschwerden die Patienten bei Simulation angeben.

5.3 Resultate

Patienten, die im SFSS als Antwortverzerrer eingestuft worden sind, und Patienten, die im WMT als unkooperativ eingestuft wurden, geben hochsignifikant mehr Beschwerden an, haben tendenziell häufiger entgegen ihren Angaben in den Bereichen gar keine tatsächliche Störung und zeigen eine bis zu höchst signifikante Korrelation zu erhöhten Beschwerdenschilderungen.

Die VEI-Validitätsskala „Ungünstiger Eindruck“ ist die einzige starke Skala im Bezug auf Antwortverzerrung. Sie zeigt eine hochsignifikante Korrelation zum SFSS und zum WMT.

5.4 Konklusion

Aus der Studie lassen sich zwei Schlussfolgerungen ableiten:

Anhand der Beschwerdenschilderung kann man nicht erkennen, ob jemand definitiv die Antworten verzerrt oder Beschwerden simuliert. Aber es ist möglich das Beschwerdemuster als Hinweis zu gebrauchen. Sollte also in der Beschwerdvalidierung eine Person auftauchen, die ein ähnliches Beschwerdemuster aufzeigt, wie das in dieser Studie, so kann es sinnvoll sein weitere Tests zu veranlassen, um in einem Gesamtbild verschiedener Tests eine höhere Aussagekraft erzielen zu können. Innerhalb der Beschwerdvalidierung ist es mittlerweile auch gefordert mehrere Tests zu verwenden und ihrer Kombination eine Trefferquote zu erzielen, die Fehler kaum noch zulässt (Merten et al. 2007).

Da die VEI-Validitätsskala „Ungünstiger Eindruck“ eine so hoch signifikante Korrelation zum SFSS und WMT hat und diese schon als sehr gute Tests anerkannt sind, kann diese Skala durchaus auch als Erkennungsskala von Simulation gebraucht werden.

Literaturverzeichnis

Bagby R. M., Nicholson R. A., Bacchiochi J. R., Ryder A. G., Bury A. S. (2002):

The predictive capacity of the MMPI-2 and PAI validity scales and indexes to detect coached and uncoached feigning.

Journal of Personality Assessment, 78, S. 69-86

Benbadis, S. R. (2005):

Psychogenic symptoms: is the psychiatric community in denial?

In: Epilepsy & Behavior, 6, S. 9–14

Binder L. M., Rohling M. L. (1996):

Money matters: a meta-analytic review of the effects of financial incentives on recovery after closed-head injury

The American Journal of Psychiatry, 153, S. 7-10

Blanchard D. D., McGrath R. E., Pogge D. L., Khadivi A. (2003):

A comparison of the PAI and MMPI-2 as predictors of faking bad in college students.

Journal of Personality Assessment, 80, S. 197-205

Blaskewitz M., Merten, T. (2007):

Diagnostik der Beschwerdendvalidität

Fortschritte der Neurologie & Psychiatrie, 75, S. 140-154

Blaskewitz M. (2005):

Diagnostik der Beschwerdendvalidität, Diplomarbeit, S. 73

Boyle G. J., Lennon T. J. (1994):

Examination of the Reliability and Validity of the Personality Assessment Inventory

Journal of Psychopathology and Behavioral Assessment, Vol. 16, Nr. 3, S. 174-176

Brockhaus R., Merten T. (2004):

Neuropsychologische Diagnostik suboptimalen Leistungsverhaltens mit dem Word Memory Test.

In: Der Nervenarzt, 75

Springer Verlag, Berlin, S. 882-887

Bush S. S., Ruff R. M., Tröster A. I. et al. (2005):

Symptom validity assessment: practice issues and medical necessity.

NAN Policy and Planning Committee. Archives of Clinical Neuropsychology, 20, S. 419–426

Butcher J. N., Dahlstrom W. G., Graham J. R., Tellegen A., Kaemmer B. (1989):

Manual for the restandardized Minnesota Multiphasic Personality Inventory: MMPI-2. An interpretative and administrative guide.
University of Minnesota Press, Minneapolis

Cima M. et al. (2003):

Strukturierter Fragebogen Simulierter Symptome
In: Der Nervenarzt, 11
Springer Verlag, Berlin, S. 977

Duellman R.M., Bowers T.G. (2004):

Use of the Personality Assessment Inventory (PAI) in forensic and correctional settings: evidence for concurrent validity.
International Journal of Forensic Psychology, 1, S. 42-57

Cochrane H. J., Baker G.A., Meudell P.R. (1998):

Simulating a memory impairment: can amnestics implicitly outperform simulators?
British Journal of Clinical Psychology, 37, S. 31–48

Edens J.F., (2007):

Detection of Malingering in Psychiatric Unit and General Population Prison Inmates: A Comparison of the PAI, SIMS, and SIRS
Journal of Personality Assessment, 88(1), S. 39

Goldberg H.E., Back-Madruga C., & Boone K.B. (2007)

The Impact of Psychiatric Disorders on Cognitive Symptom Validity Test Scores.
In: K.B. Boone: Assessment of Feigned Cognitive Impairment. A Neuropsychological Perspective.
Guilford, New York, S. 281-309

Green P., (2005):

Green's Word Memory Test User's Manual, Edmonton, Green's Publishing.

Green P., Merten T., Brockhaus R. (2004):

Reportpsychologie, 5, S. 29

Green P., Lees-Haley P.R., Allen L. M. (2002):

The Word Memory Test and the validity of neuropsychological test scores.
Journal of Forensic Neuropsychology, 2, S. 97-124

Groves J. A. (2009):

Untersuchungen zur Konstruktvalidität des Verhaltens- und Erlebensinventars (VEI) an einer klinischen Stichprobe.
Dissertation, München, S. 108

Grow R., McVaugh W., Eno T. D. (1980):

Faking and the MMPI.

Journal of Clinical Psychology, 36, S. 910–917

Hawes S., Boccaccini M.T. (2009):

Detection of Overreporting of Psychopathology

on the Personality Assessment Inventory: A Meta-Analytic Review

Psychological Assessment, Vol. 21, No. 1, S.120ff

Lynch W. J. (2004):

Determination of effort level, exaggeration, and malingering in neurocognitive assessment.

The Journal of Head Trauma Rehabilitation, 19, S. 277-283

Martelli M. F., Zasler N. D., Nicholson K., Hart R. P., Heilbronner R. L. (2002):

Masquerades of brain injury part III: Critical evaluation of symptom validity testing and diagnostic realities in assessment.

Journal of Controversal Medical Claims, 9, S. 19-22

Merten T., Dettenborn H. (2009):

Diagnostik der Beschwerdenuvalidität

Deutscher Psychologen Verlag, Berlin, S. 7

Merten T., Friedel E., Stevens A. (2007):

Die Authentizität der Beschwerdenschilderung

in der neurologisch-psychiatrischen Begutachtung

Praxis der Rechtspsychologie 17 (1), März, S. 150

Merten T. et al. (2007):

Über die Validität von Persönlichkeitsprofilen in der nervenärztlichen Begutachtung

In: Der Nervenarzt, 5

Springer Verlag, Berlin, S. 511-512

Merten T., Friedel E., Stevens A. (2006):

Eingeschränkte Kooperativität in der neurologisch-psychiatrischen Begutachtung: Schätzungen zur Auftretenshäufigkeit an einer Begutachtungspopulation.

Versicherungsmedizin 58, S.19–21

Merten T., Henry M., Hilsabeck R. (2004):

Symptomvalidierungstests in der neuropsychologischen Diagnostik: eine Analogstudie.

Zeitschrift für Neuropsychologie, 15, S. 81-90

Mittenberg W., Patton C., Canyock et al. (2002):

Base rates of malingering and symptom exaggeration.

Journal of Clinical and Experimental Psychology, 24, S. 1094–1102

Morey L.C., (1991):

Personality Assessment Inventory (PAI)

Psychological Assessment Resources

Morey L.C., (2007):

Personality Assessment Inventory (PAI)

Psychological Assessment Resources

Resnick P. J. (1997):

Malingered Psychosis

In: R. Rogers: Clinical assessment of malingering and deception.

Guilford, New York, S. 47–67

Resnick P. J. (1984):

The detection of malingered mental illness.

Behavioral Sciences & the Law, 2, S. 20–38

Rogers R., Hinds J. D., Sewell K. W. (1996):

Feigning psychopathology among adolescent offenders: Validation of the SIRS, MMPI-A, and SIMS.

Journal of Personality Assessment, 67, S. 244–257

Rogers R., Sewell K. W., Goldstein A. M. (1994):

Explanatory models of malingering: A prototypical analysis.

Law and Human Behavior, 18, S. 543–552

Rogers R., Bagby R. M., Dickens S. E. (1992):

Structured Interview of Reported symptoms (SIRS) and professional manual.

Psychological Assessment Resources, Odessa, FL

Rogers R. (1984):

Towards an empirical model of malingering and deception.

Behavioral Science and the Law, 2, S. 93–112

Schretlen D. J. (1988):

The use of psychological tests to identify malingered symptoms of mental disorder.

Clinical Psychology Review, 8, S. 451–476

Seamons D. T., Howell R. J., Carlisle A. L., Roe A.V. (1981):

Rorschach simulation of mental illness and normality by psychotic and nonpsychotic legal offenders.

Journal of Personality Assessment, 45, S. 130–135

Sellbom M., Bagby R.M. (2008):

In: R. Rogers: Response Styles on Multiscale Inventories
The Guilford Press, New York, S. 193ff

Sellbom M. et al. (2008):

Response Styles on Multiscale Inventories
In: R. Rogers: Clinical assessment of malingering and deception.
The Guilford Press, New York, S. 189

Slick D. J., Sherman E. M. S., Iverson G. L. (1999):

Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research.
The Clinical Neuropsychologist, 13, S. 545-61

Smith, G. P. (1997):

Assessment of malingering with selfreport instruments.
In: R. Rogers: Clinical assessment of malingering and deception.
Guilford, New York, S. 351–370

Smith G.P. , Burger G. K. (1997):

Detection of Malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS).
Journal of American Academy Psychiatry and the Law, 25, S. 183–189

Stieglitz R. D. (2005):

Möglichkeiten und Grenzen der Selbstbeurteilung in der Begutachtung.
In: Der Nervenarzt, 76 Supplement 1
Springer Verlag, Berlin, S. 286

Tan J. E., Slick D. J., Strauss E., Hultsch D.F (2002):

How'd they do it? Malingering strategies on symptom validity tests.
The Clinical Neuropsychologist, 16, S. 495-505

Van Gorp W.G., Humphrey L. A., Kalechstein A. L., Brumm V. L., McMullen W. J., Stoddard M. A., Pachana N. A. (1999):

How well do standard clinical neuropsychological tests identify malingering? A preliminary analysis.
Journal of Clinical and Experimental Neuropsychology, 21 (2), S. 245-250

White J. (1993):

The Personality Assessment Inventory in a forensic setting.
Paper presented at the 28th Annual Convention of the American Psychological Association, Los Angeles.

Danksagung

Ich möchte allen, die mir die Arbeit an meiner Dissertation ermöglichten und erleichterten, ganz herzlich danken:

- Tatjana Tscheuschner, meine Ehefrau, die mich immer unterstützt hat und auch Zeitmangel in Kauf genommen hat. Sie hat mich immer wieder motiviert und aufgemuntert.
- Meine Eltern Willi und Marion Tscheuschner, die mir das Studium und damit auch die Doktorarbeit durch ihre finanzielle Unterstützung überhaupt erst möglich gemacht haben
- Prof. Dr. Stevens, der mir diese interessante Arbeit möglich gemacht hat und Fragen stets beantwortet hat
- Lisa Thies, die mir immer wieder mit Rat zur Seite gestanden ist

Lebenslauf

1998-1992	Grundschule Suttrop
1992-2001	Gymnasium Warstein
2001-2002	Zivildienst beim DRK Rettungsdienst in Ludwigsburg (Ausbildung zum Rettungssanitäter)
2002-2005	Studium der Theologie am Seminar Schloss Bogenhofen in Braunau (Österreich)
2005-2011	Studium der Medizin an der Eberhard Karls Universität Tübingen
4-6.10.2011	Schriftliche Prüfung des 2. Abschnittes der Ärztlichen Prüfung
8-9.11.2011	Mündliche Prüfung des 2. Abschnittes der Ärztlichen Prüfung