

6. Die Validierung des Computergestützten Adaptiven Tests zur Angstmessung (Angst-CAT)

6.1. Einleitung

Zur Beurteilung der psychometrischen Güte eines Tests ist nach der Testkonstruktion die Validierung des entwickelten Instruments unabdingbar. Die vorliegende empirische Studie widmet sich der Validierung des Angst-CATs, dessen Testentwicklung im vorangegangenen Kapitel 5 beschrieben wurde.

Unter Validierung wird die Überprüfung der Validität eines Tests verstanden (Lienert & Raatz, 1994). Die Validität - nach Bortz und Döring (1995) das wichtigste Testgütekriterien überhaupt – gibt an, „wie gut ein Test in der Lage ist, genau das zu messen, was er zu messen vorgibt“ (S.185). Um die Validität eines Tests zu bestimmen, existieren verschiedene Validitätsansätze, welche unterschiedliche Untersuchungsmethoden erfordern (Cronbach, 1990).

Die Ziele der hier dargestellten Validierungsstudie und die zur Zielerreichung genutzten Validitätsansätze werden in Kapitel 6.2., die untersuchten Hypothesen in Kapitel 6.3 expliziert, gefolgt von einer Beschreibung der untersuchten *Stichprobe* (Kapitel 6.4.) und der an ihr erhobenen *Instrumente* (Kapitel 6.5.). Anschließend werden die zur Untersuchung der verschiedenen Validitätsansätze genutzten statistischen *Methoden* dargestellt (Kapitel 6.6.) und die *Ergebnisse* zusammengefasst und erörtert (Kapitel 6.7.).

6.2. Ziele

Seit Beginn der Validierungsforschung (Anfang des 20. Jahrhunderts durch Spearman, 1904) spielt die *Konstruktvalidität* eine dominierende Rolle. Kennzeichnend für diese Art der Validität ist die Erhebung von Konstrukten (z. B. mittels psychometrischer Instrumente) und deren Beziehung zum Testwert des zu validierenden Instruments (hier: das Angst-CAT). Unter der Voraussetzung, dass die ausgewählten und erfassten Konstrukte selbst repräsentativ, reliabel, valide und für die Validierung adäquat sind, können durch die empirische Untersuchung dieser Zusammenhänge Rückschlüsse auf die Gültigkeit des untersuchten Tests gezogen werden.

Das Ziel vorliegender Studie ist die Bestimmung der *Konstruktvalidität* im Sinne einer *Übereinstimmungsvalidität* (*konkurrente Validität*; Lienert & Raatz, 1994, S. 224) von Variablen, von denen aufgrund von theoretischen und empirischen Forschungsbefunden erwartet wird, dass sie in unterschiedlicher Konstruktnähe

zum Angst-CAT positioniert werden können, und welche praktisch zeitgleich – jedoch unabhängig voneinander - mit dem Angst-CAT erhoben werden.

Von Belang ist hierbei nicht nur die Überprüfung, ob mehrere Methoden (psychometrische Tests, Interviews etc.), mit einem ähnlichen Messbereich (Erfassung von Angst), jedoch mit unterschiedlichen Operationalisierungen dieses Messbereichs zu *ähnlichen* Messergebnissen kommen (*konvergente Validität*), sondern auch, ob Hinweise auf *gering* ausgeprägte Zusammenhänge zwischen Tests, welche die Erfassung differierender Konstrukte intendieren, eruierbar sind, so dass Rückschlüsse auf die Fähigkeit des Angst-CATs zur Diskrimination zwischen unterschiedlichen Konstrukten möglich sind (*divergente bzw. diskriminante Validität*; Campbell & Fiske, 1959).

6.3. Hypothesen

In der vorliegenden Studie wurden zur Überprüfung der Validität des Angst-CATs neben dem zu validierenden Angst-CAT verschiedene psychometrische Inventare zur Angst- und Depressionserfassung sowie zur Messung verschiedener Persönlichkeitskonstrukte und ein strukturiertes diagnostisches Interview zwischen den Jahren 2002 und 2003 an Patienten der Medizinischen Klinik mit Schwerpunkt Psychosomatik der Charité Berlin angewandt.

Es wird im Sinne einer guten konvergenten Validität erwartet, dass das Angst-CAT mit den erhobenen Angstinventaren in einem engen Zusammenhang steht (hoch korreliert), sowie Patienten mit der Diagnose einer Angststörung im Angst-CAT höhere Werte erzielen als Patienten ohne eine psychische Störung. Weiterhin wird erwartet, dass sich eine gute divergente Validität des Angst-CATs in Form einer hohen Diskrimination zu anderen Eigenschaftskonstrukten, welche mit den Persönlichkeitsinventaren erfasst werden, und in Form einer Diskrimination zwischen verschiedenen Diagnosegruppen ausdrückt.

Angeichts einer Fülle von theoretischen Forschungsdiskursen (siehe Kapitel 2.5) und einer ausgeprägten empirischen Befundlage (siehe Kapitel 6.5.1.), die darauf hinweist, dass sich die psychometrische Diskrimination zwischen den Konstrukten Angst und Depression bzw. Neurotizismus schwierig gestaltet, wird vermutet, dass eine solche Diskriminationsleistung auch mit dem Angst-CAT nicht hinreichend gelingt.

6.4. Stichprobe

Die Validierungsstichprobe umfasst insgesamt N = 102 Patienten, die in der Medizinischen Klinik mit Schwerpunkt Psychosomatik der Charité Berlin zur Diagnostik oder Therapie in den Jahren 2002 bis 2003 stationär behandelt wurden. Tabelle 19 fasst die wesentlichsten soziodemografischen und klinischen Charakteristika der Stichprobe zusammen.

Tabelle 19: Soziodemografische und klinische Charakteristika der Validierungsstichprobe.

Charakteristika	Kategorie / Parameter	Angaben
Geschlecht	Weiblich	79,4%
	Männlich	20,6%
Alter	Arithmetischer Mittelwert (\bar{X})	42,28 Jahre
	Standardabweichung (SD)	15,53 Jahre
	Altersspanne	18-77 Jahre
Familienstand	verheiratet	45,1%
	ledig (mit Partner)	14,7%
	ledig (ohne Partner)	25,5%
	geschieden / Verwitwet	14,7%
Diagnosen⁸¹	Angststörungen (F.40-41)	56,8%
	Depressive Störungen (F.32-34)	58,8%
	Somatoforme Störungen (F.45)	50,0%
	Essstörungen (F.50)	6,9%
	Primär somatische Erkrankungen (nicht F)	9,8%

Leider war es bisher nicht möglich, das Angst-CAT einer gesunden Probandenstichprobe vorzulegen, welche für die Bevölkerung des deutschsprachigen Raumes repräsentativ ist. Jedoch liegen uns von einer Gruppe von N = 35 Medizinstudenten (der Humboldt-Universität zu Berlin) Theta-Werte des Angst-CATs vor, welche im laufenden Sommersemester 2003 erhoben wurden. Diese werden im Folgenden als eine vorläufige Vergleichsstichprobe genutzt.

⁸¹ Die Prozentwerte der Diagnosen summieren sich nicht zu 100%, da Komorbidität zwischen einzelnen Störungen häufig ist.

6.5. Validierungsinstrumente

Zur Validierung wurden im Rahmen der klinisch-psychologischen Routinediagnostik (Testbatterien) das Angst-CAT und fünf psychometrische Verfahren sowie ein strukturiertes diagnostisches Interview angewandt. Diese sollen eine Überprüfung der *Konstruktvalidität* des Angst-CATs ermöglichen. Es wird angenommen, dass sie selbst ausreichend valide Repräsentanten der Konstrukte der Angst / Depression und anderer Persönlichkeitsfaktoren darstellen. Folgende psychometrischen Instrumente, welche sich in der klinischen Diagnostik bewährt haben, wurden an oben beschriebener Patientenstichprobe erhoben:

- **das Beck-Angst-Inventar**
(BAI; Margraf & Ehlers, in Druck),
- **die Hospital Anxiety and Depression Scale**
(HADS; Hermann, Buss & Snaith, 1995),
- **das Beck-Depressions-Inventar**
(BDI; Hautzinger, Bailer, Worall & Keller, 1994),
- **das NEO-Fünf-Faktoren-Inventar**
(NEO-FFI; Borkenau & Ostendorf, 1993) und
- **der Gießen-Test**
(GT; Beckmann, Brähler & Richter, 1991).

Der Einsatz des Angst-CATs erfolgte an einem stationären Computer; alle weiteren psychometrischen Instrumente wurden computergestützt mittels Handcomputer, sogenannter PDA's (Personal Digital Assistants; Psion), deren Einsatz bereits erprobt ist, erhoben (Rose et al., 1999, 2003; siehe Kapitel 5.2.).

Desweiteren wurde eines der international am weitesten verbreiteten, strukturierten klinischen Interviews (siehe Kapitel 2.7.1. und 4.1.) an oben beschriebener Stichprobe computergestützt angewandt: das M-CIDI (als Papierversion: DIA-X) von Wittchen und Pfister (1996). Dieses unter der Schirmherrschaft der World Health Organization (WHO) und dem National Institute of Mental Health (NIMH) an dem Max-Planck-Institut für Psychiatrie in München entwickelte Instrument dient der strukturierten klinischen Diagnostik der Angst als psychischer Störung nach den Kriterien des ICD-10 (Dilling et al., 2000) und DSM-IV (Saß et al., 1996; siehe Kapitel 6.5.3.).

6.5.1. Klinische Instrumente zur Angst und Depressionsmessung

6.5.1.1. Beck-Angst-Inventar (BAI)

Das Beck-Angst-Inventar (Margraf & Ehlers, 1995) ist ein weit verbreitetes und in vielfältigen klinischen Zusammenhängen eingesetztes Selbstbeurteilungsinstrument zur Erfassung des Schweregrads klinisch relevanter Angst in Patientengruppen und der Allgemeinbevölkerung (ab 12 Jahren).

Das Instrument, welches 21 Items mit 4-stufigem Antwortformat umfasst, wurde entwickelt, um Angst hinsichtlich der Schwere ihres Auftretens in den letzten 7 Tagen in Anlehnung an die Symptomlisten des DSM-IV (Saß et al., 1996) für Panikanfälle und generalisierte Angst möglichst exakt und ökonomisch zu messen. Die Items repräsentieren weitestgehend somatische Korrelate der Angst (Westhoff, 1993).

Das BAI basiert auf der amerikanischen Originalversion (Beck & Steer, 1993), welche die Erfassung der Ängstlichkeit möglichst unabhängig von depressiver Symptomatik intendiert. Dieser Anspruch wird nur teilweise eingelöst (Korrelationen mit Depressionsmaßen liegen zwischen $r = 0,43$ (CCL-D⁸²) bis $r = 0,47$ (BDI⁸³), $N = 281$ bzw. $N = 287$, Margraf & Ehlers, in Druck).

Das BAI (Originalversion) korreliert mit der RCMAS⁸⁴ ($N = 80$ psychiatrische erwachsene Patienten) in einer Höhe von $r = 0,58$ und mit der Angst-Skala des MMPIs⁸⁵ nach statistischer Kontrolle des Zusammenhangs zu den BDI-Scores in einer Höhe von $r = 0,30$ ($N = 125$ Jungen) bzw. $r = 0,56$ ($N = 115$ Mädchen; Osman et al., 2002). Der Übereinstimmungskoeffizient zu Fremdratings der Angst von Klinikern liegt an einer psychiatrischer Stichprobe bei $r = 0,40$.

Der deutschen Version des BAIs wird eine sehr gute bis gute interne Konsistenz von $\alpha = 0,92$ ($N = 291$ Patienten mit psychischen und / oder organischen Diagnosen) bis $\alpha = 0,88$ ($N = 3.000$ Personen aus der Allgemeinbevölkerung) und eine mäßig bis hohe Retest-Reliabilität ($N = 1.000$; $r = 0,68$ bei 14 Tagen; $r = 0,9$ bei 48h) zugeschrieben, so dass eine Sensitivität für Therapieeffekte angenommen werden kann. Sie korreliert mit Angstmaßen zu

⁸² CCL-D : Cognition Checklist-Depression (Tönnies, 1995).

⁸³ BDI : Beck-Depressions-Inventar (Hautzinger, Bailer, Worall & Keller, 1994).

⁸⁴ RCMAS: Revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 1978).

⁸⁵ MMPI: Minnesota Multiphasic Personality Inventory for Adolescents (Butcher et al., 1992).

$r = 0,45$ (STAI-State⁸⁶; $N = 154$), $r = 0,48$ (STAI-Trait; $N = 227$), $r = 0,50$ (CCL⁸⁷; $N = 289$) und $r = 0,73$ (SCL-90-R-Angst⁸⁸, $N = 675$; Margraf & Ehlers, in Druck). Populationsnormen für klinische Stichproben und die Allgemeinbevölkerung liegen vor ($N = 2.000$).

6.5.1.2. Hospital Anxiety and Depression Scale (HADS)

Die Hospital Anxiety and Depression Scale (Hermann, Buss & Snaith, 1995) ist ein kurzer Selbstbeurteilungsfragebogen zur Erfassung von Angst und Depressivität bei Erwachsenen. Er wurde gezielt zum Einsatz bei körperlich Kranken konstruiert (Zigmond & Snaith, 1983) und soll im Kontext somatischer Medizin dazu beitragen, Patienten mit psychischer Morbidität zu identifizieren (Brähler, Holling, Leutner & Petermann, 2002). Die HADS besteht aus 14 Items mit 4-stufigem Antwortformat, aus denen je eine Angst- und Depressivitäts-Subskala (HADS-A /-D) gebildet wird. Angst wird in Anlehnung an die Generalisierte Angststörung (DSM-IV; Saß et al., 1996) und Depressivität wird hinsichtlich „endogenomorpher“ Symptome (Freudlosigkeit, Interessenverlust etc.) bezüglich ihres Auftretens in der letzten Woche erfasst.

Die interne Konsistenz der Angst-Subskala liegt bei $\alpha = 0,80$, die der Depressivitäts-Subskala bei $\alpha = 0,81$ ($N = 6.200$ Patienten). Die Retest-Reliabilitäten betragen zwischen $r = 0,7$ (> 6 Wochen) und $r = 0,84$ bzw. $r = 0,85$ (2 Wochen). Korrelationen zu anderen Angst- bzw. Depressionsskalen an $N = 1.815$ Patienten liegen zwischen $r = 0,48$ bis $r = 0,86$ ($\bar{r} = 0,66$ ⁸⁹, HADS-A) bzw. $r = 0,46$ bis $r = 0,78$ ($\bar{r} = 0,59$ ⁹⁰, HADS-D; Hinz & Schwarz, 2001). Interkorrelationen zwischen der Angst- und der Depressionsskala des HADS liegen bei $r = 0,53$. Es existieren Normen für $N = 5.579$ kardiologische Patienten und vorläufige Normen für $N = 278$ Gesunde.

⁸⁶ STAI: State-Trait-Angst-Inventar (Laux et al., 1981).

⁸⁷ CCL: Cognition Checklist (Tönnies, 1995).

⁸⁸ SCL-90-R: Die Symptom-Checkliste von Derogatis (Franke, 1995).

⁸⁹ Gemittelte Korrelation der in der Validierungsstudie des HADS (Hermann et al., 1995) eingesetzten folgenden acht Angstskalen: Angstskala des General Health Questionnaire (GHQ-28); Linear-Analog-Angstskala; Irritability, Depression and Anxiety-Scale (IDA); Zung Angst- und Depressionsskala; Crown-Crisp Experiential Index; Arthritis Impact Measurement Scale (AIMS), de Bonis-Angstskala; State Trait Anxiety Inventory (State).

⁹⁰ Gemittelte Korrelation der in der Validierungsstudie des HADS (Hermann et al., 1995) eingesetzten folgenden 6 Depressionsskalen: Depressionsskala des General Health Questionnaire (GHQ-28); Irritability, Depression and Anxiety-Scale (IDA); Zung Angst- und Depressionsskala; Arthritis Impact Measurement Scale (AIMS); Crown-Crisp Experiential Index; Depressionsskala (D-S).

6.5.1.3. Beck-Depressions-Inventar (BDI)

Das Beck-Depressions-Inventar (Hautzinger, Bailer, Worall & Keller, 1994) ist ein seit 30 Jahren international und national weit verbreitetes Selbstbeurteilungsinstrument zur Erfassung des Schweregrades depressiver Symptomatik bei Jugendlichen ab 16 Jahren und bei Erwachsenen. Es entstand vor dem Hintergrund klinischer Beobachtungen depressiver Patienten und erfasst mit 21 Items die häufigsten depressiven Symptome. Seine innere Konsistenz liegt zwischen $\alpha = 0,73$ und $\alpha = 0,95$, die Retest-Reliabilitäten für eine Woche betragen $r = 0,60$ bzw. $r = 0,86$.

Korrelationen mit anderen Depressionsinventaren liegen zwischen $r = 0,61$ bis $r = 0,86$ (HAMA)⁹¹, $r = 0,57$ bis $r = 0,83$ (SRDS)⁹² und $r = 0,41$ bis $r = 0,70$ (MMPI-D)⁹³. Patienten mit *Angststörungen* haben in der Regel ebenfalls erhöhte BDI-Werte, wenngleich Patienten mit Depressionen im BDI meist signifikant höhere Werte als Angstpatienten zeigen (Beck, 1994). Korrelationen mit Angstinventaren liegen auch in spezifischen Artikeln zur Differenzierbarkeit von Angst und Depression nicht vor (Steer, Beck, Riskind & Brown, 1986). Es liegen Normen einer klinischen Stichprobe (N = 477) depressiver Patienten vor.

6.5.2. Persönlichkeitsinventare

6.5.2.1. NEO-Fünf-Faktoren-Inventar (NEO-FFI)

Das NEO-Fünf-Faktoren-Inventar von Borkenau und Ostendorf (1993; Originalversion: NEO-FFI von Costa & McCrae, 1985) ist ein multidimensionaler Persönlichkeitsstrukturtest für Erwachsene, welcher sowohl für Forschungszwecke, als auch für Anwendungen in der Klinischen Psychologie, der Schullaufbahn-, Studien- und Berufsberatung sowie in der Organisationspsychologie genutzt wird.

Er geht auf den sogenannten psycholexikalischen Ansatz zurück (Allport & Odbert, 1936; Cattell, 1943; Angleitner, Ostendorf & John, 1990). Umfangreiche faktorenanalytische Studien zu individuellen Unterschieden in der Persönlichkeit zeigen, dass der Einschätzung von Personen fünf robuste Dimensionen

⁹¹ HAMA: Hamilton-Angst-Skala (Hamilton, 1959).

⁹² SRDS: Self-Rating Depression Scale (Zung, 1965).

⁹³ MMPI-D: Minnesota Multiphasic Personality Inventory-Depression Scale (Hathaway & McKinley, 1983).

(„Big Five“) zugrunde liegen, welche das NEO-FFI mit 60 fünfstufigen Items (pro Skala: 12 Items) mit den folgenden fünf Skalen erfasst:

1. die „*Neurotizismus*“-Skala erfasst emotionale Stabilität versus Labilität, d. h. inwiefern ein Proband z. B. dazu neigt, nervös, ängstlich, traurig, unsicher und verlegen zu sein und sich Sorgen um seine Gesundheit zu machen;
2. mit der „*Extraversions*“-Skala kann das Ausmaß der Geselligkeit, Selbstsicherheit, Aktivität, Gesprächigkeit und der Optimismus einer Person erhoben werden;
3. die Skala „*Offenheit für Erfahrung*“ misst das Interesse an neuen Erfahrungen, Erlebnissen und Eindrücken;
4. die Skala „*Verträglichkeit*“ erfasst die Neigung zu altruistischem Verhalten und das zwischenmenschliche Vertrauen bzw. Harmoniebedürfnis und Nachgiebigkeit und
5. die Skala „*Gewissenhaftigkeit*“ misst das Ausmaß der Impuls- und Selbstkontrolle (Ordentlichkeit, Zuverlässigkeit, Ehrgeiz, Disziplin).

Die internen Konsistenzen der Skalen liegen bei $\alpha = 0,78$ ($N = 2.112$), die Retest-Reliabilitäten von zwei Jahren liegen zwischen $r = 0,65$ (Verträglichkeit) und $r = 0,81$ (Extraversion). Eine Reihe von Studien zur faktoriellen Validität belegen durch hohe Kongruenzkoeffizienten ($r = 0,91$ bis $r = 0,98$) die Stabilität der Faktorenstruktur über unterschiedliche Stichproben. Zur kriteriumsbezogenen Validität werden im Testhandbuch keine Studien erwähnt. Es liegen keine bevölkerungsrepräsentativen Normen, jedoch statistische Kennwerte einer Standardisierungsstichprobe ($N = 2.112$) vor.

6.5.2.2. Gießen-Test (GT)

Der Gießen-Test von Beckmann, Brähler und Richter (1991) ist ein Selbstbeurteilungsverfahren, welches den Probanden die Gelegenheit gibt, sich selbst in ihrem Realselbst- und Idealselbstbild einzuschätzen. Er dient der klinischen Diagnostik und Therapieverlaufsevaluation und findet unter anderem auch Anwendung in der sozialpsychologischen Forschung. Bei der Erfassung des Realselbst- und Idealselbstbildes werden vor allem die innere Verfassung einer Person und seine psychosozialen Umweltbeziehungen fokussiert.

Der GT besteht aus 40 bipolar formulierten Feststellungen, die auf einer siebenstufigen Skala nach ihrem Zutreffen beantwortet werden sollen, und die zu den folgenden sechs Skalen zusammengefasst werden:

1. die Skala „*Soziale Resonanz*“ dient der Selbsteinschätzung einer Person bezüglich ihrer Wirkung auf die Umwelt. Dazu gehören sowohl äußere Merkmale (Aussehen, Attraktivität) als auch das selbsteingeschätzte eigene Maß an Beliebtheit, Wertschätzung, Achtung und Durchsetzungsfähigkeit;
2. die Skala „*Dominanz*“ bildet Merkmale wie Aggressivität, Eigensinn und Impulsivität versus Gefügigkeit bzw. Unterordnungstendenzen ab;
3. die Skala „*Kontrolle*“ erfasst das Ausmaß der Selbstkontrolle im Sinne von Ordentlichkeit, Stetigkeit, Eifer und Genauigkeit im Umgang mit Objekten;
4. die Skala „*Grundstimmung*“ dient der Erfassung der allgemeinen Stimmung (u. a. Depressivität, Ängstlichkeit und Ärger);
5. die Skala „*Durchlässigkeit*“ erfasst die zwischenmenschliche Aufgeschlossenheit, Vertrauensseeligkeit und die Fähigkeit, psychosoziale Bedürfnisse im Kontakt mit anderen Menschen zu äußern;
6. die Skala „*Soziale Potenz*“ erhebt das Ausmaß an sozialen Fähigkeiten wie Geselligkeit, Hingabefähigkeit, Konkurrenzfähigkeit etc., welche eine Person sich selbst zuschreibt.

Die mittlere interne Konsistenz der Skalen liegt bei $\alpha = 0,86$ (N = 235 „neurotische“ Patienten); die Restest-Reliabilitäten für sechs Wochen liegen zwischen $r = 0,65$ und $r = 0,76$ (N = 204 „neurotische“ Patienten). Da die Autoren eine konzeptuelle Validität aufgrund der gezielten Auswahl tiefenpsychologisch und sozialpsychologisch relevanter Feststellungen annehmen, liegen Ergebnisse von Kriteriumsvalidierungsstudien an N = 2.182 Probanden vor (zwei eingesetzte Vergleichsinstrumente: a) zum Erziehungsverhalten und b) zu interpersonellen Problemen; Brähler, Schumacher & Brähler, 1999). Eine aktuelle Normierung (1999) an N = 1.008 Ostdeutschen und N = 995 Westdeutschen findet sich bei Brähler und Richter (2000).

6.5.3. Diagnostisches Interview: M-CIDI (DIA-X)

Das Munich Composite International Diagnostic Interview (M-CIDI; Wittchen & Pfister, 1996) ist ein voll standardisiertes computergestütztes Interview-

verfahren zur diagnostisch-klassifikatorischen Erfassung psychischer Störungen, welches sich zum Einsatz in der klinischen Praxis und Forschung (v. a. in epidemiologischen Studien) bei Probanden im Alter von 14 bis 65 Jahren eignet. Als Papierversion nennt es sich *DI*Agnostisches *EX*pertensystem psychischer Störungen (DIA-X).

Es wurde unter der Schirmherrschaft der Weltgesundheitsorganisation (WHO) und dem National Institute of Mental Health (NIMH, U.S.A.) entwickelt und erlaubt die Diagnostik von 64 Störungen nach den Kriterien des ICD-10 (Dilling, et al., 2000) und DSM-IV (Saß et al., 1996). Folgende psychische Störungen werden in 12 Interviewsektionen erfragt: a) Störungen durch Tabak, b) somatoforme und dissoziative Störungen, c) Phobien und andere Angststörungen, d) depressive Störungen und Dysthymie, e) Manie und bipolare affektive Störungen, f) Schizophrenie und andere psychotische Störungen, g) Essstörungen, h) Störungen durch Alkohol, i) Zwangsstörungen, j) Drogenmissbrauch und –abhängigkeit, k) organisch bedingte psychische Störungen, l) posttraumatische Belastungsstörungen.

Das M-CIDI-Programmpaket ermöglicht sowohl eine simultan zum Interview verlaufende computergestützte Dateneingabe sowie eine automatische *Auswertung* des Interviews nach den diagnostischen Kriterien des ICD-10 und DSM-IV. Der Diagnosenausdruck umfasst Angaben zu den vorliegenden psychischen Störungen, deren erstes und letztes Auftreten, dem jeweiligen Schweregrad und der Komorbiditätsstruktur.

Die Interrater-Reliabilitäten liegen zwischen $\kappa = 0,81$ und $\kappa = 1,0$ (*symptombezogene* Interrater-Reliabilitäten) bzw. $\kappa = 0,82$ und $\kappa = 0,98$ (*diagnosenbezogene* Interrater-Reliabilitäten). Die Restest-Reliabilitäten von 1-14 Tagen (N = 142 Fälle) liegen zwischen $\kappa = 0,49$ (undifferenzierte somatoforme Störung) und $\kappa = 0,83$ (Anorexia nervosa); für Angststörungen beträgt sie $\kappa = 0,57$ (soziale Phobie) bis $\kappa = 0,92$ (Panikattacken).

Die Validität variiert stark zwischen unterschiedlichen Diagnosegruppen. Im Vergleich zu klinischen Konsensus-Diagnosen erfahrener Psychiater ergaben sich Übereinstimmungswerte zu der strukturierten computergestützten Interviewdiagnostik von $\kappa = 0,39$ (psychotische Störungen), $\kappa = 0,39 / 0,43$ (somatoforme Störungen) bis $\kappa = 0,82$ (Panikstörungen).

6.6. Methodisches Vorgehen

Die folgende Beschreibung des methodischen Vorgehens orientiert sich an der Reihenfolge der Darstellung der Ergebnisse.

Der Ergebnisteil, welcher die Validierung des Angst-CATs beinhaltet (Kapitel 6.7.), gliedert sich in einen *ersten* allgemein deskriptiven Ergebnisteil (6.7.1.), einen *zweiten* Teil, welcher der konvergenten Validierung (6.7.2.), und einen *dritten* Teil, welcher der diskriminanten Validierung des Angst-CATs (6.7.3.) dient.

Im *ersten* Teil (Kapitel 6.7.1.) werden die Itemselektion, d. h. die im Angst-CAT dargebotene Anzahl der Items in Abhängigkeit von den geschätzten Theta-Werten mit deskriptiven Statistiken untersucht, mögliche Zeitersparnisse bei Einsatz des Angst-CATs gegenüber herkömmlichen Instrumenten analysiert und Verteilungsparameter der Theta-Werte des Angst-CATs in Abhängigkeit von soziodemografischen Variablen exploriert. Als inferenzstatistische Prüfmethode werden zur Überprüfung von Mittelwertsunterschieden t-Tests für unabhängige Stichproben (z. B. zur Untersuchung eines Geschlechtseffekts) sowie einfaktorielle Varianzanalysen (zur Untersuchung von Alters- bzw. Familienstandseffekten) durchgeführt.

Der *zweite* Teil (Kapitel 6.7.2.) umfasst die *konvergente* Validierung des Angst-CATs, d. h. es wird die Beziehung zu Instrumenten, deren Messbereiche konstruktnah bzw. -identisch mit dem des Angst-CATs sind, untersucht. Dieser Teil gliedert sich in zwei Unterkapitel. Zunächst wird die konvergente Validität in Bezug auf andere psychometrische Testverfahren und anschließend in Bezug auf das mit dem strukturierten Interview (M-CIDI) erhobene Fremdurteil untersucht.

Zur Überprüfung der konvergenten Validität bezüglich verschiedener Testverfahren wurden Produkt-Moment-Korrelationen (Pearson's Korrelationskoeffizient) mit den erhobenen Summenscores der Angst-Inventare (Angst-CAT, BAI, HADS-A) berechnet. Die konvergente Validität in Bezug auf das Fremdurteil wird untersucht, indem die Mittelwertsunterschiede der Theta-Schätzungen verschiedener Stichproben, welche mit dem M-CIDI klassifikatorisch erfasst wurden (Patienten ohne bzw. mit Angststörungen, Referenzgruppe: Medizinstudenten), inferenzstatistisch mittels einfaktorieller Varianzanalysen überprüft werden.

Der dritte Teil (Kapitel 6.7.3.), welcher der Untersuchung der *diskriminanten* Validität dient, d. h. der Zusammenhangsuntersuchung des Angst-CATs zu Instrumenten, welche die Messung unterschiedlicher Konstrukte intendieren, gliedert sich - wie Kapitel 6.7.2. - in eine Untersuchungsphase zur Überprüfung der Validität in Bezug auf andere Testverfahren und in Bezug auf das diagnostische Fremdurteil (M-CIDI).

Die Überprüfung der diskriminanten Validität in Bezug auf andere Testverfahren geschieht mittels korrelativer Statistiken (Pearson's Korrelationskoeffizient). Hier wird zunächst die Diskriminationsfähigkeit des Angst-CATs zwischen den Konstrukten Angst und Depression untersucht, indem die erhobenen Angst- und Depressionsinventare (BAI, HADS, BDI) in korrelative Beziehung gesetzt werden. Anschließend wird die psychometrische Diskriminationsfähigkeit des Angst-CATs zu anderen Persönlichkeits-konstrukten mittels der Ergebnisse der zwei eingesetzten Persönlichkeits-inventare (NEO-FFI, GT) korrelations- und regressionsstatistisch exploriert.

Die diskriminante Validität in Bezug auf das diagnostische Fremdurteil (M-CIDI) wird bestimmt, indem die Mittelwerte der Theta-Schätzungen von verschiedenen mit dem M-CIDI ermittelten Diagnosegruppen (Patienten mit Angst-, depressiven, Ess- und somatoformen Störungen) verglichen werden. Da bei psychosomatischen Patienten hohe Komorbiditätsraten zwischen den einzelnen Störungsgruppen zu erwarten sind, erfolgt eine Überprüfung der Mittelwertsunterschiede der Theta-Werte zunächst zwischen den in der Realität am häufigsten vorkommenden Diagnosegruppen, bei denen Patienten mehrere Störungen aufweisen (Komorbidität), d.h. aufgrund von Komorbidität kommt es zu Personenüberschneidungen zwischen den Diagnosegruppen. Um den die diskriminante Validität des Angst-CATs möglicherweise beeinträchtigenden Einfluss dieser Komorbidität zu eliminieren, werden anschließend die verschiedenen Diagnosegruppen unter Ausschluss von Komorbidität gebildet (d. h. ohne Überschneidungen zwischen den Diagnosegruppen) und inferenzstatistisch auf Mittelwertsunterschiede in den Theta-Werten überprüft. Als globale Prüfmethode wurde die einfaktorielle Varianzanalyse, als Einzelvergleichsmethode der Scheffé-Test eingesetzt. Alle erläuterten Berechnungen⁹⁴ erfolgten unter Einsatz des Programms SPSS 10.0.

⁹⁴ Allen erörterten Berechnungen ist gemein, dass sowohl das Intervallskalenniveau als auch

6.7. Ergebnisse

6.7.1. Allgemeine Ergebnisse zum Angst-CAT

6.7.1.1. Die Itemselektion

Die Stoppfunktion des IRT-basierten Itemselektionsalgorithmus (Kapitel 4.3.3.6.) des Angst-CATs wurde im Rahmen vorliegender Validierungsstudie auf eine Reliabilität von $Rel(\theta) = 0,9$ festgelegt. Die Anwendung des Angst-CATs an $N = 102$ psychosomatischen stationären Patienten (Kapitel 6.4.) zeigte, dass...

1. eine Erfassung der Angstaussprägung mit im Durchschnitt $5,3 \pm 1,9$ Items ($\bar{X} \pm SD$) auf diesem Messpräzisionsniveau möglich ist;
2. die durchschnittliche Testdurchführungszeit der Patienten⁹⁵ 1 min. und 40 sek. beträgt ($SD = 49s$)⁹⁶;
3. zwischen 4 und 14 Items pro Testdurchlauf zur Angsterfassung genutzt werden (Verteilungsspannweite);
4. die darzubietende Itemanzahl bei einer durchschnittliche Angstaussprägung gering ist, jedoch zu den Extremausprägungen hin zunimmt (siehe Abbildung 21).

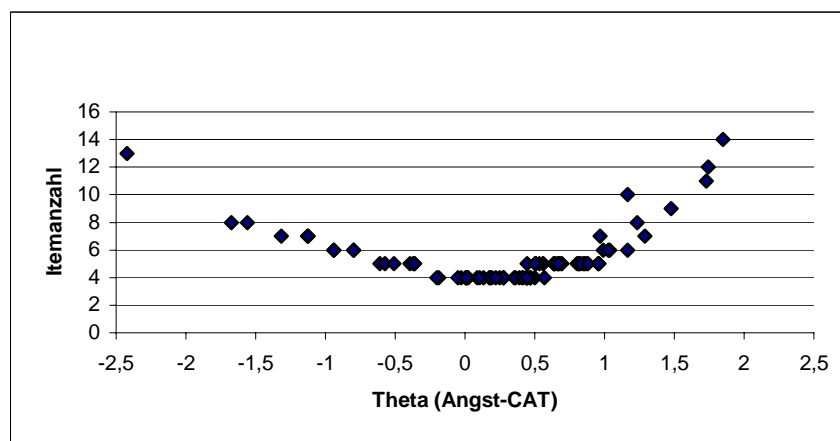


Abbildung 21: Verteilung der im Angst-CAT dargebotenen Anzahl der Items in Abhängigkeit von den durch das Angst-CAT geschätzten Theta-Werten (N = 102 psychosomatische Patienten).

die Normalverteilung vorausgesetzt wird.

⁹⁵ Zum Vergleich: Die durchschnittliche Testdurchführungszeit des Angst-CATs beträgt bei Studenten: 1min., 25sek., $SD = 46sek.$

⁹⁶ Zum Vergleich: Die durchschnittliche Testdurchführungszeit des STAls mit 40 Items wird im Testhandbuch (Laux et al., 1981) auf 6-10 min. geschätzt.

Letzterer Befund resultiert aus der Beschaffenheit der Itembank, welche aus vielen hoch informativen Items besteht, die eine *durchschnittliche* Angstaussprägung (bezogen auf das psychosomatische Kollektiv) gut erfassen, jedoch weniger Items aufweist, welche in den *Extrembereichen* der Angstaussprägung eine hohe Iteminformation aufweisen, so dass – wird eine konstante Messgenauigkeit über alle Merkmalsausprägungsbereiche hinweg angestrebt (zur Stoppfunktion siehe Kapitel 4.3.3.6.) – in den Extrembereichen mehr Items dargeboten werden müssen, um diese zu gewährleisten.

Die dargestellten Befunde replizieren Ergebnisse einer Simulations-Vorstudie zur Güte des Angst-CATs (Walter et al., eingereicht), in der computergestützt konventionell (d. h. nicht adaptiv) erhobene psychometrische Daten von $N = 2.348$ psychosomatischen Patienten mit einem simulierten adaptiven Itemselektionsalgorithmus so reanalysiert wurden, dass für jeden Patienten eine IRT-basierte Theta-Schätzung mit dem Angst-CAT erfolgte. Der in Abbildung 21 veranschaulichte Zusammenhang offenbarte sich bereits in dieser Simulations-Vorstudie. Im Rahmen der Vorstudie konnte weiterhin die im Angst-CAT darzubietende Itemanzahl bei unterschiedlichen Stoppfunktionen simuliert werden. Bei einer Stoppfunktion von $Rel(\theta) = 0,9$ zeigten sich ähnliche Ergebnisse wie in der Validierungsstudie: $6,9 \pm 2,6$ Items ($\bar{X} \pm SD$) wurden zur Schätzung der Theta-Werte vom Angst-CAT in Simulationen genutzt. Wurde das Angst-CAT mit einer Stoppfunktion von $Rel(\theta) = 0,8$ simuliert, so benötigte es zur Angsterfassung nur $3,1 \pm 0,8$ Items ($\bar{X} \pm SD$).

Erste Hinweise aus dieser Studie widersprechen der naheliegenden Vermutung eines deutlichen Informationsverlusts bei der Darbietung dieser geringen Anzahl von Items. Abbildung 22 veranschaulicht den hohen korrelativen Zusammenhang ($r = 0,97$)⁹⁷ zwischen der simulierten Theta-Schätzung des Angst-CATs auf der Grundlage der gesamten Itembank und der des Angst-CATs (Walter et al., eingereicht).

⁹⁷ Bei der Interpretation dieses Ergebnisses muss berücksichtigt werden, dass sich die hohe Ausprägung der Korrelation teilweise aus einer sich überlappenden Itemmenge ergibt.

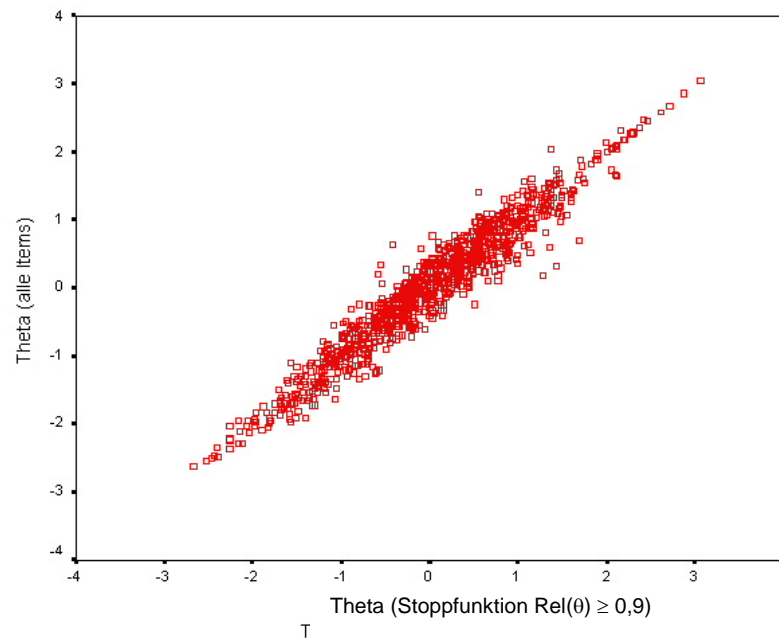


Abbildung 22: Beziehung zwischen der Theta-Schätzung auf der Grundlage aller Items der Itembank und der Theta-Schätzung des Angst-CATs (Stoppfunktion $\text{Rel}(\theta) \geq 0,9$).

6.7.1.2. Statistische Kennwerte in Abhängigkeit von soziodemografischen Variablen

Im Folgenden wurde die Verteilung der Theta-Werte des Angst-CATs von $N = 102$ psychosomatischen stationären Patienten bezüglich soziodemografischer Kennwerte wie des Geschlechts, Alters und Familienstands untersucht.

Ein durchgeführter t-Test für unabhängige Stichproben zur zufallskritischen Überprüfung möglicher *geschlechtsbedingter* Mittelwertsunterschiede führt zu keinem signifikanten Ergebnis (Tabelle 20).

Tabelle 20: Statistische Kennwerte des Angst-CATs in Abhängigkeit vom Geschlecht.

	Geschlecht	N	\bar{X}	SD	SE	Mittlere Differenz	SE Differenz	t-Wert	df	p
Theta Angst-CAT	weiblich	81	,326	,735	,082	-0,019	0,177	-0,106	100	0,916
	männlich	21	,345	,674	,147					

Ebenfalls keine signifikanten Ergebnisse resultierten aus einfaktoriellen Varianzanalysen zur Überprüfung der Mittelwertsunterschiede zwischen verschiedenen Altersgruppen ($Q_{\text{between}} = 1,31$; $df = 5$; $\bar{Q} = 0,26$; $F = 0,49$;

$p = 0,78$), obgleich Patienten der Altersgruppe der 26-35-Jährigen und der über 75-Jährigen leicht geringere Theta-Werte im Angst-CAT aufweisen als Patienten sonstiger Altersgruppen (siehe Tabelle 21).

Tabelle 21: Statistische Kennwerte des Angst-CATs unterschiedlicher Altersgruppen.

	Alter	N	\bar{X}	SD	SE
Theta Angst-CAT	18-25 Jahre	20	,466	,650	,145
	26-35 Jahre	16	,149	,819	,205
	36-45 Jahre	23	,388	,808	,169
	46-55 Jahre	20	,358	,819	,183
	56-65 Jahre	15	,319	,480	,124
	> 75 Jahre	8	,132	,611	,216

Auch die Überprüfung der Mittelwertsunterschiede zwischen Gruppen unterschiedlichen *Familienstandes* (einfaktorielle Varianzanalyse) führte zu keinen signifikanten Ergebnissen ($Q_{\text{between}} = 0,33$; $df = 3$; $\bar{Q} = 0,11$; $F = 0,21$; $p = 0,89$; Abbildung 23).

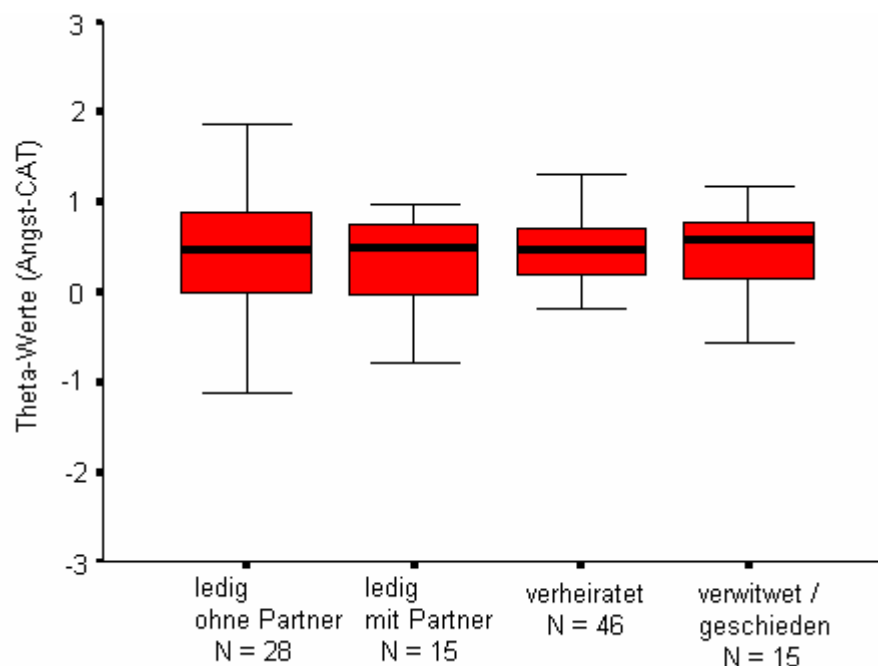


Abbildung 23: Die Theta-Werte-Verteilung des Angst-CATs in Abhängigkeit vom Familienstand.

6.7.2. Konvergente Validierung

In der bereits erwähnten Simulations-Vorstudie zur Güte des Angst-CATs (Walter et al., eingereicht) wurden die Theta-Werte des Angst-CATs psychosomatischer Patienten bereits in Beziehung zur State-Skala des STAI (Laux et al., 1981) gesetzt, um erste Hinweise auf die konvergente Validität des Instruments zu erhalten. Die Simulationsexperimente führten zu einer Korrelation zwischen dem Angst-CAT und der State-Skala des STAI von $r = 0,88$. Da die Itembank des Angst-CATs jedoch 15 der 20 Items der State-Skala des STAI umfasst, kann dieser Befund einer überlappenden Itemmenge geschuldet sein, so dass es der tiefergehenden Untersuchung der Validität - wie sie mit folgender Validierungsstudie realisiert wird - bedurfte.

6.7.2.1. Konvergente Validität in Bezug auf die Angst-Inventare

In der vorliegenden prospektiven Validierungsstudie wurde zunächst das Angst-CAT in korrelationsstatistischen Zusammenhang mit zwei Angstskalen gesetzt (HADS-A und BAI; Kapitel 6.5.1.) . Da aus organisatorischen Gründen nur die Hälfte der Patienten ($N = 50$) sowohl das Angst-CAT als auch die anderen psychometrischen Instrumente innerhalb von 48 Stunden beantworten konnten (bei der anderen Hälfte der Patienten liegt die Differenz zwischen den Messzeitpunkten bei bis zu 14 Tagen), werden hier Ergebnisse dieser Teilstichprobe ($N = 50$) und der Gesamtstichprobe ($N = 102$) berichtet.

Tabelle 22: Korrelationen zwischen dem Angst-CAT und den zwei Angst-Skalen.

	Zeitdifferenz zwischen den Testerhebungen	N	HADS-Angst	BAI
Theta Angst-CAT	< 14 Tage	102	,66*	,51*
	48 h	davon: 50	,76*	,55*

Die Korrelationen ($r = 0,51-0,76$) deuten - verglichen mit der Interkorrelation der eingesetzten etablierten Instrumente ($r_{\text{HADS-A / BAI}} = 0,66$; $N = 102$) oder der bekannten Interkorrelationen dieser Angstinventare zu anderen Angstskalen (Kapitel 6.5.1.1./2.: $r_{\text{BAI / STAI (S/T)}}^{98} = 0,45/0,48$; $r_{\text{HADS-A / Angstskalen}}^{99} = 0,48-0,86$;

⁹⁸ In der Validierungsstudie des BAIs (Margraf & Ehlers, in Druck) errechnete Korrelationen.

⁹⁹ Gemittelte Korrelation der folgenden in der Validierungsstudie des HADS (Hermann et al., 1995) eingesetzten acht Angstskalen: Angstskala des General Health Questionnaire (GHQ-28); Linear-Analog-Angstskala; Irritability, Depression and Anxiety Scale (IDA); Zung Angst- und

$\bar{r} = 0,66$) - auf eine *mittelmäßige* bis *gute* konvergente Validität des Angst-CATs hin. Die höheren Korrelationen bei einer zeitnäheren Erhebung ist vor dem Hintergrund der Intention der Messung einer zeitlich variablen *Zustands-Angst* zu erwarten. Insgesamt liegen die Korrelationen des Angst-CATs zur Angstskaala des HADS höher als die zum BAI. Für den Unterschied in der Korrelationshöhe sind die Iteminhalte der verschiedenen Angstskalen verantwortlich. Während die Angstskaala des HADS inhaltlich eine hohe Itemtextähnlichkeit zu Items des Angst-CATs aufweist (erfragt werden Gefühle der An- und Entspannung, Rastlosigkeit, beunruhigende Gedanken, Zukunfts-sorgen und Panikzustände), erfragen 13 (von 21) der Items des BAIs somatische Korrelate der Angst (z. B. Taubheits-, Hitze-, Schwindel, Erstickungs- und Schwächegefühle), welche im Rahmen der Konstruktion des Angst-CATs größtenteils aufgrund von Verletzungen der Unidimensionalitätsannahme im Rahmen der statistischen Itemanalyse und -selektion aus der Itembank ausgeschlossen wurden (siehe Kapitel 5.4.1.). Insofern unterscheidet sich die im Angst-CAT realisierte Konzeptualisierung der Angst stärker von derjenigen des BAIs als von derjenigen des HADS.

6.7.2.2. Konvergente Validität in Bezug auf das diagnostische Fremdurteil

Neben psychometrischen Instrumenten wurde ein strukturiertes diagnostisches Interview (M-CIDI; Wittchen & Pfister, 1996; siehe Kapitel 6.5.3.) zur Diagnostik psychischer Störungen an der psychosomatischen Stichprobe eingesetzt. Als Variablen für eine diagnostische Überprüfung der konvergenten Validität werden die im M-CIDI ermittelte Diagnose einer Angststörung (F.40-41.9 nach ICD-10, Dilling et al., 2000; siehe Kapitel 2.6.1.) bzw. das Fehlen der Diagnose einer psychischen Störung (keine F-Kodierung im ICD-10) herangezogen. Abbildung 24 veranschaulicht die Mittelwerte der Theta-Werte der Patientienstichproben ($N_{\text{Angstdiagnose}} = 58$, $N_{\text{keine F-Diagnose}} = 10$), sowie einer nicht diagnostizierten Vergleichsstichprobe von Medizinstudenten ($N = 35$).¹⁰⁰

Depressionsskala; Crown-Crisp Experiential Index; Arthritis Impact Measurement Scale (AIMS); de Bonis-Angstskaala; State Trait Anxiety Inventroy (State).

¹⁰⁰ Eine Normierung des Angst-CATs an gesunden Probanden ist nicht Gegenstand dieser Arbeit, wird jedoch in naher Zukunft erfolgen.

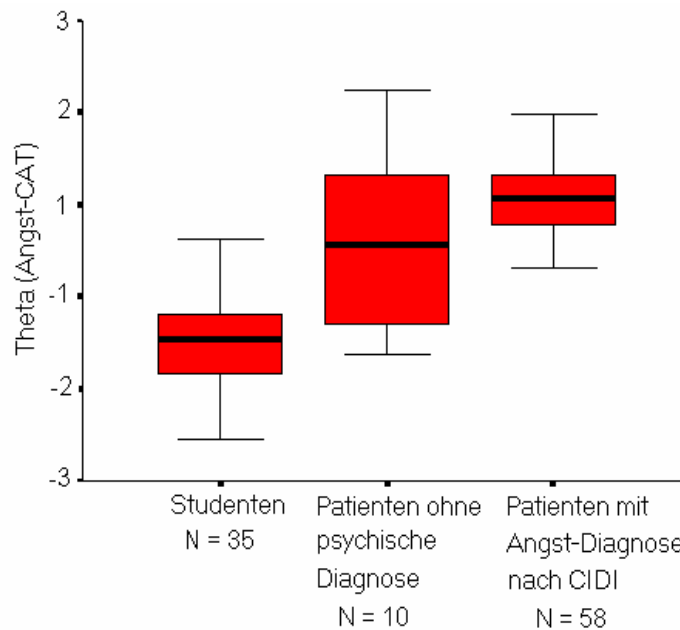


Abbildung 24: Die Theta-Werte-Verteilung des Angst-CATs verschiedener Vergleichsgruppen.

Tabelle 23 berichtet die statistischen Kennwerte des in Abbildung 24 dargestellten Befundes. Eine einfaktorielles Varianzanalyse zur zufallskritischen Absicherung der in der Tabelle dargestellten Mittelwertsunterschiede zeigt, dass sich Patienten mit einer Angststörung von Patienten ohne eine psychische Störung sowie von Studenten in den Theta-Gruppenmittelwerten statistisch bedeutsam unterscheiden ($Q_S = 41,53$; $df = 2$; $\overline{QS} = 20,763$; $F = 35,58$, $p \leq 0,001$).

Tabelle 23: Statistische Kennwerte verschiedener Vergleichsgruppen.

	Gruppe	N	\bar{X}	SD	SE
Theta Angst-CAT	Patienten mit Angst-Diagnose	58	,445	,715	,094
	Patienten ohne F-Diagnose	10	,043	,939	,297
	Studenten	35	-,932	,791	,134

6.7.3. Diskriminante Validierung

6.7.3.1. Diskriminante Validität in Bezug auf andere Testverfahren

Zur Exploration der diskriminanten Validität wurden zwei Depressionsinventare (HADS-Depressionsskala; BDI) und zwei Persönlichkeitsinventare (NEO-FFI, GT) an der psychosomatischen Stichprobe (siehe Kapitel 6.4. und 6.5.) angewandt. Diese werden im Folgenden in korrelationsstatistischen Zusammenhang mit dem Angst-CAT gesetzt.

6.7.3.1.1. Angst und Depression

Tabelle 24 veranschaulicht die korrelativen Beziehungen zwischen den eingesetzten Angst- und Depressionsinventaren.

Tabelle 24: Korrelationsgrid: Angst- und Depressionsinventare (N = 102 psychosomatische Patienten).

		Angst			Depression	
		Angst-CAT	HADS-A	BAI	HADS-D	BDI
Angst	Angst-CAT	1,000	,663*	,514*	,598*	,593*
	HADS-Angst		1,000	,658*	,608*	,619*
	BAI			1,000	,470*	,563*
Depression	HADS-Depression				1,000	,711*
	BDI					1,000

Signifikante Korrelationen: *: $p \leq 0,05$.

In Einklang mit den theoretischen Ausführungen zur Diskrimination von Angst und Depression (Kapitel 2.5.) und den empirischen Ergebnisse aus anderen Validierungsstudien ($r_{\text{HADS-A/HADS-D}} = 0,53$; $r_{\text{BAI/BDI}} = 0,47$; $r_{\text{BAI/CCL-D}} = 0,43$; siehe Kapitel 6.5.1.) zeigt Tabelle 24, dass eine psychometrische Diskrimination zwischen den Konstrukten „Angst“ und „Depression“ nicht gelingt. Während die Korrelation zwischen dem Angst-CAT und der Angst-Skala des HADS (HADS-A) die Korrelationen zu den Depressions-Skalen (HADS-D, BDI) übersteigt, gilt dies nicht für die Korrelation zwischen dem Angst-CAT und dem BAI. Dieser Befund kann - wie im vorangegangenen Kapitel 6.7.2.1. bereits vermutet - wahrscheinlich durch die unterschiedliche Konzeptualisierung der Konstrukte dieser beiden Skalen (Angst-CAT / BAI) erklärt werden.

Vergleicht man die Korrelationsspannweite (range) des Angst-CATs zu den Depressionsinventaren mit derjenigen der zwei anderen Angst-Skalen

(HADS-A, BAI), so zeigt sich, dass die Korrelationsspannweite des Angst-CATs zu den Depressionsinventaren ($r_{A-CAT/HADS-D} = 0,59$; $r_{A-CAT/BDI} = 0,60$) im Mittelfeld zwischen derjenigen des BAIs ($r_{BAI/HADS-D} = 0,47$; $r_{BAI/BDI} = 0,56$) und derjenigen der HADS-Angstskala zu den Depressionsinventaren ($r_{HADS-A/HADS-D} = 0,61$; $r_{HADS-A/BDI} = 0,62$) liegt, und damit mit diesen Instrumenten vergleichbar ist.

6.7.3.1.2. Angst und Persönlichkeitskonstrukte

Nach der Erörterung der Korrelationen zwischen den Angst- und Depressionsinventaren, werden nun die korrelativen Beziehungen des Angst-CATs zu den Skalen von zwei Persönlichkeitsinventaren: dem NEO-Fünf-Faktoren-Inventar (NEO-FFI) und dem Gießen-Test (GT) beschrieben (siehe Kapitel 6.5.2.). Diese sind in Tabelle 25 abgebildet. Zur besseren Einordnung des psychometrischen „Standorts“ des Angst-CATs im Gesamt der psychometrischen Instrumente sind die Korrelationen der beiden Angstinventare (BAI, HADS-A) mit in der Tabelle aufgeführt.

Insgesamt sind die korrelativen Beziehungen der Angstinventare (Angst-CAT, BAI, HADS-A) zum NEO-FFI etwas stärker ausgeprägt als die zum GT.

Betrachtet man zunächst die Korrelationen der Angstinventare zum NEO-FFI, so zeigt sich, dass alle drei Angstinventare (Angst-CAT, BAI und HADS-A) insbesondere das Angst-CAT hoch mit der Skala „Neurotizismus“ korrelieren ($r = 0,51$ bis $r = 0,63$). Dass mit dem Angst-CAT keine bessere Differenzierung zwischen Angst und Neurotizismus gelingt als mit herkömmlichen Instrumenten, ist nicht erstaunlich, da das Angst-CAT (bislang) ausschließlich aus Items etablierter Fragebogen besteht.

Die Berechnung einer einfachen linearen Regression mit dieser Skala führt zu folgender Regressionsgleichung:

$$\text{Angst-CAT} = 0,637 * \text{Neurotizismus-Skala (NEO-FFI)} - 1,147;$$
$$\text{QS}_{\text{Regression}} = 20,53; \text{QS}_{\text{Residuen}} = 31,83; R^2 = 0,39; F = 64,48; p \leq 0,001.$$

Diese verdeutlicht, dass die Skala „Neurotizismus“, welche von Costa und McCrae (1985) als stabile Eigenschaft („Trait“) konzipiert wurde, knapp 40% der Varianz der Theta-Werte des Angst-CATs aufzuklären vermag.

**Tabelle 25: Korrelationsgrid: Angst- und Persönlichkeitsinventare
(N = 102 psychosomatische Patienten).**

	A - CAT	BAI	HADS - A	NEO-FFI					GT ¹⁰¹					
				Neu	Ex	Off	Ver	Gew	SoRe	Do	Ko	Stim	Dur	SoPo
A-CAT	1,000	,514*	,663*	,626*	-,304*	-,086	-,130	-,322*	-,206*	-,053	-,118	,122	,000	-,107
BAI		1,000	,658*	,506*	-,218*	-,174	-,159	-,202*	-,096	-,066	,021	,074	,004	,017
HADS-A			1,000	,591*	-,288*	-,066	-,118	-,226*	-,185*	-,053	-,057	,139	,059	,006
NEO-FFI	Neu			1,000	-,547*	-,138	-,263*	-,546*	-,335*	-,075	-,180	,260*	,159	,072
	Ex				1,000	,276*	,165	,408*	,341*	-,174	-,017	-,099	-,424*	-,410*
	Off					1,000	,129	,122	-,030	-,229*	,058	,181	-,017	-,163
	Ver						1,000	,208*	-,020	,279*	,084	,027	-,033	-,042
	Gew							1,000	,371*	,067	,319*	-,184	-,292*	-,251*
GT	SoRe								1,000	,152	,120	-,625*	-,662*	-,633*
	Do									1,000	,132	-,351*	-,099	,076
	Ko										1,000	,110	,026	,006
	Stim											1,000	,448*	,393*
	Dur												1,000	,661*
	Sopo													

Farbmarkierung: Korrelationshöhe: hellgrau: $r > 0,4$; mittelgrau: $r > 0,5$; dunkelgrau: $r > 0,6$;

Signifikante Korrelationen: *: $p \leq 0,05$;

Abkürzungen: NEO-FFI: Neu: Neurotizismus; Ex: Extraversion; Off: Offenheit für Erfahrungen;

Ver: Verträglichkeit; Gew: Gewissenhaftigkeit; GT: SoRe: Soziale Resonanz; Do: Dominanz;

Ko: Kontrolle; Stim: Grundstimmung; Dur: Durchlässigkeit; SoPo: Soziale Potenz.

Der Messbereich dieser Skalen steht konzeptuell insofern in einem engen Zusammenhang, als bei der Erfassung der emotionalen Stabilität einer Person (Neurotizismus) Items genutzt werden, welche das Erleben negativer Gefühlszustände erfragen. Als negative Gefühlszustände werden Erschütterung, Betroffenheit, Beschämung, Traurigkeit, Sorgen, Unsicherheit, Nervosität und Ängstlichkeit erfragt (Borkenau & Ostendorf, 1993, S. 27), d. h. Begrifflichkeiten verwendet, die zum Teil auch bei der Erfassung der Zustands-Angst eine Rolle spielen. State- und Trait-Angst werden im STAI zwar als zwei Dimensionen konzipiert, die Interkorrelation dieser Skalen liegt jedoch mit $r = 0,43$ bis $r = 0,75$ (bei unterschiedlichen Stichproben) recht hoch (Laux et al., 1981). Dieser Befund und schon im Testmanual des STAI's berichtete Ergebnisse (Laux et al., 1981) deuten darauf hin, dass die Zustands- und Eigenschafts-Angst nicht (statistisch) unabhängig voneinander sind (siehe Kapitel 2.4.1. und 2.7.3.2./3.).

¹⁰¹ Hohe Werte auf den Skalen 1.-4. indizieren eine hohe Ausprägung von 1. Sozialer Resonanz, 2. Dominanz, 3. Kontrolle und 4. (positiver) Grundstimmung, ein hoher Wert auf der Skala 5. indiziert emotionale Verschlussenheit, ein hoher Wert auf der Skala 6. indiziert geringe soziale Kompetenz.

Die geringste Korrelation eines Angstinventars mit der Neurotizismus-Skala findet sich beim BAI. Dies gründet sich wahrscheinlich in der Fokussierung dieser Skala auf der Erfassung somatischer Angstkorrelate (siehe Kapitel 6.5.1.1., 6.7.2.1. und 6.7.3.1.). Die hohe Korrelation zum Angst-CAT und zur HADS deutet auf eine größere Konstruktnähe dieser Skalen hin.

Der in den Korrelationen des Angst-CATs zu den anderen vier Skalen des NEO-FFIs aufgezeigte Zusammenhang lässt sich regressionsstatistisch durch die folgende multiple lineare Regression näher explorieren.

$$\text{Angst-CAT} = 0,686 * \text{NEU} + 0,071 * \text{EX} - 0,023 * \text{OFF} + 0,070 * \text{VER} + 0,023 * \text{GEW} - 1,578;$$
$$\text{QS}_{\text{Regression}} = 20,73; \text{QS}_{\text{Residuen}} = 31,64; R^2 = 0,40; F = 12,58; p \leq 0,001.$$

Ogleich sich signifikante Korrelationen der Angstinventare zu den Skalen Extraversion und Gewissenhaftigkeit zeigen, offenbart oben dargestellte Regressionsgleichung, dass das Hinzufügen der anderen Skalen des NEO-FFI als zusätzliche Prädiktorvariablen zu keiner deutlichen Verbesserung der Varianzaufklärung im Vergleich zur einfachen linearen Regression führt.

Die Korrelationen sind somit wahrscheinlich vor dem Hintergrund relativ hoher Interkorrelationen (im NEO-FFI) zwischen den Skalen „Neurotizismus“ und „Extraversion“ bzw. „Gewissenhaftigkeit“ zu sehen ($r = 0,55$). Die negative Beziehung des Konstrukts der Ängstlichkeit zu dem der Extraversion ist konstrukttheoretisch durch den Zusammenhang zwischen *sozialer* Ängstlichkeit und der Neigung zu Introversion (Borkenau & Ostendorf, 1993, S. 28) erklärbar. Die negative Beziehung des Konstrukts der Ängstlichkeit zu dem der Gewissenhaftigkeit ist vor dem Hintergrund verständlich, dass „die Unfähigkeit, Impulsen oder Versuchungen zu widerstehen, im NEO-Modell als ein Indikator für Neurotizismus gewertet wird“ (Borkenau & Ostendorf, 1993, S. 28).

Die geringen Korrelationen der Angst-Skalen, eingeschlossen des Angst-CATs, zu den Skalen „Offenheit für Erfahrungen“ und „Verträglichkeit“ sprechen für die diskriminante Validität dieser Skalen.

Die Betrachtung der Korrelationen zum NEO-FFI abschließend, sei angemerkt, dass die hier eruierten Interkorrelationen der fünf Skalen des NEO-FFIs deutlich höher ausfallen als die im Testhandbuch referierten Interkorrelationswerte einer *gesunden* Gesamtstichprobe ($r = -0,33$ und $0,16$; $N = 2.112$; Borkenau &

Ostendorf, 1993, S. 15). Es bleibt offen, ob dieser Befund – falls generalisierbar – eine Besonderheit psychosomatischer Stichproben sein könnte.

Die Korrelationen der Angstinventare zu *fünf* der sechs Skalen des GT, weisen durch geringe Werte ($r = -0,12$ bis $r = 0,18$) auf eine hohe diskriminante Validität zwischen den durch die Skalen erfassten Konstrukten hin. Jedoch bestehen signifikante Korrelationen zwischen der Skala „Soziale Resonanz“ und dem Angst-CAT ($r = -0,21$) bzw. dem BAI ($r = -0,19$). Konstrukttheoretisch reflektiert mag dies daraus resultieren, dass diese Skala unter anderem auch eine „narzisstische Gratifikation“, d. h. positive, soziale, selbstwertstärkende Erlebnisse erfasst (Beckmann et al., 1991, S. 39). Vor dem Hintergrund des bereits oben erörterten negativen Zusammenhangs zwischen Extraversion und Ängstlichkeit erscheint eine geringe negative Korrelation plausibel, wenn man zusätzlich annimmt, dass Extrovertierte (Personen mit einer geringeren *sozialen* Ängstlichkeit) mehr soziale Gratifikation erfahren.

Erwähnenswert ist ferner, dass das Angst-CAT und die anderen Angstinventare nur niedrige, nicht signifikante Korrelationen zur Skala allgemeine „Grundstimmung“ des GT aufweisen, so dass es scheint - obgleich eine Diskrimination zwischen den Konstrukten „Angst“ und „Depression“ schwierig ist -, dass sich Angst als ein spezifisches emotionales Konstrukt von einer allgemeinen Grundstimmung psychometrisch diskriminieren lässt.

Betrachtet man die Interkorrelationen der Skalen des GT, so fällt auf, dass vier der sechs Skalen (Soziale Resonanz, Durchlässigkeit, Stimmung und Soziale Potenz) hoch miteinander korrelieren ($r > 0,6$). Auch die Testautoren (Beckmann et al., 1991) fanden zwischen diesen Skalen Interkorrelationen (bei gesunden und „neurotischen“ Patienten), allerdings in einer deutlich geringeren Ausprägung ($r = -0,24$ bis $r = -0,56$). Wahrscheinlich fordert hier die Schwerpunktlegung bei der Testkonstruktion des GT (Beckmann et al., 1991) auf das Erfassen psychosozialer Aspekte durch eine schlechte Diskriminationsleistung zwischen den einzelnen Konstrukten ihren Tribut.

Fokussiert man abschließend die Korrelationen der Skalen beider Persönlichkeitsinventare (NEO-FFI und GT) untereinander, fällt vor allem ein mittelstark ausgeprägter Zusammenhang ($r = -0,42$ bis $r = 0,34$) zwischen der Skala „Extraversion“ des NEO-FFIs und den „psychosozialen“ Skalen

(„Soziale Resonanz“, „Durchlässigkeit“¹⁰² und „Soziale Potenz“¹⁰³) des GT ins Auge. Dieser erscheint durch eine relative Konstruktnähe verständlich.

6.7.3.2. Diskriminante Validität in Bezug auf das diagnostische Fremdurteil

Wie bereits zur Überprüfung der konvergenten Validität (siehe Kapitel 6.7.2.2.), werden auch zur Überprüfung der diskriminanten Validität die durch den Einsatz des strukturierten diagnostischen Interviews M-CIDI (siehe Kapitel 6.5.3.) an der psychosomatischen Stichprobe erhobenen Diagnosen psychischer Störungen gemäß ICD-10 genutzt. Tabelle 26 gibt einen ersten groben Überblick über die statistischen Kennwerte der Theta-Werte des Angst-CATs der gesamten psychosomatischen Stichprobe (N = 102), welche sich in Patienten mit bzw. ohne Diagnose einer psychischen Störung nach ICD-10 untergliedern lässt. Desweiteren werden als Vergleichsstichprobe die statistischen Kennwerte von einer Gruppe von Medizinstudenten (N = 35) aufgeführt.

Tabelle 26: Statistische Kennwerte der Theta-Werte des Angst-CATs verschiedener Vergleichsgruppen.

Gruppe	N ¹⁰⁴	\bar{X}	SD	SE
Patienten mit der Diagnose einer psychischen Störung (F)	92	,361	,692	,072
Patienten ohne die Diagnose einer psychischen Störung (kein F)	10	,043	,939	,297
Studenten	35	-,932	,791	,134

Eine einfaktorielle Varianzanalyse zeigt, dass das Angst-CAT gut zwischen Patienten mit der Diagnose einer psychischen Störung und ohne eine solche Diagnose (nach ICD-10) bzw. gesunden Personen (Medizinstudenten) zu differenzieren vermag (QS = 42,43; df = 2; \overline{QS} = 21,22; F = 39,08; p ≤ 0,001).

Ogleich das Angst-CAT nicht zur diagnosenspezifischen Differenzierung entwickelt wurde, wurde zusätzlich die Diskriminationsfähigkeit des Angst-CATs bezüglich verschiedener Patientengruppen mit unterschiedlichen Diagnosen einer psychischen Störung untersucht. Hierzu wurden die Patienten der psychosomatischen Gesamtstichprobe (N = 102) mittels der klassifikatorischen

¹⁰² Hohe Werte auf der Skala „Durchlässigkeit“ indizieren emotionale Verschlussenheit.

¹⁰³ Hohe Werte auf der Skala „Soziale Potenz“ indizieren eine geringe soziale Kompetenz.

¹⁰⁴ Die Diagnosegruppengrößen summieren sich nicht zur Gesamtstichprobengröße (N = 102), da eine hohe Komorbidität zwischen den Störungen vorliegt.

Diagnostik des M-CIDI verschiedenen diagnosenspezifischen Subgruppen zugeordnet (siehe Tabelle 27).

Tabelle 27: Statistische Kennwerte der Theta-Werte des Angst-CATs verschiedener Diagnosegruppen (mit Komorbidität).

Gruppe	N ¹⁰⁵	\bar{X}	SD	SE
Patienten mit einer Angststörung (F.40-41.9)	58	,445	,715	,094
Patienten mit einer depressiven Störung (F.32-34)	60	,491	,638	,082
Patienten mit einer Essstörung (F.50)	7	,407	,661	,250
Patienten mit einer somatoformen Störung (F.45)	51	,369	,644	,090

Tabelle 27 zeigt die statistischen Kennwerte der Theta-Werte des Angst-CATs der verschiedenen diagnostischen Subgruppen, wie sie im psychosomatischen Kollektiv geschätzt wurden. Aufgrund einer hohen Komorbidität zwischen den einzelnen Störungsgruppen (somatoforme Störungen & Angststörungen: 70,6%;¹⁰⁶ Essstörungen & Angststörungen: 42,86%;¹⁰⁷ depressive Störungen & Angststörungen: 83,33%)¹⁰⁸ ist die auf den ersten Blick geringe Diskriminationsfähigkeit des Angst-CATs zwischen den einzelnen Diagnosegruppen nicht erstaunlich.

Um zu überprüfen, wie stark die vorliegende Komorbidität des Patientenkollektivs die Diskriminationsfähigkeit des Angst-CATs beeinträchtigt, wurden Patientengruppen ohne Komorbidität gebildet. Die in Abbildung 25 veranschaulichten Patientengruppen bestehen demnach aus Patienten, welche Störungen aus jeweils nur einer Diagnosegruppe aufweisen, d. h. Patienten mit einer Angststörung *und* einer weiteren Diagnose einer psychischen Störung (= Komorbidität) wurden aus den Patientenkollektiven ausgeschlossen.

Abbildung 25 deutet darauf hin, dass das Angst-CAT bei Patienten ohne Komorbidität besser zwischen verschiedenen Störungen zu differenzieren vermag als bei Vorliegen von Komorbidität. Tabelle 28 gibt Aufschluss über die statistischen Kennwerte der Theta-Werte des Angst-CATs dieser verschiedenen diagnostischen Subgruppen nach Ausschluss von Komorbidität.

¹⁰⁵ Die Diagnosegruppengrößen summieren sich nicht zu der Gesamtstichprobengröße (N = 102), da eine hohe Komorbidität zwischen den Störungen vorliegt.

¹⁰⁶ N = 36 von 51 Patienten mit einer somatoformen Störung haben auch eine Angststörung.

¹⁰⁷ N = 3 von 7 Patienten mit einer Essstörung haben auch eine Angststörung.

¹⁰⁸ N = 50 von 60 Patienten mit einer depressiven Störung haben auch eine Angststörung.

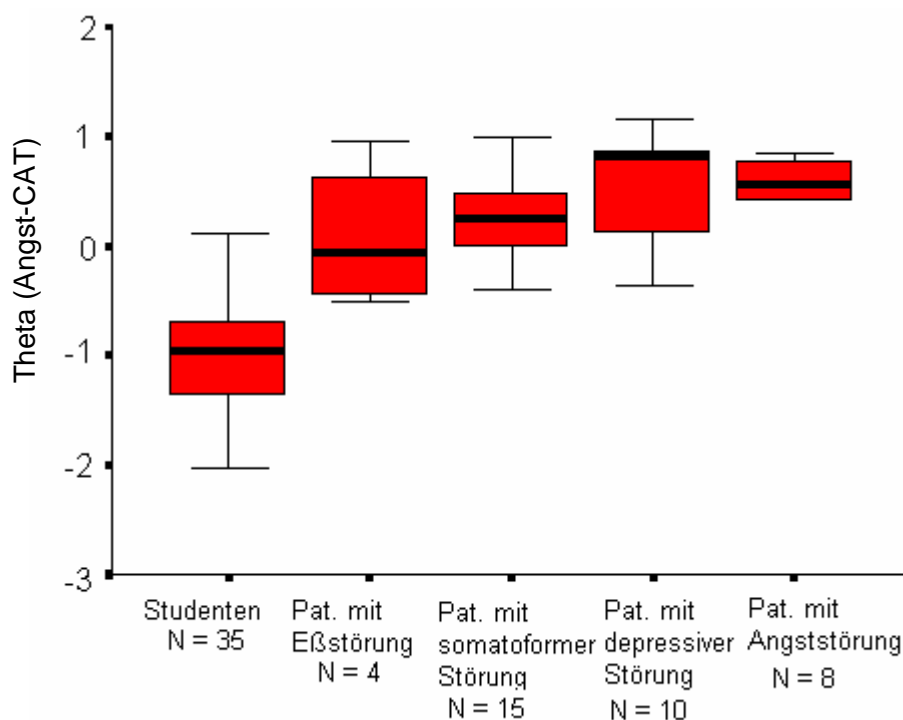


Abbildung 25: Die Theta-Werte-Verteilung des Angst-CATs im Vergleich verschiedener Diagnosegruppen ohne Komorbidität.

Vergleicht man Tabelle 27 (Patientenkollektiv mit Komorbidität) und Tabelle 28 (Patientenkollektiv ohne Komorbidität) so zeigt sich, dass – bei Ausschluss von Komorbidität – die Unterschiede in den Mittelwerten des Angst-CATs verschiedener Patientengruppen deutlicher zu Tage treten. Der zu diskutierende Befund, dass depressive Patienten im Falle von Komorbidität im Durchschnitt etwas höhere Theta-Werte erzielen als ängstliche Patienten (Tabelle 27), wird nach Ausschluss von Komorbidität *nicht* repliziert.

Tabelle 28: Statistische Kennwerte der Theta-Werte des Angst-CATs verschiedener Diagnosegruppen (ohne Komorbidität).

Diagnosegruppe	N ₁₀₉	\bar{X}	SD	SE	95% Intervall für den Mittelwert:	
					Ober-Grenze	Unter-Grenze
Pat. nur mit einer Essstörung (F.50)	4	,091	,675	,337	-,983	1,164
Pat. nur mit einer somatoformen Störung (F.45)	15	,156	,620	,160	-,187	,500
Pat. nur mit einer depressiven Störung (F.32-34)	10	,552	,514	,171	,157	,947
Pat. nur mit einer Angststörung (F.40-41.9)	8	,553	,762	,270	-,085	1,190

Pat.= Psychosomatische stationäre Patienten.

¹⁰⁹ Die Diagnosegruppengrößen summieren sich nicht zur Gesamtstichprobengröße (N = 102), da eine hohe Komorbidität zwischen den Störungen besteht.

Eine über die Gruppen durchgeführte einfaktorielle Varianzanalyse zur globalen Bewertung der Unterschiede in den Theta-Gruppenmittelwerten des Angst-CATs ergibt, dass sich die Gruppen insgesamt auf einem Signifikanzniveau von $p \leq 0,001$ (QS = 30,07; df = 4; $\overline{QS} = 7,52$; F = 14,50) unterscheiden. Ein anschließend durchgeführter Scheffé-Test zur genaueren Untersuchung der Mittelwertsunterschiede zwischen den einzelnen Gruppen zeigt, dass sich die Patienten mit einer somatoformen bzw. depressiven bzw. Angststörung auf einem Signifikanzniveau von $p \leq 0,001$ signifikant von gesunden Personen bzw. der Gruppe der Patienten mit Essstörungen unterscheiden.

Die Unterschiede der Theta-Mittelwerte des Angst-CATs zwischen den drei oben erläuterten Diagnosegruppen (somatoforme, depressive bzw. Angststörung) sind – obgleich sie bei Ausschluss von Komorbidität insgesamt größer sind (siehe Tabelle 27 / 28 im Vergleich) – nicht signifikant. Es bleibt zu diskutieren, ob dies aus den geringen Stichprobengrößen resultiert. Zusammenfassend lässt sich resümieren, dass – obgleich das Angst-CAT nicht dafür konstruiert wurde, verschiedene diagnostische Gruppen voneinander zu trennen – die Ergebnisse in Tabelle 28 und Abbildung 25 – insbesondere die klare Trennung der gesunden Personen von dem psychosomatischen Kollektiv – als Hinweis auf eine gute diskriminante Validität interpretiert werden dürfen.

6.7.4. Zusammenfassung der Validierungsergebnisse

Das Angst-CAT erweist sich in vorliegender Validierungsstudie als ein valides, psychometrisches Verfahren zur Erfassung der Zustands-Angst in einem psychosomatischen Patientenkollektiv.

Eine *mittelmäßige* bis *gute konvergente Validität* des Angst-CATs konnte in Form von mittelmäßig bis hohen Korrelationen zu anderen Angstinventaren (BAI, HADS-A; $r = 0,51$ bis $r = 0,76$) belegt werden. Die Höhe der Korrelationen steht im Einklang mit konvergenten Validierungsergebnissen bereits etablierter Angstinventare ($r = 0,45$ bis $r = 0,86$). Eine konvergente (diagnosenspezifische) Validität ist insofern gegeben, als Patienten mit der Diagnose einer Angststörung signifikant höhere Theta-Werte im Angst-CAT aufweisen als Patienten ohne die Diagnose einer psychischen Störung bzw. gesunde Personen ($p \leq 0,001$).

Die *diskriminante Validität* des Angst-CATs unterscheidet sich im Hinblick auf die untersuchten Konstrukte. Die psychometrische Diskrimination von *Angst* und *Depression* gestaltet sich – wie theoretisch und empirisch in der Literatur ($r = 0,43$ bis $r = 0,62$) bereits vielfach diskutiert – auch im Angst-CAT (BDI, HADS-D; $r = 0,59$ bis $r = 0,60$) *schwierig*.

Dagegen kann aufgrund geringer Korrelationen des Angst-CATs zu Skalen von zwei Persönlichkeitsinventaren (NEO-FFI, GT) auf eine *gute* diskriminante Validität bezüglich *anderer Eigenschaftskonstrukte* geschlossen werden. Das Konstrukt der Angst lässt sich von einem allgemeinen Konstrukt der Grundstimmung ($r = 0,12$) und allen weiteren Skalen der Persönlichkeitsinventare ($r = -0,21$ bis $r = 0,12$) gut differenzieren.

Einzige Ausnahme ist die Diskrimination zum Konstrukt „*Neurotizismus*“, welche – angezeigt durch eine Korrelation von $r = 0,63$ – nicht gelingt, und eine Korrelation des Angst-CATs mit den Skalen „*Extraversion*“ und „*Gewissenhaftigkeit*“ ($r = 0,3$) mitbedingt. Dieser Befund steht in Einklang mit Forschungsbefunden aus der Literatur und ist in die Forschungsdiskussion um eine mögliche Differenzierbarkeit zwischen einer Eigenschafts- und einer Zustands-Angst einzuordnen (zur Diskussion siehe Kapitel 7.5.1.).

Obgleich das Angst-CAT nicht zur diagnosenspezifischen Diskrimination entwickelt wurde, legen die berichteten Ergebnisse nahe, dass eine diskriminante Validität bezüglich verschiedener Diagnosegruppen bedingt gegeben ist. Eine Differenzierung zwischen verschiedenen Diagnosegruppen ist tendenziell möglich, jedoch nur bei Patienten, welche keine Komorbidität aufweisen. Insofern sollte das Angst-CAT stets im Zusammenhang weiterer klinischer Diagnostik interpretiert werden.