

---

## 4. Diskussion

### 4.1. Ziele der eigenen Arbeiten

Die Arbeiten, die hier zur Habilitation vorgelegt werden, befassen sich mit der Entwicklung und Verbesserung psychologisch diagnostischer Verfahren in der Integrierten Psychosomatik und ihren angrenzenden medizinischen und gesundheitspsychologischen Disziplinen. Ein gemeinsames Ziel aller Arbeiten war es, Instrumente zu entwickeln, zu optimieren sowie Referenzwerte für ausgewählte Krankengruppen zu ermitteln. Inhaltlich sollten die Verfahren ausgewählte Konstrukte aus den Bereichen gesundheitsbezogene Lebensqualität, Stress und Ressourcen erfassen, die in vielfältige psychosomatischen und gesundheitspsychologischen Erklärungsmodellen eingegangen sind und die für die Therapieevaluation und Prognose herangezogen werden. Im stärker auf die Klinik bezogenen Teil der Arbeiten wurde die Diagnostik autodestruktiver Syndrome vertieft.

Die Entwicklung testpsychologischer Verfahren für die Klinik war – neben inhaltlichen Gesichtspunkten – an mehreren in der aktuellen Forschung relevanten Desideraten orientiert. Es galt, Verfahren, die sich in angloamerikanischen Studien bewährt und sich in der internationalen Diskussion behauptet haben, für den deutschen Sprachraum zu übertragen und zu validieren, um Instrumente für Studien mit dem Anspruch internationaler Vergleichbarkeit zur Verfügung zu haben. Mit Blick auf die Klinikroutine galt es, den Umfang psychologischer Fragebögen und Testbatterien so zu kürzen, dass der Bearbeitungsaufwand für Patienten in einem noch zu rechtfertigenden Maß bleibt und die Qualität der Daten nicht durch Ermüdungseffekte oder motivationale Probleme gemindert wird. Für KTT-basierte Instrumente ist bekannt, dass Reliabilität und dimensionale Struktur stark stichprobenabhängig sind. Daher sollten KTT-basierte Instrumente nicht nur auf ihre Gütekriterien, sondern auch daraufhin überprüft werden, ob sich ihre dimensional Strukturen nach Kürzung und in neuen Stichproben replizieren lassen. Schließlich wurden neue Verfahren entwickelt, die die Vorteile Probabilistischer Testtheorie beim computer-adaptiven Testen nutzen sollten.

## 4.2 Funktionen psychologischer Diagnostik

In der Einleitung wurde skizziert, dass psychologische Diagnostik in der Psychosomatik und ihren Nachbardisziplinen grundsätzlich fünf Funktionen erfüllen kann: *Deskription*, *Klassifikation*, *Erklärung*, *Prognose* und *Evaluation* [18;19]. Es soll kurz erörtert werden, inwieweit die hier vorgestellten diagnostischen Verfahren diese Funktionen erfüllen.

*Deskription.* Alle Verfahren, die in den eigenen Arbeiten entwickelt oder weiterentwickelt wurden, erfüllen in gewissem Umfang die Funktion der Deskription des jeweiligen Merkmals. Die psychometrischen Fragebögen und Tests (ALL, SWOP, PSQ, URS-CED, D-CAT, A-CAT) richten sich auf die standardisierte und ökonomische Erfassung der Ausprägung eines psychologischen oder psychopathologischen Konstruktes, wie zum Beispiel Stress, Selbstwirksamkeit oder Depression. Für die KTT-basierten Instrumente lässt sich der ermittelte Testwert in Bezug zu einer der verfügbaren Referenzgruppen bringen. Für die IRT-basierten Verfahren steht der Testwert in Bezug zu einem latenten Merkmalskontinuum, das stichprobenunabhängig angewandt werden kann.

Die Instrumente zur Erfassung von Selbstschädigung sind noch stärker deskriptiv ausgerichtet. Die beiden Selbstbeurteilungsbögen DSHI und SHBQ sollen keine abstrakten psychologischen Konstrukte erfassen. Sie sollen vielmehr im Sinne von Registrierverfahren das Vorliegen konkreter Verhaltensweisen dokumentieren (z.B. sich selbst schneiden, brennen, verätzen). Diese Verhaltensweisen entstammen zwar einer übergeordneten Verhaltenskategorie (selbstverletzendes Verhalten im DSHI [108] bzw. selbstverletzendes und suizidales Verhalten im SHBQ [109]). Von ihnen wird jedoch nicht auf ein abstraktes psychologisches Konstrukt und seine Ausprägung geschlossen. Die Verhaltensweisen werden lediglich in ihrer Bandbreite und Anzahl erfasst. Relevant kann dies in vielen Kontexten sein, unabdingbar ist es unter epidemiologischen Gesichtspunkten. Während insbesondere aus britischen und auch US-amerikanischen Studien differenzierte Ergebnisse zur Verbreitung selbstverletzenden Verhaltens in der Bevölkerung und in unterschiedlichen Gruppierungen vorliegen (z.B. [112-

117]), sind solche Daten für Deutschland und andere europäische Staaten gar nicht oder nur in sehr geringem Umfang verfügbar.

In vergleichbarer Weise ist der Fremdbeurteilungsbogen Selbstschädigung ein klinisches Dokumentationsschema, das der systematischen Registrierung und Abbildung bestimmter Verhaltenstypen, ihrer Ausprägung und ihrer Vorgeschichte dient, ohne dass das Instrument den Rückschluss auf ein abstraktes Konstrukt implizieren soll.

*Klassifikation.* Gerade die Instrumente zur Erfassung von autodestruktivem Verhalten können aber zugleich eine *klassifikatorische Funktion* erfüllen. Eine differenzierte Deskription des Merkmals bildet hierfür die Voraussetzung. Die beiden Selbstbeurteilungsverfahren „Deliberate Self-harm Inventory“ [108] und „Self-harm Behavior Questionnaire“ [109] liefern nicht nur Angaben darüber, ob selbstverletzendes oder suizidales Verhalten vorliegt oder nicht, sie helfen mit ihrer differenzierten Abbildung typischer Formen selbstschädigenden Verhaltens bei der Klassifikation. In der Literatur gibt es zwar bereits Vorschläge zur Klassifikation selbstschädigenden Verhaltens [118;119], jedoch noch einen Mangel an Studien, die solche Vorschläge mit Hilfe standardisierter und validierter Erhebungsverfahren empirisch unterlegen. Die hier validierten Instrumente können auch für solche Fragestellungen herangezogen werden.

Der klinische Fremdbeurteilungsbogen zur Selbstschädigung [56] kann ebenfalls eine klassifikatorische Funktion erfüllen, auch wenn er streng genommen keinen klassifikatorischen, sondern einen typologischen Ansatz verfolgt. Dazu ist anzumerken, dass bislang kein allgemein akzeptierter Ansatz zur Klassifikation autodestruktiven Verhaltens vorliegt. Hinsichtlich der gängigen Klassifikationssysteme psychischer und Verhaltensstörungen, ICD-10 Teil F und DSM IV, sehen verschiedene Autoren Entwicklungsbedarf [56;100;120] – eine Sichtweise, die der Autor dieser Schrift teilt. Für die Kritik spricht auch, dass offen selbstverletzendes Verhalten diesen Systemen folgend als Kriterium einer so genannten Borderline-Persönlichkeitsstörung gefasst wird und allenfalls noch als Symptom einer Belastungsstörung (im DSM) genannt wird, obgleich empirische Studien [116;118], klinische Beobachtungen [121] und auch die eigenen Befunde dafür sprechen, dass

Selbstverletzung auch vorkommen kann, ohne dass die Person eine Borderline-Persönlichkeitsstörung nach den einschlägigen Kriterien aufweisen muss.

Der typologische Ansatz der DKPM-Arbeitsgruppe [55;56;101] eignet sich meines Erachtens als ein Ausgangspunkt für die Bildung einer übergeordneten Kategorie „Autodestruktionssyndrome“, bei der direkte Selbstverletzung, artifiziell induzierte Erkrankung und vorgetäuschte Symptome (mit an das Medizinalsystem delegierter Autoaggression) relevante Subtypen sein könnten.

Ein weiteres Instrument, das unter einer klassifikatorischen Fragestellung mit eingesetzt werden kann, ist der Depressions-CAT. Zwar werden hier keine diagnostischen Kriterien für die Diagnose einer Depression nach einem gängigen Klassifikationsschema erfasst. Aber wie in klassischen Selbstbeurteilungsverfahren depressiver Symptome [122;123] lässt sich für den Depressions-CAT ein klinisch relevanter Cut-Off-Wert bestimmen. Teilt man eine gemischt psychosomatisch-internistische Patientengruppierung mittels dem standardisierte klinischem strukturierten Interview DIA-X nach ihrer DSM-Diagnose in Depressive und Nicht-Depressive ein, so leistet der Depressions-CAT eine treffsichere Zuordnung zu einer dieser Gruppen mit 92% Sensitivität und 78% Spezifität [124]. Damit eignet sich das Instrument dazu, eine klinische Diagnosestellung in ihrer Treffsicherheit zu stützen.

*Erklärung.* In aktuellen psychosomatischen Modellen zur *Erklärung* von Krankheitsentstehung und –aufrechterhaltung nimmt Stress eine zentrale Rolle ein. Dieses lässt sich mit Hilfe des „Perceived Stress Questionnaire“ reliabel und valide bestimmen. Er wurde für den deutschen Sprachraum adaptiert und validiert. Mit dem Original-Instrument ließ sich beispielsweise erhöhtes Stresserleben als ein prognostisch ungünstiger Faktor einer chronisch entzündlichen Darmerkrankung nachweisen [89]. Mit dem hier validierten PSQ konnten unter anderem Zusammenhänge zwischen erhöhtem Stresserleben und immunologischen Imbalancen bei Frauen, die eine Fehlgeburt erlitten hatten, belegen [125]. Das Instrument ist derzeit u.a. in prospektiven Studien zu Stress und Schwangerschaft im Einsatz.

*Prognose* und *Evaluation*. In den Studien zur Evaluation stationärer psychosomatischer Behandlung ließen sich Verbesserungen sowohl für subjektive Erfolgsparameter (Körperbeschwerden, Stimmungsqualitäten, Lebenszufriedenheit, globale Lebensqualität, Alltagsbewältigung) als auch für objektivierbare Parameter (Krankenhaustage, Arbeitsunfähigkeitszeiten) nachweisen. Änderungen in den subjektiven Parametern erfolgten zeitlich unmittelbar. In den Katamnesen zeigten sich sowohl stabile Verbesserungen als auch solche, die anschließend wieder leicht oder auch deutlich rückläufig waren. Die Änderungen in den objektiven Parametern erfolgten teilweise deutlich zeitversetzt zueinander und zu den subjektiven Parametern. Subjektive und objektive Parameter waren nahezu unkorreliert. Dies spricht dafür, dass sie je spezifische Merkmalsbereiche erfassen und in der Therapieevaluation nicht wechselseitig ersetzt werden können.

Die generalisierten Erwartungshaltungen Optimismus, Pessimismus und Selbstwirksamkeit, wie sie mit dem SWOP erfasst wurden, erwiesen sich sowohl für Änderungen in den subjektiven als auch in den objektiven Parametern als prognostisch relevante Faktoren.

Insbesondere für die Veränderungsmessung im Rahmen der Evaluation psychosomatischer Behandlung stellt sich allerdings ein messtheoretisches Problem für Instrumente, die auf der KTT basieren. Diese vereinen nämlich die jeweiligen Messfehler mehrerer Zeitpunkte auf sich und sind somit in der Veränderungsmessung weniger reliabel als bei Messungen zu nur einem Zeitpunkt [126]. Bisher ist keine Studie bekannt, in der klinische Änderungen, etwa die Effekte unter psychosomatischer Therapie, mit Hilfe von IRT-basierten Verfahren erfasst worden wären. In dieser Hinsicht besteht also noch einiges Potential für die Verbesserung psychometrischer Diagnostik.

### **4.3 Datenebenen und Datenquellen**

Die je spezifischen Resultate bei einer Mehrebenenerfassung in der Therapieevaluation wurden bereits im vorangegangenen Abschnitt erwähnt.

Selbstauskünfte und Fremdbeurteilungen über den gleichen Merkmalsbereich miteinander zu kombinieren, wurde in der Studie zu autodestruktiven Syndromen realisiert. Dabei fand sich, dass die Selbst- und Fremdbeurteilungsverfahren zwar in einem ausreichenden Umfang übereinstimmen, um auf Konstruktvalidität der Instrumente schließen zu lassen, aber auch je spezifische Ergebnisse erbrachten. So gab es sowohl Patienten, die zwar im Fragebogen selbstverletzendes Verhalten berichteten, dies aber im Arzt-Patient-Gespräch offenkundig nicht mitteilten, als auch solche, die zwar selbst kein selbstschädigendes Verhalten berichteten, für die ein solches aber durch den klinischen Diagnostiker konstatiert wurde. Auch wenn man einen Messfehler (für beide Datenquellen) in Rechnung stellt, so ist darüber hinaus anzunehmen, dass jede Datenquelle ihre spezifischen Vorteile und auch Nachteile mit sich bringt. So könnten die Selbstauskunftsverfahren für autodestruktives Verhalten dort eine bessere „Trefferquote“ haben, wo es den Betroffenen leichter fällt, sich in einer schriftlichen Weise, ohne den direkten Kontakt zu einer anderen Person zu offenbaren. Dies scheint selbst dann der Fall zu sein, wenn diese andere Person in die Behandlung involviert ist und die Patienten davon ausgehen können, dass der Behandler die Angaben im Fragebogen liest, also keine im eigentlichen Sinne anonyme Messung vorliegt. Fremdbeurteilungsverfahren hingegen könnten besonders relevant in Fällen sein, in denen die Autodestruktion durch die Betroffenen verheimlicht wird und in denen aus klinisch-anamnestischen Informationen und gezielten Beobachtungen eindeutige Hinweise auf ein solches Verhalten gewonnen werden können. Zusätzlich können Fremdbeurteilungsbögen in der Klinikroutine die Aufgabe erfüllen, den diagnostischen Blick der Behandler zu schärfen. Unterschiede in den diagnostizierten Quoten über die beiden Jahre der Studie – mit und ohne gezielte Supervision – verweisen darauf, dass die gelenkte Aufmerksamkeit der Diagnostiker eine Rolle für die Wahrnehmung dieser Phänomene spielt – wenngleich nicht geklärt ist, ob damit auch der Anteil falsch positiver Diagnosen ansteigt.

Darüber hinaus bestätigen die eigenen Befunde zu den Selbstbeurteilungsverfahren DSHI und SHBQ und dem Fremdbeurteilungsbogen frühere Befunde, wonach die Übereinstimmung

---

zwischen Selbst- und Fremdbeurteilung mit fortschreitendem Behandlungsverlauf größer wird [22].

#### 4.4 Fazit und Ausblick

Zusammenfassend lassen die eigenen Studien in Hinblick auf psychometrische Diagnostik für die Psychosomatische Medizin und benachbarte Disziplinen folgende Schlüsse zu:

- International etablierte psychometrische Verfahren für ausgewählte klinische und gesundheitspsychologische Konstrukte (am Beispiel Stresserleben, autodestruktives und suizidales Verhalten) ließen sich für den deutschen Sprachraum adaptieren und anhand klinischer Stichproben sowie gesunder Vergleichsgruppen validieren.
- Für die Adaptation KTT-basierter Instrumente war es erforderlich, die psychometrischen Gütekriterien neu zu belegen und die Stabilität der dimensionalen Struktur zu prüfen. Letzteres konnte am Beispiel des Stressfragebogens mit Hilfe von linearen Strukturgleichungsmodellen erfolgen, wozu sich konfirmatorische Faktorenanalysen und Multi-Sample-Analysen zur Absicherung der Stichprobeninvarianz eigneten.
- Im Zuge der dimensionalen Überprüfung und Restrukturierung ließen sich KTT-basierte Verfahren anhand item-statistischer Merkmale kürzen und damit ökonomisieren, ohne dass die Messpräzision nennenswert beeinträchtigt wurde.
- Insbesondere auf der Grundlage probabilistischer Testtheorie und mit Hilfe eines computer-adaptiven Messalgorithmus ließ sich eine ökonomisierte und zugleich präzise Messung realisieren. Am Beispiel der Depressionsmessung konnte gezeigt werden, dass die Anwendung der IRT zu einer präziseren Merkmalsschätzung in den Extrembereichen des Merkmalskontinuums führte als dies mit Hilfe KTT-basierter Instrumente der Fall war.

- Informationen auf mehreren Ebenen und aus mehreren Quellen zu erfassen, wurde in der katamnestischen Evaluation stationärer psychosomatischer Therapie realisiert. Dass subjektive Parameter (Körperbeschwerden, Stimmungsmerkmale, Lebenszufriedenheit, Alltagsbewältigung) und objektivierbare Merkmale (Krankenhaustage, Arbeitsunfähigkeitszeiten) kaum korrelierten und zeitlich versetzte Verbesserungen zeigten, belegt deren je spezifische Relevanz für die Therapieevaluation.
- Anhand der Diagnostik autodestruktiver Syndrome wurde gezeigt, dass Fremdbeurteilungsverfahren und Selbstbeurteilungsverfahren zwar Übereinstimmungen aufweisen, aber nicht deckungsgleich sind. Abgesehen von den etwas unterschiedlichen inhaltlichen Zielsetzungen der Instrumente gibt es bei diesem Störungsspektrum offenkundig Informationen, die sich eher aus der klinischen Beobachtung erschließen, wie auch solche, die eher im schriftlichen Fragebogen als im Arzt-Patienten-Kontakt mitgeteilt werden.

Als ein Fazit lässt sich also festhalten, dass Messgenauigkeit und Ökonomie psychometrischer Instrumente zwar auch im Rahmen KTT-basierter Methodik gesichert werden können, dass aber probabilistischer Messmodelle unter Anwendung computer-adaptiver Testalgorithmen hierfür erhebliche Vorteile mit sich bringen. Ihre Nutzung sollte weiter ausgebaut und auch in der Veränderungsmessung ausgetestet werden. Für komplexe klinische Fragestellungen, wie die Diagnostik teils offen gehandhabter, teils verheimlichter autodestruktiver Syndrome oder die Evaluation psychosomatischer Behandlung, bringt Multimodalität, also die Einbeziehung unterschiedlicher Datenquellen und Datenebenen, keine redundante, sondern je spezifische Information und sollte – wo immer möglich – realisiert werden.