

Von A wie Angst bis Z wie Zwang – Störungsspezifische Messinstrumente zur Evaluation von Psychotherapie (Erwachsene)

Edgar Geissner & Stefan Koch

Zusammenfassung: Psychometrische Verfahren zur Evaluation psychotherapeutischer Behandlungen werden immer mehr zum Standard und sichern die Bestimmung des Erfolgs eigener Maßnahmen ab – neben Patientenbericht und klinischem Urteil. Sie werden überdies vom Gesetzgeber zunehmend verlangt; und Kammern sowie Gremien haben diesbezüglich Empfehlungen vorgelegt. Nach einem Kriterienkatalog haben wir die Landschaft klinischer Verfahren, die die Bedingung „störungsspezifisch“ und „veränderungssensitiv“ erfüllen, gesichtet, eine Charakterisierung und Bewertung durchgeführt und möchten für die folgenden neunzehn ICD-F-Diagnosen Vorschläge zum praktischen Einsatz (Erwachsenenbereich) machen: Agoraphobie/Panikstörung, Alkohol, ADHS im Erwachsenenalter, Bipolare Störung, Borderline-Persönlichkeitsstörung, Computerspielabhängigkeit, Depression, Dissoziative Störungen, Essstörungen, Hypochondrie, Posttraumatische Belastungsstörung, Psychosen, Schlafstörungen, Schmerz, Sexuelle Funktionsstörungen, Somatoforme Störungen, Soziale Phobie, (Anhaltende) Trauerstörung und Zwangserkrankungen. Anmerkungen zu geeigneten Messzeitpunkten, zur Honorierung und zu künftigen Entwicklungen runden den Beitrag ab.

Einleitung

Die Evaluation durchgeführter Psychotherapien mittels psychometrischer Verfahren ist nicht nur sinnvoll und lohnenswert, sondern wird auch zunehmend geradezu verbindlich gefordert (Kasten 1). Über alle durchgeführten Behandlungen eines Jahres hinweg bildet sie eine Bilanzierung der eigenen Maßnahmen ab – i. S. von Qualitätssicherung. Ergebnisse dienen auch als individuelles (Zwischen-)Fazit in der Besprechung mit Patient*innen, sind insofern Bestandteil der Psychotherapie. Darüber hinaus haben sie ihren Stellenwert in der Objektivierung des Behandlungsverlaufs an der Schnittstelle beispielsweise von ambulanter und stationärer Versorgung bzw. weiterführender Maßnahmen.

Hier mag manch eine*r einwenden, dass Veränderungen im Verlauf einer Therapie auch so sichtbar sein sollten – eine Wertung, die nicht immer richtig sein muss, denn (auch) Psychotherapeut*innen können Fehleinschätzungen unterliegen: So stuften sie beispielsweise in einer Studie die Therapiezielerreichung ihrer Patient*innen am Behandlungsende zu 45 % höher ein als diese selbst (Geissner et al., 2010) oder missdeuteten – in einer anderen Untersuchung – zu 40 % eine eigentliche Verschlechterung als Symptomverbesserung (Lutz et al., 2019).¹

Eine Einschätzung durch die Patient*innen ist also empfehlenswert, kann ein wichtiges Korrektiv sein, einen zu positiven globalen Eindruck zurechtzurücken und sichert die Aussagekraft eigener Therapieerfolge.

Glücklicherweise verfügt die Psychologie über ein Arsenal an Instrumenten, um Erfolg (serhaltung) messbar zu machen. Weniger sind hierbei reine Screening-Instrumente zu Beginn der Therapie gemeint, vielmehr Messverfahren, mit deren Hilfe Vergleiche zwischen zwei oder mehr Zeitpunkten möglich sind, die also Veränderungsaussagen erlauben. Solche Verfahren dürfen somit nicht traitorientiert konzipiert sein, sondern benötigen einen umgrenzten Zeitrahmen (für Psychotherapie geeignete Einheiten wie z. B. „letzte Woche“ oder „im letzten Monat“). Lutz et al. (2019) unterscheiden hierbei direkte und indirekte Verfahren. Die direkten Verfahren thematisieren Veränderung explizit und erfragen z. B. „im Vergleich zu früher sehe ich meine Probleme heute in einem anderen Licht“. Sie sind den Autoren zufolge nicht mehr aktuell, denn die Gefahr der Überschätzung eigener Bemühungen liegt dort nahe: Welche*r Patient*in sieht nach Wochen der Behandlung die Probleme nicht „in einem anderen Licht“? Indirekte Verfahren thematisieren dagegen Veränderung nicht explizit. Die Untersuchung besteht hier aus der zwei- oder mehrmaligen Vorgabe des gleichen Instruments und dem Vergleich der dabei erhaltenen Resultate. Sie entsprechen dem gegenwärtigen methodischen Stand besser und bieten sich insofern für die Outcome-Bestimmung an. Eine kleine Zusammenstellung der oben erwähnten offiziellen Regelungen und Empfehlungen bezüglich Dokumentation und Evaluation findet sich in Kasten 1.

¹ Zu den hier kursiv ausgewiesenen Kurztiteln finden Sie bibliographische Angaben am Ende des Artikels, das vollständige Literaturverzeichnis auf der Homepage der Zeitschrift unter www.psychotherapeutenjournal.de.

- Psychotherapeutengesetz vom 15. November 2019: In § 7 Absatz 3 Ziffer 3 wird ausgeführt, dass bei der Erteilung der Approbation eine Befähigung gegeben sein muss, eigene oder von anderen durchgeführte Maßnahmen der psychotherapeutischen Versorgung zu dokumentieren und zu evaluieren (Bundesgesetzblatt 2019, S. 1604 ff.).
- Approbationsordnung, Bundesministerium für Gesundheit am 4. März 2020, Anlage 2, Ziffer 5: „Bei der Zulassung zur psychotherapeutischen Prüfung sind Kenntnisse in Dokumentation, Evaluierung und Organisation psychotherapeutischer Behandlungen nachzuweisen“ (Bundesgesetzblatt 2020, S. 474)
- Berufsordnung, insbesondere § 9 Absatz 2: „Psychotherapeutinnen und Psychotherapeuten sind verpflichtet, ... wesentliche Maßnahmen und deren Ergebnisse aufzuzeichnen insbesondere Anamnese, Diagnosen, Untersuchungen, Untersuchungsergebnisse, Befunde, Therapien und ihre Wirkungen...“ (PTK Bayern, 2014, S. 12).
- Musterweiterbildungsordnung, Abschnitt „Tätigkeitsprofile als Ziele der Weiterbildung“: „Fachpsychotherapeut*innen sind an ... Maßnahmen zu Therapieverlaufs- und -erfolgskontrolle beteiligt“ (38. Deutscher Psychotherapeutentag 24. April 2021, Dokument „Tätigkeitsprofile“, S. 4).
- Im Einklang hierzu Empfehlungen am 37. Deutschen Psychotherapeutentag 2020, Stichworte: „Untersuchungsergebnisse“ (S. 3), „eingesetzte Testverfahren“ (S. 6) bzw. „Besprechung erreichter Veränderungen am Therapieende gemeinsam mit dem/der Patient/in“ (S. 9).
- Einzelne Kammern: Z. B. PTK Bayern – Mitgliederrundschreiben 1, 2021, S. 8–9; PTK Berlin – Psychotherapeutenjournal 3/2020, S. 292; PTK Hessen – Psychotherapeutenjournal 3/2020, S. 303 f.
- Deutscher Ethikrat in „Patientenwohl als ethischer Maßstab für das Krankenhaus“: Betonung der Notwendigkeit von Prozessqualität, Ergebnisqualität und Qualitätsmessung im Sinne sowohl subjektiver als auch objektiver Indikatoren im Behandlungsverlauf (Deutscher Ethikrat 2016, S. 51–53).

Kasten 1: Auswahl offizieller Regelungen bzw. Empfehlungen zu Dokumentation und Evaluation

Evaluation regulärer Psychotherapien ist noch nicht lange Thema jenseits von forschungsbezogenen Studien (zur praxisgeeigneten Auswertung solcher „naturalistischen Settings“: Hiller et al., 2009, 2011). Dabei geht es weniger um unspezifische, oft nur formale Zufriedenheitsindizes einer klinischen Behandlung (Schmidt et al., 1989; Schmidt, 2008) als vielmehr um die Feststellung des Therapieergebnisses orientiert an der jeweiligen Störung. Vertreter*innen der Richtlinienverfahren haben diesen Gedanken – auch konkret in Beiträgen im Psychotherapeutenjournal – aufgegriffen und sich ihrerseits für diese Art der Dokumentation ausgesprochen, etwa Stangier (2019; VT), Strauß (2019; TP) oder Schiepek et al. (2019; systemisch).

Das Routine Outcome Monitoring i. S. einer Sitzung-für-Sitzung-Erhebung (Lutz et al., 2019, S. 38), stößt in der Versorgungspraxis sicher noch an Grenzen, aber Erhebungen vor

Beginn und nach Abschluss (bei längeren Behandlungen auch zwischendurch) scheinen machbar und für Patient*innen zumutbar, siehe hierzu auch den Abschnitt „Untersuchungszeitpunkte“ im Diskussionsteil. Zumutbar insbesondere auch dann, wenn nicht ganze Fragebogen-Batterien verabreicht werden, sondern gezielt nur dasjenige Verfahren zum Einsatz kommt, welches inhaltlich sinnvolle quantitative Kennwerte zum Vorher-Nachher-Störungsausmaß ermöglicht. Damit ist für die gute Evaluationspraxis ein wesentlicher Schritt getan, ohne dass eine Anzahl „auch interessanter“, aber letztlich weniger zentraler Merkmale miterhoben würde.

Wir möchten im Folgenden zu den wichtigen psychischen Störungen „von A wie Angst bis Z wie Zwang“ jeweils ein typisches, wissenschaftlich bewährtes und für die klinische Anwendung gut geeignetes Verfahren zur Evaluation von Psychotherapie – Erwachsenenbereich – vorschlagen (vgl. Tabelle auf S. 319). Wir decken damit die ICD-Gruppen F1 bis F5 ab und stellen aus der Gruppe F6 ein Messinstrument – zum Thema Borderline-Persönlichkeitsstörung – vor. Dazu haben wir einige Beurteilungsgesichtspunkte formuliert (siehe Kasten 2) und diese der Auswahl und Beschreibung zugrunde gelegt. Wesentlich waren zusammengefasst vor allem drei Punkte: 1. Die Passung des jeweiligen Untersuchungsinstrumentes spezifisch auf die betreffende Störung musste gegeben sein – was sich bei der Sichtung manchmal als nicht selbstverständlich erwies. 2. Auch musste eine Eignung im Sinne der Beurteilung von Therapiefortschritt, der Reduktion im Ausmaß der Symptomatik, bestehen (Veränderungssensitivität). 3. Schließlich sollten die Messinstrumente (in Bezug auf die grundlegenden Psychotherapie-Verfahren) schulübergreifend einsetzbar sein.²

Unsere Recherchen ergaben für die gesichteten Verfahren hinsichtlich Entwicklungs-, Untersuchungs- und Analyseaufwand eine große Bandbreite – von methodisch hochsophistiziert und mit guter Stichprobenbasis bis hin zu reinen Übersetzungen aus dem Englischen und/oder mit noch relativ geringer Datenbasis. Dessen ungeachtet sind die vorgestellten Messinstrumente in ihrem jeweiligen Bereich nach derzeitigem Stand fachlich als „first choice“ einzustufen und für die Psychotherapieevaluation zu empfehlen.

Für viele Untersuchungsinstrumente besteht darüber hinaus eine konsensuelle Verständigungsgrundlage aus Fachgesellschaften/Arbeitsgruppen, die sich inhaltlich der jeweiligen Störung widmen; teils existieren auch Empfehlungen der WHO, Nennungen in den AWMF-Leitlinien oder gesicherte Erfahrungswerte aus publizierten Therapievergleichsstudien.

² Die Autoren weisen darauf hin, dass handlungsleitend bei der Texterstellung der (internationale) wissenschaftsbasierte Erkenntnisstand klinisch-psychologischer Diagnostik/Testentwicklung war. Eine Verknüpfung mit etwaigen Themen aktueller berufspolitischer Diskussion stand daher nicht im Vordergrund, darüber hinaus auch keine Wertung der vorgestellten Verfahren zu wirtschaftlichen, lizenzrechtlichen oder honorierungsbezogenen Aspekten.

- Enger inhaltlicher Bezug zur Störung. Erfahrungswerte im Einsatz bei *Psychotherapiepatient*innen* vorhanden – nicht bei (studentischen) Freiwilligen aus Analogstudien.
- Eignung für Wiederholungsmessungen. Kein reines Screening-Instrument. Umgrenzter Zeitrahmen für die Beurteilung eingetretener Veränderungen in/nach Psychotherapie. Keine Trait-Instrumente (Persönlichkeitstests).
- Dem aktuellen wissenschaftlichen Stand entsprechend, keine Aufnahme allein aufgrund historischer Verdienste.
- Relevante Verbreitung des Verfahrens bei der jeweiligen Störung. In der Praxis bewährt. In Publikationen als anerkannt erwähnt.
- In der Anwendung: Gute Einsetzbarkeit, fair, transparent, übersichtlich, kurz genug, die allgemeinen Praxisabläufe nicht zu sehr beeinträchtigend.
- Selbstrating-Verfahren (Anm.: In einem Fall nicht gegeben, s. u.). Items von den Patientinnen als prinzipiell nachvollziehbar und erkennbar mit der Störung im Zusammenhang stehend verstanden.
- Aus psychometrisch-testtheoretischer Sicht (dimensionaler Aufbau, Item- und Skaleneigenschaften) vertretbar. Validitätsbefunde gegeben, ebenso Hinweise zur Veränderungsbeurteilung.
- Unaufwändige Auswertung, klare Ergebnisinterpretation.
- Normen bzw. klinische Richtwerte vorhanden. Referenzwerte aus Kriteriumsgruppen, evtl. auch Vergleichswerte von Personen ohne die Störung.
- Kompatibilität mit der klinischen Klassifikationssystematik der Gesundheitsversorgung.
- Übergreifende Eignung für die PT-Richtlinienverfahren.
- Ein charakteristischer Vertreter für den jeweiligen Bereich – nicht „alles, was es hierzu gibt“.

Kasten 2: Gesichtspunkte bei der Auswahl störungsbezogener Verfahren zur Evaluation

Hinzuweisen ist an dieser Stelle, dass ergänzend zum Text des Beitrags ein ausführliches elektronisches Supplement erstellt wurde, in welchem alle Angaben zu den nachfolgenden Messverfahren aufgeführt sind (siehe Hinweis am Ende des Artikels).

Die störungsspezifischen Messinstrumente „von A bis Z“

Agoraphobie/Panikstörung

Das Messverfahren *AKV – Aengste, Kognitionen, Vermeidung* (Ehlers et al., 2001) besteht aus den drei Fragebogenteilen Körperliche Symptome der Angst (*BSQ – Body Sensations Questionnaire*, die englische Bezeichnung ist im Deutschen so beibehalten), Angstbezogene Kognitionen (*ACQ – Anxiety*

Cognitions Questionnaire) und Mobilitätseinschränkung aufgrund der Ängste (*MI – Mobility Index*). Es umfasst 18 bzw. 15 bzw. 28 Aussagen. Der Fragebogen *MI* soll in zwei Varianten bearbeitet werden: Mobilitätseinschränkung „wenn allein“ und „wenn in Begleitung“. Die Angstkognitionen ließen sich in Skalenanalysen nochmals in „körperliche Krise“ und „Kontrollverlust“ gliedern, in der Praxis werden sie jedoch aufsummiert. Einsatzbereich sind die ICD-Diagnosen F40.0, F40.2 und F41.0. Zahlreiche Studien belegen die drei Komponenten des *AKV* als valide und veränderungssensitiv. Auch die Differenziertheit bei gleichzeitig guter Handhabbarkeit spricht für das Instrumentarium. Das Manual bietet sämtliche datenbezogenen Hinweise zu Skalen und Items. Die Eignung zur Outcome-Messung in der Praxis wird gezeigt. Normen (Stanine mit $M = 5$ und $SD = 2$; auch Prozentränge) basieren auf über 1.200 Agoraphobie-/Panikstörungspatient*innen, überdies auch vergleichend auf 400 anderen Angstpatient*innen. Die Bearbeitungszeit beträgt beim ersten Mal ca. 20 Min. (später deutlich kürzer).

Alkohol

Aus den Empfehlungen der WHO zu alkoholbezogenen Störungen (Barbor et al., 2001) und basierend auf Übersichtsarbeiten (Reinert & Allen, 2007) resultiert ein deutliches Votum zugunsten des *AUDIT – Alcohol Use Disorders Identification Test*³. Ungeachtet etablierter Verfahren hat sich dieses zunächst für Screening konzipierte 10-Item-Instrument (Bearbeitungsdauer: 1–2 Minuten) heute durchgesetzt. Die zwei faktorenanalytisch ermittelten Dimensionen lauten „consumption“ und „adverse consequences“, in der Praxis wird aber ein einfacher Gesamtsummenwert gebildet. Reliabilitäts- und Validitätsbefunde liegen in großer Zahl vor. Die Sensitivität/Spezifität beträgt 90 % bzw. 80 %. Auf die Eignung als Evaluationsmaß in Präventionsprogrammen wird hingewiesen (Babor et al., 2001, S. 27). Für die ICD-Diagnosen F10.1 und F10.2 gilt als relevanter Outcome aber nicht allein der *AUDIT*-Wert, sondern eine Merkmalsmatrix aus Abstinenz, etwaiger Rückfallrate pro Zeit (z. B. nach einem Jahr) und gesellschaftlicher und beruflicher Wiederteilhabe (DGPPN/DG-SUCHT, 2020, S. 356). Der *AUDIT* bietet die Möglichkeit eines raschen Überblicks, unter anderem auch in all jenen Fällen, in welchen eine andere Primärdiagnose als F10.1 vorliegt. Eine Anzahl weiterer Kennwerte berichtet Dybek (2008, S. 93). International gilt ein Cut-off-Wert von 8 Punkten (Summe über die 10 Items; min. = 0, max. = 40) als Kriterium für bedenklichen Konsum. In Deutschland wird eine Reduktion auf 5 Punkte z. Zt. diskutiert.

Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen (ADHS) im Erwachsenenalter

Zur Diagnostik der Aufmerksamkeitsdefizit-/ Hyperaktivitätsstörung (ADHS; F90) im Erwachsenenalter wird ein mehr-

³ Verfügbar unter: https://www.bundesaeztekammer.de/fileadmin/user_upload/downloads/AlkAuditFragebogen.pdf (17.10.2021).

schrittiges Vorgehen empfohlen (AWMF, 2017). Hierfür bieten die *Homburger ADHS-Skalen für Erwachsene (HASE)*; Rösler et al., 2021) das am weitesten verbreitete Instrumentarium. Zur Verbesserung der Differenzialdiagnostik der ADHS im Erwachsenenalter kann die Kombination von *WURS-k* und *ADHS-SB* als Bestandteile des *HASE* empfohlen werden (Paucke et al., 2018).

Die deutsche Kurzform der *Wender Utah Rating Scale (WURS-k)* erfasst die retrospektive Selbstbeurteilung von Symptomen der ADHS im Kindesalter (8.–10. Lebensjahr, 21 Items plus 4 Items zur Bewertung einer Antworttendenz). Bei gesicherter Reliabilität kann ein Gesamtwert bestimmt und anhand eines gut validierten Cut-offs (≥ 30 ; Wertespanne: 0–84) interpretiert werden. Der retrospektiven Zielsetzung des Instruments entsprechend ist keine Veränderungsmessung anhand der *WURS-k* sinnvoll.

Die *ADHS-Selbstbeurteilung (ADHS-SB)* erlaubt sodann die Bestimmung der Schwere der aktuellen ADHS-Symptomatik. Zur Auswertung können ein Gesamtwert (0–18 Kriterien) sowie Summenwerte für die Teilbereiche Unaufmerksamkeit (9 Kriterien), Überaktivität (5 Kriterien) und Impulsivität (4 Kriterien) bestimmt werden, deren Reliabilität Mindestanforderungen erfüllt. Als Belege der Validität werden beispielsweise signifikante Zusammenhänge mit anderen Selbst- und Fremdeinschätzungsmaßen berichtet. In der aktualisierten Version liegen Prozentrang- und Stanine-Normen für den Gesamtwert und Normwerte für die Subskalen vor. Außerdem werden Empfehlungen zu Cut-off-Werten gegeben. Die Instruktion („innerhalb der letzten 7 Tage“) legt eine Eignung der *ADHS-SB* zur Veränderungsmessung nahe (Rösler et al., 2006).

Bipolare Störungen

Die bipolaren Erkrankungen F30.0, F30.1 und F31.x mit einer Lebenszeitprävalenz von etwas mehr als 1 % kommen – etwa in Relation zu Depression und Ängsten – in der Psychotherapiepraxis vergleichsweise selten vor (Überblick Meyer, 2021a). Das Selbstbeurteilungsinstrument der Wahl für Verlauf und Outcome ist die *Altman Self-Rating Mania Scale (ASRM)*; Altman et al., 1979; Meyer, 2021b; Meyer et al., 2020). Wenngleich wegen der mitunter geringeren Einsichtsfähigkeit dieser Patient*innen ein reines Fragebogen-prozedere gelegentlich angezweifelt wurde, konnte sich die *ASRM* dennoch überraschend gut bewähren und erwies ihre Eignung als Outcomemaß in der Praxis. Eine Beschreibung findet sich bei Bernhard und Meyer (2012; dort auch vollständiger Skalenabdruck). Es handelt sich dabei um fünf Fragengruppen zu jeweils fünf in der Intensität zunehmenden Beschreibungen (max. fünf Minuten). Cut-off ist ein Wert von 6 (min. = 0, max. = 20), dies deutet auf einen hypomanen

respektive manischen Zustand hin. Sinnvoll ist ein wöchentliches oder monatliches Bearbeitungsintervall (Bernhard & Meyer, 2012). Altman et al. (1979) berichten eine gute Skalenhomogenität und eine Diskriminationsfähigkeit „Bipolar vs. Nichtpatient*innen“ von 87 %.

Borderline-Persönlichkeitsstörungen

Ziel der *Borderline Symptom List (BSL-23)*; Bohus et al., 2009) ist die Bestimmung des Schweregrades der spezifischen Symptomatik von Borderline Persönlichkeitsstörungen (F60.3). Die *BSL-23* wurde als ökonomische Kurzform (23 Items) der *BSL-95* entwickelt, welche die Selbsteinschätzung von sieben Merkmalen/Subskalen der Borderline-Störung erlaubt: Selbstwahrnehmung, Affektregulation, Autoaggressivität, Dysthymie, soziale Isolation, Intrusionen, Feindseligkeit (Bohus et al., 2001). Ergänzend werden eine allgemeine Befindlichkeitseinschätzung (von 0 bis 100 %) sowie Fragen zu selbstschädigendem Verhalten vorgelegt. Das Verfahren inklusive Auswertungsbogen, Auswertungshinweisen und Prozentrang-Normen ist in Deutsch und 18 (!) weiteren Sprachen online über die Website des Zentralinstituts für Seelische Gesundheit Mannheim verfügbar. Zur Bewertung der

— **Auf Basis der Empfehlungen der WHO und von Übersichtsarbeiten hat sich der *AUDIT – Alcohol Use Disorders Identification Test* bei alkoholbezogenen Störungen als maßgebliches Messinstrument durchgesetzt.** —

Gütekriterien liegen unter anderem Angaben zur Reliabilität, zur Differenzierung von Patient*innen mit anderen Störungen und zur Veränderungssensitivität vor (Bohus et al., 2009). Anwendung der *BSL-23* wie auch deren Auswertung sind sehr einfach. Normwerte liegen auf Basis von 560 Borderline-Patient*innen vor. Bei der Interpretation ist zu berücksichtigen, dass die Schwere der Borderline-Symptomatik in Bezug auf eine klinische Vergleichsgruppe bestimmt wird. Das Messverfahren hat sich in einer Reihe von Evaluationsstudien zur dialektisch-behavioralen Therapie (DBT) als veränderungssensitiv erwiesen.

Computerspielabhängigkeit

Die Computerspielabhängigkeit (Internet Gaming Disorder) umfasst abhängiges Verhalten von Online-Computerspielen ohne Geld (d. h. kein pathologisches Glückspiel) als neue Diagnose des DSM-5 (in der ICD-11 künftig: 6C51; vgl. Müller, 2021). Die *Computerspielabhängigkeitsskala (CSAS)*; Rehbein et al., 2015) erlaubt die Erfassung der diagnostischen Kriterien als Selbst- wie auch Fremdbeurteilungsinstrument. Mit dem Verfahren liegen spezifische Versionen für die Altersbereiche Erwachsene (*CSAS-E*); bzw. Fremdeinschätzung durch Partner*innen: *CSAS-FP*) und Jugendliche (*CSAS-J*); bzw. Fremdeinschätzung durch Eltern: *CSAS-FE*) vor. Anhand von

jeweils 18 Items werden hierin u. a. die gedankliche Vereinnahmung durch das Computerspielen, Entzugserscheinungen, Toleranzentwicklung, Kontrollverlust und negative psychosoziale Folgen erfragt. Zur Auswertung wird ein Gesamtwert bzgl. Erfüllung der diagnostischen Kriterien bestimmt. Zur Interpretation werden alters- und geschlechtsspezifische Normwerte angeboten (Stanine und Prozentrangwerte). Ferner liegen gute Reliabilitätskennwerte vor sowie Validitätsbelege z. B. zur Übereinstimmung mit der Dauer der Computerspielnutzung. Zwar werden kritische Differenzen berichtet, aufgrund des langen Beurteilungszeitraums („innerhalb der letzten 12 Monate“) ist bei Verwendung des Instrumentes zur Veränderungsmessung gegebenenfalls die Instruktion anzupassen.

Depression

Für Depression (ICD-Gruppen F32.x/F33.x) – sicherlich derzeit häufigster Anlass zur Aufnahme einer Psychotherapie – geht kein Weg am *Beck-Depressions-Inventar II* vorbei (*BDI-II*; Beck et al., 1996; Hautzinger et al., 2009). Dieses etablierte Untersuchungsinstrument erfüllt im Prinzip alle Kriterien, es misst inhaltlich valide die relevanten Depressionsmerkmale (von Version I zu II nochmals verbessert), ist für Evaluation geeignet, verfügt über eine umfangreiche Datenbasis und ist international anerkannt in klinisch-psychologischer und psychiatrischer Behandlungspraxis sowie in der Forschung. Differenzierte Richtwerte, Prozentrangnormen und kritische Unterschiedswerte sind angegeben (klinisch bedeutsam ist eine Messzeitpunkte-Differenz von 8 Punkten; Gesamtränge 0 bis 63). Informationen zu den verschiedenen Stichproben, zur Teststatistik, zur Validität (auch bzgl. Fremdratingverfahren), zur Retest-Zuverlässigkeit und zur Eignung für Veränderungsmessung finden sich im Manual (Hautzinger et al., 2009). Die Entdeckungsgenauigkeit (Sensitivität) und die daraus abgeleitete Stufeneinteilung von „keine“ bis „schwere“ Depression ist aus dem Amerikanischen übernommen. Dimensionsanalysen ergaben korrelierte Faktoren, so dass der Summenwert über den Gesamtfragebogen verwendet wird.

Dissoziative Störungen

Der *Fragebogen zu Dissoziativen Symptomen – Kurzform (FDS-20)*; Spitzer et al., 2021) ist eine deutschsprachige Adaptation der *Dissociative Experiences Scale (DES)*; Bernstein & Putnam, 1986). Das Verfahren erlaubt die Bestimmung der Schwere dissoziativer Symptome, wie sie im Rahmen dissoziativer Störungen (ICD-Gruppe F44.x) wie auch als Begleitsymptomatik bei vielen anderen psychischen Störungen, auftreten können (Lyssenko et al., 2018). Die 20-Item-Kurzform wurde zum Screening und als veränderungssensitives Instrument zur Verlaufsmessung entwickelt. Hierfür wurde ein Bewertungszeitraum „in den letzten zwei Wochen“ vorgeschlagen. Die Durchführung und die Auswertung mit Berechnung eines Gesamtwertmittels sind einfach möglich. Für das Verfahren wird eine hohe Reliabilität sowie Angaben zur konvergenten Validität, divergenten Validität und Verän-

derungssensitivität berichtet. Zur Interpretation des *FDS-20* und Bewertung von Veränderungseffekten werden kritische Werte des Erreichens nicht-pathologischer Werte sowie Kriterien einer statistisch signifikanten Veränderung angeboten, z. B. Cut-off 7 für Erreichen eines nicht-pathologischen Wertes bzw. Veränderungen um mindestens 11,6 im Gesamtwert (Wertespanne: 0–100; Spitzer et al., 2021).

Essstörungen

Im Bereich Essstörungen (ICD-Gruppe F50.x) liegt als ausgereiftes und international in Praxis und Forschung bewährtes Instrument das *Eating Disorder Inventory – Version II* vor (*EDI-II*; Garner, 1991; Paul & Thiel, 2005). Das Messverfahren gibt es in einer 64-Item-Standardform sowie erweitert um 30 Items. Die Autoren betonen die gute Eignung in der Veränderungsmessung bei der Magersucht- und Bulimiebehandlung. Alle Störungs-Charakteristika sind abgebildet, sie stehen im Einklang mit den Konzepten der einflussreichen deutsch-amerikanischen Psychoanalytikerin Hilde Bruch: Schlankheitsstreben, Bulimie, Unzufriedenheit mit dem Körper, Ineffektivität, Perfektionismus, Misstrauen, Interozeptive (Fehl-)Wahrnehmung sowie Angst vor dem Erwachsenwerden. Auswertung und Interpretation sind skalenbezogen. Der *Version II* des *EDI* liegen Daten von 500 Magersucht- und Bulimiepatient*innen zugrunde, ferner Daten aus *Version I* auch zur Binge Eating Disorder und atypischen Essstörungen. Studien zur Zuverlässigkeit des Fragebogens und seiner inhaltlichen Eignung untermauern die Einsatzmöglichkeiten im klinischen Feld. Prozentränge liegen für restriktive Magersucht und Magersucht/purging Type, für Bulimie und für männliche und weibliche Kontrollpersonen vor (dadurch Cut-off-Bestimmung möglich).

Hypochondrie

Für die Messung der beharrlichen Beschäftigung mit einer möglichen Krankheit oder der Überzeugung, eine solche zu haben (F45.2), existieren die Hypochondriefragebögen *Whiteley-Index* und *Illness Attitude Scales (WI, IAS)*; im gemeinsamen Manual: Hiller & Rief, 2004). Beide Messinstrumente sind gründlich untersucht und in ihrer Nützlichkeit international wie auch im deutschsprachigen Raum durch Studien belegt. Aus der vergleichenden Schlussbetrachtung (Hiller & Rief, 2004, S. 61–62) für die beiden relativ ähnlichen Instrumente (*WI*, Pilowsky, 1997; *IAS*, Kellner, 1986) resultiert eine leichte Überlegenheit des *Whiteley-Index*: Praxisrelevant sind z. B. Kürze (*WI* 14 Items; *IAS* 29 Items), einfacheres Vorgebformat und schnellere Auswertung. Beide Verfahren sind nicht im strengen Sinne eindimensional, allerdings sind die Faktoren Krankheitsängste, somatische Beschwerden und Krankheitsüberzeugungen im *WI* inhaltlich stark aufeinander bezogen und lassen die Bildung eines Gesamtsummenwerts als Indikator für Hypochondrie zu. Für die deutsche Fassung wurden 800 Patient*innen rekrutiert, überdies auch eine bevölkerungsrepräsentative Stichprobe (N = 2.000). Reichhaltige Reliabilitäts- und Validitätsbefunde liegen vor (die Veränderungssensitivität ist an 260 Psychosomatikpatient*innen

belegt), und es werden Angaben zu Sensibilität und Spezifität, zum Cut-off und zu Klassifikationspunktwerten berichtet (Prozentrangnormen).

Posttraumatische Belastungsstörung (PTBS)

Die *Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5; deutsch: Krüger-Gottschalk et al., 2017 – Appendix)* erlaubt die Bestimmung der Schwere einer posttraumatischen Belastungsstörung auf Basis der aktuellen Diagnosekriterien nach DSM-5. Anhand von 20 Selbsteinschätzungen kann die Schwere von Intrusionen (Items 1–5), Vermeidungen (Items 6–7), negativen Kognitionen und Emotionen (Items 8–14) und Hyperarousal (Items 15–20) sowie ein Gesamtwert (Wertespanne: 0–80) bestimmt werden. Für die deutsche Version wird eine hohe Homogenität beschrieben. Der Cut-off liegt bei 33. Als Validitätsbelege liegen u. a. Übereinstimmungen mit Ergebnissen strukturierter Interviews vor. Ferner sind bislang vor allem für die englische Originalversion (Weathers et al., 2013) weitere Belege für Validität und Veränderungssensitivität verfügbar (Blevins et al., 2015; Wortmann et al., 2016).

Ergänzend ist anzumerken, dass sich zur Erfassung der in der ICD-11 neu etablierten Diagnose einer Komplexen Posttraumatischen Belastungsstörung (KPTBS)⁴ derzeit das *International Trauma Questionnaire (ITQ; Cloitre et al., 2018)* als Standard entwickelt. Dieses 18-Item-Instrument liegt in 28 Sprachen vor (deutsch: Lueger-Schuster et al., 2018). Eine KPTBS kann anhand von je zwei Kriterien der drei Symptombereiche Wiedererleben im Hier und Jetzt, Vermeidung und Gefühl einer aktuellen Bedrohung (bezogen auf den letzten Monat) bestimmt werden. Ferner müssen Kriterien einer Störung der Selbstorganisation (z. B. negatives Selbstkonzept) und der funktionellen Beeinträchtigung (z. B. in sozialen Kontakten) erfüllt sein.

Psychosen

Für die Outcome-Messung steht bei ICD-Gruppe F2 (Schizophrenie – Wahnhafte Störungen) ein *Fremdratingverfahren* zur Verfügung, die *Positive and Negative Syndrome Scale – PANSS* (Moritz et al., 2017), welche als Untersuchungsinstrument dort breite Anerkennung findet. Sie wurde zur typologischen Beschreibung wie auch für die dimensionale Beurteilung schizophrener Erscheinungsbilder entwickelt, „ihre Änderungssensitivität macht sie zum Goldstandard in Interventionsstudien“ (Moritz et al., 2017, S. 150). Die Skala besteht aus 30 Items, die auf die Bereiche Positivsymptomatik, Negativsymptomatik und Globale Symptomatik aufgeteilt sind (Items bei Moritz et al., 2017, S. 149). Von den PANSS-Originalautoren Kay et al. (2000) wird geraten, zur Stützung mit den Patient*innen auch das strukturierte klinische Interview

SCI-PANSS durchzuführen: Dabei sollen die Verhaltensbeobachtungen während des Interviews und die entsprechenden Antworten in der *PANSS* eingetragen werden (Opler et al., 1999). Erfahrene Kliniker*innen kommen jedoch auch ohne dieses Tool aus (Moritz, 2021). Angaben zu Reliabilität und Validität finden sich bei Kay et al. (2000) und Van der Gaag et al. (2006). Für den von min. = 30 bis max. = 210 variierenden Punktwertbereich liegt eine Schweregradeinteilung vor („leicht – mittelgradig – deutlich – schwer“). Studien der Hamburger Arbeitsgruppe an fast 6.000 Patient*innen dienen als fundierte Datenbasis (auch zur Beurteilung hinsichtlich Besserung bzw. Remission; Moritz et al., 2017).

Schlafstörungen

Der *Insomnia Severity Index (ISI; Dieck et al., 2018; Morin, 1993)* ist eine kurze und leicht auswertbare Skala zur Bestimmung der Schwere von Schlafstörungen, wie sie sowohl als eigene Symptomatik (nicht-organische Insomnie, ICD-Gruppe F51.x) als auch als Begleitsymptomatik anderer psychischer Störungen sehr verbreitet sind. Anhand von sieben Items wird bezogen auf die vergangenen zwei Wochen das Ausmaß von Einschlafstörungen, Durchschlafstörungen, Früherwachen, Zufriedenheit mit dem Schlaf sowie subjektiver Folgen der Schlafstörungen beurteilt. Das Messverfahren wurde unter anderem anhand von 700 Patient*innen mit primärer Insomnie sowie einer nichtklinischen Stichprobe validiert (Cut-off Gesamtwert > 10; Wertespanne: 0–28). Darüber hi-

— Für das Störungsbild der Depression geht kein Weg am Beck-Depressions-Inventar II vorbei. —

naus werden aus der englischen Version Cut-off-Werte für vier Schweregrade angeboten (vgl. Dieck et al., 2018, S. 29). Insgesamt werden eine gute Reliabilität, Übereinstimmungen mit anderen Schlafstörungs-Instrumenten und Hinweise auf die diskriminative Fähigkeit berichtet (Dieck et al., 2018; Gerber et al., 2017), sodass das Untersuchungsinstrument als Standard zur Veränderungsmessung empfohlen werden kann (Riemann et al., 2017).

Schmerz

Schmerzmessung erfolgt in bewährter Weise mittels des Systems *Schmerzempfindungsskala* und *Fragebogen zur Erfassung der Schmerzverarbeitung (SES, FESV; Geissner, 1996, 2001; Martin & Henningsen, 2021)*. Mit Hilfe des Verfahrens *SES* wird zunächst die Stärke des *affektiven* und des *sensorischen* Schmerzerlebens festgestellt. In einem weiteren Schritt wird untersucht, über welches Ausmaß an schmerzbezogenen Bewältigungsstrategien Patient*innen

⁴ Durch die Kombination dieser beiden Verfahren wird dem Einbezug der komplexen PTBS in der ICD-11 und ihrer hohen Bedeutung in der klinischen Praxis entsprochen.

Die ICD-Diagnosen F45.4 und F45.41 bieten die Grundlage für die psychotherapeutische Behandlung von Patient*innen mit chronischen Schmerzen, worunter in Deutschland 12 Millionen Menschen – dauerhaft fortbestehend oder in Abständen wiederkehrend – leiden (Nobis et al., 2020). Eine Erleichterung für die psychotherapeutische Behandlung ist die erst unlängst erfolgte Etablierung der ICD-Ziffer F45.41, wodurch eine – ohnehin fragliche – Psychogenitätsannahme die Schmerzproblematik betreffend nicht mehr Voraussetzung sein muss. Hier wird im Übrigen eine Brücke zwischen den F-Diagnosen und mehreren somatischen ICD-Kapiteln geschlagen, was Nilges et al. (2021) als „klare schmerzpsychologische Erfolgsgeschichte“ werten.

Anmerkung zu F-Diagnosen bei Schmerz

verfügen. Dies gliedert sich in handlungs- und in kognitiv orientierte Komponenten. Drei Faktoren der schmerzbedingten psychischen Beeinträchtigung komplettieren die Untersuchung der psychischen Seite des Schmerzgeschehens. Nach methodisch aufwändigen Dimensions- und Skalenanalysen resultieren für die Praxis sehr kurze Skalen (10 und 14 Aussagen für SES; je vier bzw. fünf Aussagen für die einzelnen FESV-Bereiche) mit schneller Bearbeitbarkeit. Zahlreiche Validierungsbefunde liegen manualisiert zusammengefasst vor. Die Veränderungssensitivität (Zeitraumen „in den letzten Tagen“; für die SES sind zusätzlich auch die Varianten „jetzt“ und „in den letzten 3 Monaten“ wählbar) konnte in Therapiestudien belegt werden. Die T-Wert- und Prozentrang-Normen basieren auf Eichstichproben von insgesamt fast 1.500 Patient*innen für eine Bandbreite an Schmerzlokalisationsgebieten und -diagnosen. Klinisch signifikante Mindestdifferenzwerte zwischen zwei Messzeitpunkten sind vorgegeben und werden darüber hinaus an Beispielen erläutert.

Sexuelle Funktionsstörungen

Sexualstörungen finden sich im ICD an verschiedenen Stellen – die Diagnosegruppe, die hier zur Diskussion steht, ist die ICD-Gruppe F52.x (Sexuelle Funktionsstörungen, nicht verursacht durch eine organische Störung oder Krankheit). Das Vorgehen besteht aus einer sorgfältigen Exploration im angstfreien Raum, für welches Interview-Leitfäden vorliegen (Hoyndorf et al., 1995; Strauß et al., 2013; vgl. auch Reiniger & Briken, 2021, S. 308–309). Der Fragebogen der Wahl ist das *Sexual Behaviour Questionnaire – deutsch (SBO-G)*, Müller, 2016; im elektronischen Testarchiv des Leibniz-Instituts für Psychologie verfügbar). Eine Kurzbeschreibung findet sich ebenfalls dort unter PubPsych. Die Skaleneigenschaften beschrieb der Autor andernorts (Müller, 2007). Für Männer gibt es elf Fragengruppen, für Frauen zehn, mit jeweils abgestuften Unteraussagen. Der Beurteilungszeitraum ist „die letzten 3 bis 6 Monate“; insofern besteht ab einer gewissen Psychotherapiedauer die prinzipielle Eignung für Outcome-Messung. Die Skala ist faktoriell eindimensional und weist zufriedenstellende Reliabilitäten auf. Richtwerte für kleine Stichproben (USA und Deutschland) sind angegeben.

Somatoforme Störungen

Der Fragebogen *SOMS – Screening für Somatoforme Störungen* (F45.0, 45.1, 45.3; Martin & Henningsen, 2021; Rief & Hiller, 2019) beinhaltet ungeachtet der Bezeichnung „Screening“ eine für Veränderungsmessung gut geeignete Variante *SOMS-7T*, in welcher gefragt wird, wie sehr die Person „in den vergangenen 7 Tagen“ unter den Beschwerden litt. Die Bearbeitung nimmt wenige Minuten in Anspruch, es werden 52 Items zu den verschiedensten körperlichen Beschwerden und Einschränkungen vorgelegt, bei welchen das Ausmaß des subjektiven Leidens anzugeben ist. Mehrere Maße können zur Auswertung verwendet werden, am wichtigsten für unsere Zwecke ist der *Somatisierungsindex ICD-10*. Für die Testgütebestimmung standen Daten von 500 Patient*innen zur Verfügung. Der Fragebogen trennt gut zwischen Gesunden und Patient*innen, erlaubt die Klassifikation nach Unterkriterien und liefert aussagekräftige Belege zur Änderungssensitivität („7-Tage-Version“; N = 325). Prozentrangnormen sowohl für Patient*innen als auch aus einer repräsentativen Bevölkerungsstichprobe (N = 1.300 Frauen, 1.200 Männer) sind aufgeführt.

Soziale Phobie

Die *Liebowitz-Soziale-Angst-Skala (LSAS)* (Liebowitz, 1987; Stangier & Heidenreich, 2015) wurde zunächst als Fremdbeurteilungsinstrument zur Erfassung von Angst und Vermeidung in sozialen Situationen entwickelt, jedoch wird heute auch die Verwendung als Selbstbeurteilungsinstrument empfohlen (vgl. Bandelow et al., 2021, S. 59). Die LSAS ist als Bestandteil des Skalen-Sets *Soziale Angststörung (SOZAS)* (Consbruch et al., 2016) verfügbar. In der Durchführung werden 24 Situationen (13 Leistungssituationen und 11 soziale Interaktionssituationen) sowohl in Bezug auf das Ausmaß von Angst/Furcht als auch das Ausmaß des Vermeidungsverhaltens eingeschätzt. In der Auswertung können getrennt für diese beiden Bereiche Summenwerte bestimmt werden, aber auch ein Gesamtwert. Die Interpretation ist anhand von Cutoff-Werten möglich (> 30; Wertespanne: 0–144), ermittelt in der englischen Version. Als Gütekriterien werden eine hohe Reliabilität und eine hohe Übereinstimmung mit anderen Skalen des SOZAS berichtet. Das Verfahren hat sich zur Veränderungsmessung sozialer Ängste bewährt, zum Erreichen einer signifikanten Veränderung ist eine kritische Differenz von 29 erforderlich (Gesamtwert; Consbruch et al., 2016, S. 41). Das Verfahren ist als Selbsteinschätzungsinstrument schnell und einfach durchzuführen, leicht auszuwerten und liefert im übrigen nützliche Informationen z. B. im Rahmen der Expositionsbehandlung.

Trauer

In der ICD-11 wird Langanhaltende Trauer als eigenständiges – also von Depression, Anpassungsstörung etc. verschiedenes – Krankheitsbild aufgenommen (Wagner, 2016; ICD-11-Codenummer 6B42). Das meistverbreitete Instrument ist

Störung und ICD-Code	Messinstrument (Abkürzung)	Bearbeitungsdauer (Min.)	V: Eignung für Veränderungsmessung? Z: Zeitrahmen der Beurteilung */**	Normen Richtwerte	Beschreibung (B) und konkreter Publikationsort (P) des Verfahrens ***
Agoraphobie und Panikstörung F40.0, F40.2, F41.0	Fragebogen Ängste, Kognitionen und Vermeidung (AKV)	15-20'	V: Ja Z: Situativ „wenn Angst da“ **	Stanine-Normen Prozentrang-Normen	B: Manual P: Beltz Test GmbH
Alkohol (Abhängigkeit bzw. schädlicher Gebrauch) F10.1, F10.2	Alcohol Use Disorder Identification Test (AUDIT)	2'	V: Prinzipiell möglich Z: Prinzipiell „gegenwärtige Lage“ **	Richtwerte, Cut-off, Konsensusangaben	B: WHO-Manual online; Referenzartikel P: online
Aufmerksamkeitsstörung im Erwachsenenalter F90	ADHS-Selbstbeurteilung (ADHS-SB) und Wender Utah Rating Scale (WURS-k)	10' + 10'	V: ADHS-SB ja, WURS-k nein Z: ADHS-SB „letzte 7 Tage“ **	Cut-off, Prozentrang-Normen, Stanine-Normen	B: Manual P: Hogrefe
Bipolare Störung F30.0, F30.1, F31.x	Altman Self-Rating Mania Scale (ASMR)	3-4'	V: Ja Z: Prinzipiell „gegenwärtige Lage“ **	Richtwerte, Cut-off	B: diverse Referenzartikel P: Buchkapitel
Borderline-Persönlichkeitsstörung F60.3	Borderline Symptom List (BSL 23)	10'	V: Ja Z: „In der letzten Woche“ **	Prozentrang-Normen	B: online P: online
Computerspiel- und Online-Abhängigkeit ICD-11 – 6C51	Computerspiel- und Onlineabhängigkeit (CSAS)	10'	V: Bedingt, ggf. zu adaptieren Z: „letzte 12 Monate?“ **	Stanine und Prozentrang-Normen	B: Manual P: Hogrefe
Depression F32.x, F33.x, F34.1	Beck Depressions-Inventar II (BDI II)	7-8'	V: Ja Z: „In den letzten 2 Wochen“ **	Prozentrang-Normen, Richtwerte, Schweregrade	B: Manual P: Pearson Assessment Deutschland
Dissoziative Störungen F44.x	Fragebogen zu Dissoziativen Symptomen – Kurzform (FDS-20)	10'	V: Ja Z: „In den letzten 2 Wochen“ **	Kritische Werte, Reliabile Change Index, Cut-off, Quartile	B: Manual P: Hogrefe
Essstörungen F50.x	Eating Disorder Inventory-2 (EDI-2)	15-20'	V: Prinzipiell Ja Z: Prinzipiell „gegenwärtige Lage“ **	Prozentrang-Normen für Magersucht, Bulimie, gesunde Personen	B: Manual P: Hogrefe
Hypochondrie F45.2	Whiteley-Index und Illness Attitude Scale (WI / IAS)	3' bzw. 5'	V: Ja Z: Prinzipiell „gegenwärtige Lage“ **	Prozentrang-Normen Patient*innen, gesunde Personen, 2 Altersgruppen	B: Manual P: Huber
Posttraumatische Belastungsstörung F43.1	Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5) und International Trauma Questionnaire (ITQ)	10' + 12'	V: Ja Z: „Im letzten Monat“ **	PCL-5: Cut-off; ITQ: Anzahl erfüllter Kriterien	B: online und Referenzartikel P: online und Referenzartikel
Psychosen F20, F22	Positive and negative Syndrome Scale (PANSS); beachte: Fremdratingverfahren	30' incl. Interview	V: Ja Z: Prinzipiell „gegenwärtige Lage“ **	4 Schweregrade, plus Werte für „Besserungsgraden“	B: diverse Referenzartikel u. -kapitel P: Buchkapitel
Schlafstörungen F51.x	Insomnia Severity Index (ISI)	3'	V: Ja Z: „In den letzten beiden Wochen“ **	Cut-off	B: Referenzartikel P: Referenzartikel
Schmerz F45.4, F45.41	Schmerzempfindungsskala und Fragebogen zur Erfassung der Schmerzverarbeitung (SES / FESV)	3' + 7'	V: Ja Z: „In den letzten Tagen“ **	T-Wertnormen, Prozentrang-Normen; allgem. und gruppenweise	B: Manual P: Hogrefe
Sexuelle Funktionsstörungen F52.x	Sexual Behaviour Questionnaire (SBO-G)	5'	V: Prinzipiell möglich Z: Prinzipiell „gegenwärtige Lage“ **	Einfache Richtwerte US und D: M und SD für kleine Gruppen	B: Referenzartikel P: Leibniz Inst. Psych. Online Testarchiv
Somatoforme Störungen F45.0, F45.1, F45.3	Screening für Somatoforme Störungen, Variante „7T“ (SOMS)	5-10'	V: Ja Z: „In den letzten 7 Tagen“ **	Prozentrangnormen, Patient*innen und gesunde Personen	B: Manual P: Hogrefe
Soziale Phobie F40.1	Liebowitz Soziale Angst Skala (LSAS)	15'	V: Ja Z: „In der letzten Woche“ **	Cut-off, Referenzwerte	B: Manual P: Hogrefe
Trauer (Anhaltende Trauerstörung) ICD-11 – 6B42	Inventory of complicated Grief (ICG-D)	3'	V: Ja Z: Prinzipiell „gegenwärtige Lage“ **	Richtwerte, Cut-off	B: Referenzartikel und -kapitel P: Buchkapitel Therapieanleitung
Zwang F42.x	Yale-Brown Obsessive Compulsive Inventory Selbst-rating-Version – (YBOCS-SR)	10-15'	V: Ja Z: „Innerhalb der letzten 7 Tage“ **	Cut-off, Richtwerte Schweregrad	B: Referenzartikel P: Buchkapitel, Behandlungsmanual

Tabelle: Von A wie Angst bis Z wie Zwang – Übersicht zu den störungsspezifischen und veränderungssensitiven Messverfahren für die Psychotherapieevaluation
 Anmerkungen: * Erhebungsintervall: In größeren Zeitabständen (etwa 3- oder 6-monatlich) sinnvoll. ** Erhebungsintervall: Wöchentlich/monatlich laut Autoren empfohlen bzw. möglich.
 *** Genaue Quellenangaben nach Verfahren geordnet im elektronischen Supplement zur gedruckten Version des Beitrags.

das *Inventory of Complicated Grief – ICG* (Prigerson et al., 1995), welches Lumbeck et al. (2012) in einer deutschen Version vorgelegt haben. Der 19-Item-Fragebogen erwies sich als faktorenanalytisch eindimensional, so dass ein Summenwert (min. = 0 bis max. = 76) zu bilden ist. Reliabilitätsbefunde und Angaben zur Validität sprechen für die gute Einsatzmöglichkeit des Verfahrens. Es kann klar zwischen Betroffenenengruppen unterscheiden (Cut-off = 25) und erwies sich in Therapiestudien als veränderungssensitiv (Rosner et al., 2011). Das Verfahren ist Bestandteil des Behandlungsmanuels „Anhaltende Trauerstörung“ (Rosner et al., 2015; Kapitel 3: Fragebogenbeschreibung; Anhang: Abdruck der Skala).

Zwangsstörungen

Die *Yale-Brown Obsessive-Compulsive Scale – Selbstrating* (*YBOCS-SR*; Hand & Büttner-Westphal, 1991; verfügbar in Büttner-Westphal & Hand, 1991; Lakatos & Reinecker, 1999; Ertle, 2012) bildet das international am weitesten verbreitete Selbsteinschätzungsinstrument zur ökonomischen Bestimmung der Schwere einer Zwangserkrankung (ICD-Gruppe F42.x). Anhand von jeweils fünf Items können die zwei Teilbereiche Zwangsgedanken und Zwangshandlungen erhoben sowie ein Gesamtwert bestimmt werden. Reliabilität und Validität des Verfahrens werden als gut bewertet (Jacobsen et al., 2003). Ein Gesamtwert von 16 (Wertespanne: 0–40) gilt als etablierter kritischer Wert. Die *YBOCS-SR* hat sich sowohl zur Evaluation individueller Therapieverläufe wie auch in Forschungsstudien als sehr gut veränderungssensitiv bewährt.

— Die Yale-Brown Obsessive-Compulsive Scale – Selbstrating bildet das international am weitesten verbreitete Selbsteinschätzungsinstrument zur Bestimmung der Schwere einer Zwangserkrankung. —

Kriterien für die Bewertung der Therapieresponse liegen darüber hinaus vor (Kordon et al., 2013, S. 105).

Bei der Anwendung der *YBOCS-SR* ist zu berücksichtigen, dass im Vergleich zum halbstandardisierten *YBOCS-Interview* das Ausmaß der Zwänge vor allem im Bereich der Zwangsgedanken tendenziell unterschätzt wird (Federeci et al., 2010). Daher ist bei Bedarf die ergänzende Anwendung des standardisierten Interviews zu empfehlen, ebenso wie die darüber hinaus verfügbare Checkliste (*YBOCS-CL*; in Büttner-Westphal & Hand, 1991), welche gezielt ein breites Spektrum möglicher Zwangsgedanken und Zwangshandlungen erfragt.

Diskussion

Moritz et al. (2017) apostrophierten das Untersuchungsinstrument *PANSS* (Kay et al., 2020) als „Goldstandard“ in der Veränderungsmessung der Psychosen. Wir greifen dies einmal auf und hoffen, dass auch die anderen vorgestellten

Instrumente ein Goldstandard-Potential in der Outcome-Bestimmung von Psychotherapie entwickeln können. Ausschlaggebend dafür ist sicher die Bewährung in der Praxis und die Rückkopplung der gewonnenen Erfahrungswerte an die Autor*innen der einzelnen Messverfahren. Dies schließt stellenweise auch eine weitere Verbesserung der Datenbasis mit ein. Zudem wäre eine Vereinheitlichung der Qualitätskriterien hinsichtlich Mindeststandards – auch in der Weiterentwicklung der Instrumentarien – von Vorteil, wozu die Normierung und klare Effekt-Beurteilungsmöglichkeit gehören.

Untersuchungszeitpunkte: Ausgehend von der interessanten Beobachtung, dass Messwerte unmittelbar am Therapiebeginn im Vergleich zu Messwerten bei der davor stattgefundenen ersten Anmeldung/Kontaktaufnahme gar nicht mehr so negativ ausfielen, analysierten Geissner et al. (2021a, b) Gründe hierfür, etwa Hoffnung auf Erfolg, erwartete Sicherheit im geschützten Therapiekontext etc. und postulierten ein System multipler Untersuchungszeitpunkte, um den Erfolg der eigenen psychotherapeutischen Maßnahmen sowohl methodisch sauber als auch fair, dabei praxisgeeignet zu bestimmen. Ein Plädoyer für die Messung *bereits zum Anmeldezeitpunkt* wird dort ausgesprochen, und konkrete Umsetzungshinweise und Beispiele finden sich an Ort und Stelle.

Ein Seitenblick: Wie eingangs dargestellt, ist es nicht zwingend notwendig, weitere Fragebögen, die „auch noch interessant“ wären, einzusetzen, um die Compliance der Patient*innen und den Therapieverlauf nicht über Gebühr zu beanspruchen. Nimmt man dennoch zusätzliche Facetten in den Blick und verlässt die Kategorie der störungsspezifischen Instrumente, so bieten sich etwa präliminar oder im Verlauf einige Maße an,

die den (Miss-)Erfolg moderieren können: begünstigende bzw. hinderliche motivationale Faktoren (Geissner & Ivert, 2019, 2021; Grosse-Holtforth & Grawe, 2002; Schulte, 2015), ebenso wie bei den Zielgrößen Merkmale zum lebensweltlichen Kontext. Die Auswahl zusätzlicher Instrumente zur Erfassung therapeutischer Wirkfaktoren oder Moderatoren des Behandlungserfolgs wird sich jedoch, so ist zu erwarten, im Unterschied zu den störungsspezifischen Instrumenten mehr am gewählten therapeutischen Ansatz orientieren.

Honorierung innerhalb der Psychotherapie: Lutz et al. (2019, S. 69f.) schildern den nicht ganz unwichtigen Gesichtspunkt der finanziellen Erlöseseite. Der Einsatz der hier vorgestellten störungsspezifischen Instrumente für Diagnostik und Evaluation gilt als Bestandteil von Psychotherapie und rechtfertigt die Abrechenbarkeit – ausschlaggebend ist dabei der Zeitaufwand, den der*die Psychotherapeut*in hat. Es sind stets mehrere Messungen genehmigungsfähig, sodass – im Einklang mit dem zuvor erwähnten System an multiplen Untersuchungszeitpunkten – in Summe störungsspezifisch

mit 40 bis 150 €/Pat. gerechnet werden kann (vgl. auch KBV, 2020; Lutz & Weinmann-Lutz, 2021; Schulte, 2013).

Einschränkungen: Nicht zu allen psychischen Störungen oder spezifizierten Unterformen konnten Verfahren vorgestellt werden, teils auch aus Platzgründen. Der Bereich der Persönlichkeitsstörungen etwa – ICD-10-Codierung F60/F61 – gestaltete sich in der Bearbeitung schwierig: Zum einen wegen der durchweg traitartigen und damit nicht für Outcome-Messung infrage kommenden Ausrichtung der existierenden Verfahren (Beispiel: *ADP IV*; Doering et al., 2007; MedUni Wien, o. J., online), zum anderen aufgrund der Umbruchsituation hin zur ICD-11 (o. J., online; Hauser et al., 2021) mit einer gänzlich neuen Struktur.

Im Bereich der in der Praxis keineswegs selten vorkommenden Diagnose-Gruppen Z73/Z56, also mit Bezug zu Burnout, Stress und Arbeitswelt, führte die Inspizierung verfügbarer Instrumente nicht zu überzeugenden Empfehlungen. In Bereichen wie etwa Angst wiederum liegen mehrere Messverfahren vor, so dass dort aus Gründen der Übersichtlichkeit, vor allem auch der Gleichbehandlung mit anderen Störungen, nur eines ausgewählt werden konnte. Nochmals anders gelagert ist die Situation im Bereich Alkohol: Durch eine deutliche Empfehlung der WHO zugunsten des Fragebogens *AUDIT* werden bestehende Instrumente trotz einer Bewährung im Feld etwas in den Hintergrund rücken.

Aber auch dies ergaben die Recherchen: Bei nicht wenigen Störungen lag nur ein einziges für Therapieevaluation diskutables Verfahren vor: Beispiel bipolare Störung, Beispiel sexuelle Funktionsstörungen. Die Breite und Tiefe dortiger Entwicklungsarbeiten kommt noch nicht an jene ausgereifter, manualisierter Instrumente heran, trotz klarer künftiger Perspektiven.

Hinzuweisen ist schließlich auf Spezialfälle wie etwa die Autismusspektrumstörung F84.5 (Erwachsene), der zurzeit eine recht intensive Forschungsaktivität zu attestieren ist: Der diagnostische Prozess ist nicht unaufwändig und verlangt Detailkenntnisse, die eher an Spezialambulanzen gegeben sind (vgl. Dziobek & Tebartz van Elst, 2021).

Ausblick: Für die Zukunft ist sicher auch im Bereich der psychodiagnostischen Verfahren zunehmend mit digitalen Lösungen zu rechnen – mit entsprechenden Vorteilen wie der Vermeidung von Auswertungsfehlern und dem zeitlichen Aufwand zur Auswertung klassischer „analoger“ Paper-pencil-Fragebögen. Beispiele hierfür sind das Hogrefe Testsystem HTS, Klenico oder TestOS. Grenzen dieser Systeme liegen aktuell unter anderem darin, dass sie jeweils nur eine Auswahl der für die Praxis relevanten Fragebogeninstrumente abbilden. Für die Zukunft entscheidend wird ferner sein, welche Schnittstellen diese Systeme zu Abrechnungssystemen, zu einer direkten Übernahme der Testergebnisse, z. B. in Be-

fundberichte, oder zur Rückmeldung der Testergebnisse an Patient*innen angeboten werden.

Für das wohl wichtigste Kriterium, die Validität des zu messenden Gegenstands, gilt realistischerweise generell, wenn gleich etwas banal: Praxis der Evaluation von Psychotherapie kann nur auf tatsächlich Vorhandenes, sprich auf bestehende Untersuchungsverfahren zurückgreifen. Oft erweist sich die inhaltlich intendierte Zielgröße auf der Messebene jedoch als unscharf umgesetzt, d. h. das Konstrukt ist für die Untersuchung nicht optimal getroffen (Flake & Fried, 2020). Auf der anderen Seite ist das Feld der klinischen Symptome, Syndrome und Diagnosen auch Gegenstand von Wertungen und Einigungen in Wissenschaft und Praxis, es unterliegt Wandlungen über die Zeit. Die Validitätsfrage mag also immer ein wenig vorläufig sein. Dass es sich hierbei um einen Prozess und eine immer wieder erforderliche neue Konsensfindung handelt, sehen wir derzeit beispielsweise wie erwähnt bei der Neufassung der Systematik von Persönlichkeitsstörungen (Hauser et al., 2021). Aber auch komplett neue Strukturierungsvorschläge für das Gesamtsystem psychischer Störungen stehen im Raum, etwa die Hierarchical Taxonomy of Psychopathology – HiTOP (Conway et al., 2021), die in der nordamerikanischen klinischen Psychologie gegenwärtig stark diskutiert wird und in vielerlei Hinsicht ein ausgesprochen spannendes Entwicklungspotenzial bietet.

Praxisempfehlungen

Für Psychotherapie nach Richtlinienverfahren steht mehr und mehr die verbindliche Verwendung standardisierter Instrumente für Individualdiagnostik und Qualitätssicherung zur Diskussion. Für die Patient*innen sollten dabei störungsspezifische Verfahren zur Objektivierung und Verlaufsdagnostik bezogen auf die Behandlungsdiagnose(n) genutzt werden. Diese Messungen sollten aussagekräftige Beurteilungen der erzielten Veränderung (Verbesserung) ermöglichen, also zwischen **einerseits** Anmeldung bzw. Behandlungsaufnahme und **andererseits** Abschluss bzw. Katamnese.

Zusätzlich kann der Einsatz von Zwischenmessungen hilfreiche Anregungen bei therapeutischen Entscheidungen oder auch in der Supervision liefern. Als gute Praxis wäre hier vorzuschlagen, die Ergebnisse transparent mit dem*der Patient*in zu besprechen und Therapiefortschritte zu thematisieren bzw. weiteren Behandlungsbedarf zu klären.

Hinweis

Ein ausführliches Literaturverzeichnis zu diesem Beitrag mit allen Angaben zu den besprochenen Messverfahren sowie zu Einleitung und Diskussion findet sich als kostenloser Download im elektronischen Supplement des Psychotherapeutenjournals. Gehen Sie hierzu bitte auf www.psychotherapeutenjournal.de, dort auf „Aktuelle Ausgabe“ (späterhin auf „Archiv“ – mit der Heftangabe 4/2021) und klicken Sie „Literaturverzeichnis Artikel Geissner & Koch“ an.

Literatur

Geissner, E., Ivert, P. & Schmitt, M. (2021a). Geeignete Messzeitpunkte in der Psychotherapieevaluation – Ein Vorschlag. *Verhaltenstherapie – Serie Expertendiskussion*, 31 (3), 248–254.

Geissner, E., Ivert, P. & Schmitt, M. (2021b). Proposal of Measurement Occasions for Unbiased Evaluation of Psychotherapy. *Verhaltenstherapie – Expert Discussion*. Online March, 11, 2021, DOI 10.1159 / 0005145; english fulltext version of Geissner et al. 2021a; free Access (PDF).

Lutz, W., Neu, R. & Rubel, J. (2019). *Evaluation und Effekterfassung in der Psychotherapie*. Göttingen: Hogrefe.

Rief, W., Schramm, E. & Strauß, B. (Hrsg.). (2021). *Psychotherapie – Ein kompetenzorientiertes Lehrbuch*. München: Elsevier (mehrere Kapitel aus diesem Werk werden im vorliegenden Beitrag zitiert, siehe ausführlich im elektronischen Supplement).



Prof. Dr. Edgar Geissner

Department Psychologie
Abt. Klinische Psychologie und
Psychotherapie
Ludwig-Maximilians-Universität
München, Leopoldstr. 13
80802 München
edgar.geissner@campus.lmu.de

Prof. Dr. Dipl.-Psych. Edgar Geissner ist Psychologischer Psychotherapeut und als Dozent und Supervisor in der postgradualen Aus-, Fort- und Weiterbildung wie auch als Supervisor*innenausbilder tätig; er gibt Seminare für (angehende) klinische Führungskräfte. Seine Forschungsschwerpunkte liegen in der Diagnostik, Outcome-Messung und in der Erfassung differentiell-psychologischer Effekte von Psychotherapie.



Dr. Stefan Koch

Schön Klinik Roseneck
Am Roseneck 6
83209 Prien am Chiemsee
skoch@schoen-klinik.de

Dr. Dipl.-Psych. Stefan Koch, ist Psychologischer Psychotherapeut; er arbeitet als Leitender Psychologe an der Schön Klinik Roseneck sowie als Supervisor und als Dozent in der Psychotherapieausbildung. In der Forschung beschäftigt er sich mit beruflichem Stress und psychischen Erkrankungen, mit Zwangsstörungen und mit Psychotherapieevaluation.

Von A wie Angst bis Z wie Zwang – Störungsspezifische Messinstrumente zur Evaluation von Psychotherapie (Erwachsene)

Edgar Geissner & Stefan Koch

TEIL 1

Literatur im Beitrag allgemein – Einleitungsteil, Diskussionsteil

- BPTK – Bundespsychotherapeutenkammer. (2020). Empfehlungen der BPTK für die Dokumentation psychotherapeutischer Behandlungen in der psychotherapeutischen Versorgung. 37. Deutscher Psychotherapeutentag am 14. November 2020. Verfügbar unter: www.bptk.de (26.07.2021).
- BPTK – Bundespsychotherapeutenkammer. (2021). Musterweiterbildungsordnung - Tätigkeitsprofile als Ziele der Weiterbildung. 38. Deutscher Psychotherapeutentag am 24. April 2021. Verfügbar unter: www.bptk.de (26.07.2021).
- Bundesgesetzblatt (2019). Gesetz zur Reform der Psychotherapeutenausbildung vom 15. November 2019. Köln: Bundesanzeiger Verlag, 1604–1621.
- Bundesgesetzblatt (2020). Approbationsordnung für Psychotherapeutinnen und Psychotherapeuten vom 4. März 2020. Köln: Bundesanzeiger Verlag, 448–483.
- Conway, C., Krueger, R. & HiTOP Consortium Executive Board (2021). Rethinking the Diagnosis of Mental Disorders. *Current Directions in Psychological Science*, 30 (2), 151–158.
- Deutscher Ethikrat (2016). Patientenwohl als ethischer Maßstab für das Krankenhaus, Kapitel 3.3.2 - Qualitätsmessung. Berlin: Deutscher Ethikrat. Verfügbar unter: www.ethikrat.org (26.07.2021).
- Doering, S., Renn, D., Höfer, S., Smrekar, U., Janecke, N., Schatz, D., Schotte, C., de Doncker, D. & Schüßler, G. (2007). Validierung der deutschen Version des Fragebogens zur Erfassung von DSM-IV Persönlichkeitsstörungen (ADP-IV). *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 53 (2), 111–128.
- Dziobek, I. & Tebartz van Elst, L. (2021). Autismus-Spektrum-Störungen. In W. Rief, E. Schramm & B. Strauß (Hrsg.), *Psychotherapie – Ein kompetenzorientiertes Lehrbuch* (S. 365–374). München: Elsevier Deutschland.
- Flake, J. & Fried (2020). Measurement – Schmeasurement: Questionable Measurement Practices and how to Avoid Them. *Advances in Methods and Practices in Psychological Science*, 3 (4) 456–465.
- Geissner, E., Bach, U. & Ivert, P. (2010). Globale Therapie- und Zielerreichungsbeurteilungen als unaufwändige und valide Hilfsmittel in der Evaluation psychotherapeutischer Behandlungen. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 39, Supp. 1, S. 11.
- Geissner, E., Ivert, P. & Schmitt, M. (2021a). Geeignete Messzeitpunkte in der Psychotherapieevaluation – Ein Vorschlag. *Verhaltenstherapie – Serie Expertendiskussion*, 31 (3), 248–254.

- Geissner, E., Ivert, P. & Schmitt, M. (2021b). Proposal of Measurement Occasions for Unbiased Evaluation of Psychotherapy. *Verhaltenstherapie – Expert Discussion* (engl. Volltext-Version von Geissner et al., 2021a). Verfügbar unter: <https://doi.org/10.1159/000514541> (26.07.2021).
- Hiller, W., Bleichhardt, G. & Schindler, A. (2009). Evaluation von Psychotherapien aus der Perspektive von Qualitätssicherung und Qualitätsmanagement. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 35, 225–233.
- Hiller, W., Schindler, A., Andor, T. & Rist, F. (2011). Vorschläge zur Evaluation regulärer Psychotherapien an Hochschulambulanzen. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 40 (1), 22–32.
- ICD-11 – WHO (Update Mai 2021). *International Classification of Diseases (11th Revision)*, Chapter 6, Mental and Behavioural Disorders. Verfügbar unter: <https://icd.who.int/browse11/l-m/o.int/browse11/l-m> (26.07.2021).
- KBV – Kassenärztliche Bundesvereinigung. (2020). *Praxisnachrichten* 23.04.2020. Verfügbar unter: www.kbv.de/html/1150_45852.php (26.07.2021).
- Lutz, W., Neu, R. & Rubel, J. (2019). *Evaluation und Effekterfassung in der Psychotherapie*. Göttingen: Hogrefe.
- Lutz, W. & Weinmann-Lutz, B. (2021). Abrechenbarkeit von Fragebögen in der Psychotherapie und EBM-Leistungsbeschreibung. Persönliche Mitteilung an Edgar Geissner. E-Mail vom 29.07.2021.
- MedUni Wien. (o. J.). ADP-IV-Fragebogen. www.meduniwien.ac.at_diagnostic-downloads/diagnostik-von-persoenslichkeitsstoerungen (26.07.2021).
- Moritz, S., Krieger, E., Bohn, F. & Veckenstedt, R. (2017). *MKT Plus – Individualisiertes Metakognitives Therapieprogramm für Menschen mit Psychose. Anhang Evaluationsinstrumente, PANSS* (2. Aufl.; S. 147–152). Berlin: Springer.
- PTK Bayern. (2014). *Berufsordnung. Bayerische Landeskammer der Psychologischen Psychotherapeuten und der Kinder- und Jugendlichenpsychotherapeuten*. Verfügbar unter: www.ptk-bayern.de (26.07.2021).
- PTK Bayern. (2021). *Empfehlungen für die Dokumentation psychotherapeutischer Behandlungen. Bayerische Landeskammer der Psychologischen Psychotherapeuten und der Kinder- und Jugendlichenpsychotherapeuten. Mitgliederrundschreiben, Frühjahr 2021, 8–9*. Verfügbar unter: www.ptk-bayern.de (26.07.2021).
- PTK Berlin. (2020). *Mitteilungen der Psychotherapeutenkammer Berlin – Qualität in der Psychotherapie. (Wie) sollte man sie messen?* *Psychotherapeutenjournal*, 19 (3), 292.
- PTK Hessen. (2020). *Mitteilungen der Psychotherapeutenkammer Hessen – Qualitätssicherung in der ambulanten Psychotherapie. Chancen nicht verspielen*. *Psychotherapeutenjournal*, 18 (4), 303–304.
- Rief, W., Schramm, E. & Strauß, B. (2021) (Hrsg.). *Psychotherapie – Ein kompetenzorientiertes Lehrbuch*. München: Elsevier Deutschland.
- Schiepek, G., Kratzer, L., Hülsner, Y. & Bachler, E. (2019). *Prozessmonitoring in der Psychotherapie, Anspruch und Nutzen*. *Psychotherapeutenjournal*, 18 (4), 357–363.

- Schmidt, J. (2008). Fragebogen zur Messung der Patientenzufriedenheit – ZUF 8. In J. Bengel, M. Wirtz, C. Zwingmann & L. Lyssenko (Hrsg.), *Diagnostische Verfahren in der Rehabilitation* (S. 255–258). Göttingen: Hogrefe.
- Schmidt, J., Lamprecht, F. & Wittmann, W. (1989). Zufriedenheit mit der stationären Versorgung – Entwicklung eines Fragebogens. *Psychotherapie, Psychosomatik und Medizinische Psychologie*, 39 (7), 248–255.
- Schulte, D. (2013). Lohnt sich Diagnostik? Diagnostik lohnt sich – auch finanziell! Verfügbar unter: www.dgvt.de/aktuell/details/article/lohnt-sich-diagnostik/ (26.07.2021).
- Stangier, U. (2019). Prozessbasierte Kognitive Verhaltenstherapie. *Psychotherapeutenjournal*, 18 (3), 236–243.
- Strauß, B. (2019). Innovative Psychotherapieforschung – Wo stehen wir und wo wollen wir hin? *Psychotherapeutenjournal*, 18 (3), 4–10.

TEIL 2

Literatur zu den einzelnen Messinstrumenten

Agoraphobie und Panikstörung

- Ehlers, A., Margraf, J. & Chambless, D. (2001). AKV – Fragebogen zu körperbezogenen Ängsten, Kognitionen und Vermeidung. Zweite überarbeitete und neunormierte Auflage. Göttingen: Beltz Test.

Alkohol

- Barbor, T., Higgins-Biddle, J., Saunders, J. & Monteiro, M. (2001). AUDIT – The Alcohol Use Disorders Identification Test. Guidelines for Use in Primary Care. World Health Organization – Formal Publication (S. 1 – 41). Verfügbar unter: <http://hqlibdoc.who.int> (26.07.2021).
- Bundesärztekammer. (o. J.). AUDIT Fragebogen und AUDIT Auswertungsschema. Verfügbar unter: www.bundesaerztekammer.de (26.07.2021).
- DGPPN/DG-Sucht. (2020). Screening, Diagnose und Behandlung alkoholbezogener Störungen. Verfügbar unter: www.amwf.org/leitlinien/detail/II/076-001 (26.07.2021).
- Dybek, I. (2008). Screeningverfahren zur Entdeckung alkoholbezogener Störungen in Allgemeinarztpraxen. Dissertation, Universität zu Lübeck. Verfügbar unter: D-nb.info/995572267/34 (26.07.2021).

Reinert, D. & Allen, J. (2007) The Alcohol Use Disorders Identification Test: An Update of Research Findings. *Alcoholism – Clinical and Experimental Research*, 32 (2), 185–199.

Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen im Erwachsenenalter

AWMF. (2017). Langfassung der interdisziplinären evidenz- und konsensbasierten (S3) Leitlinie „ADHS bei Kindern, Jugendlichen und Erwachsenen“ (Registriernummer 028-045). Verfügbar unter: <https://www.awmf.org/leitlinien/detail/II/028-045.html> (24.06.2021).

Paucke, M., Stark, T., Exner, C., Kallweit, C., Hegerl, U. & Strauß, M. (2018). Aufmerksamkeitsdefizit-/Hyperaktivitätssyndrom (ADHS) und komorbide psychische Erkrankungen: ADHS-spezifische Selbstbeurteilungsskalen bei der Differenzialdiagnostik. *Nervenarzt*, 89 (11), 1287–1293.

Rösler, M., Retz-Junginger, P., Retz, W. & Stieglitz, R.-D. (2021). HASE - Homburger ADHS-Skalen für Erwachsene (2., überarb. u. erg. Aufl.). Göttingen: Hogrefe.

Rösler, M., Retz, W., Thome, J., Schneider, M., Stieglitz, R.-D. & Falkai, P. (2006). Psychopathological rating scales for diagnostic use in adults with attention-deficit/hyperactivity disorder (ADHD). *European Archives of Psychiatry and Clinical Neuroscience*, 256 (Suppl. 1), I/3-I/11.

Bipolare Störungen

Altman, E., Hedeker, D., Peterson, J. & Davis, J. (1979). The Altman Self Rating Mania Scale ASRM. *Biological Psychiatry*, 42, 948–955.

Bernhard, B. & Meyer, T. (2012). Bipolare Störungen. In G. Meinlschmidt, S. Schneider & J. Margraf (Hrsg.), *Lehrbuch der Verhaltenstherapie, Band 4: Materialien für die Psychotherapie* (S. 323–333). Berlin: Springer.

Meyer, T. (2021a). Manie und Hypomanie. In W. Rief, E. Schramm & B. Strauß (Hrsg.), *Psychotherapie – Ein kompetenzorientiertes Lehrbuch* (S. 99–108). München: Elsevier Deutschland.

Meyer, T. (2021b). Persönliche Mitteilung an Edgar Geissner. E-Mail vom 23.04.2021.

Meyer, T., Crist, N., LaRosa, N., Biyun, Y., Soares, J. & Bauer, I. (2020). Are Existing Self-ratings of Acute Manic Symptoms in Adults Reliable and Valid? A Systematic Review. *Bipolar Disorders*, 22 (6), 558–568.

Borderline-Persönlichkeitsstörung

Bohus, M., Kleindienst, N., Limberger, M. F., Stieglitz, R. D., Domsalla, M., Chapman, A. L. et al. (2009). The short version of the Borderline Symptom List (BSL-23): development and initial data on psychometric properties. *Psychopathology*, 42 (1), 32–39.

Bohus, M., Limberger, M.F., Frank, U., Chapman, A. L., Kühler, T. & Stieglitz, R.-D. (2007). Psychometric Properties of the Borderline Symptom List (BSL). *Psychopathology*, 40 (2), 126–132.

Bohus, M., Limberger, M. F., Frank, U., Sender, I., Gratwohl, T., Stieglitz, R. D. et al. (2001). Entwicklung der Borderline-Symptom-Liste. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 51, 201–211.

Zentralinstitut für Seelische Gesundheit Mannheim. (o. J.). Informationen und Downloads BSL-23/BSL-95 – Fragebögen. Verfügbar unter: <https://www.zi-mannheim.de/forschung/abteilungen-ags-institute/psm/psm-informationen-downloads.html> (18.06.2021).

Computerspielabhängigkeit

Müller, A. (2021). Stoffungebundene Suchterkrankungen: Verhaltenssuchte. In W. Rief, E. Schramm & B. Strauß (Hrsg.), *Psychotherapie – Ein kompetenzorientiertes Lehrbuch* (S. 227–233). München: Elsevier Deutschland.

Rehbein, F., Baier, D., Kleinmann, M. & Mößle, T. (2015). CSAS Computerspielabhängigkeitsskala: Ein Verfahren zur Erfassung der Internet Gaming Disorder nach DSM-5. Göttingen: Hogrefe.

Depression

Beck, A.T., Steer, R. & Brown, G. (1996). Beck Depression Inventory – Second Edition. Manual. San Antonio, TX: The Psychological Corporation.

Hautzinger, M., Keller, F. & Kühner, C. (2016). BDI II – Beck Depressions-Inventar Revision. Manual und Fragebogen. Frankfurt: Pearson Assessment Deutschland.

Dissoziative Störungen

Bernstein, E. M. & Putnam, F. W. (1986). Development, reliability, and validity of a dissociation scale. *Journal of Nervous and Mental Disease*, 174, 727–735.

Lysenko, L., Schmahl, C., Bockhacker, L., Vonderlin, R., Bohus, M. & Kleindienst, N. (2018). Dissociation in psychiatric disorders: a meta-analysis of studies using the Dissociative Experiences Scale. *American Journal of Psychiatry*, 175 (1), 37–46.

Spitzer, C., Stieglitz, R.-D. & Freyberger, H.J. (2021). Fragebogen zu dissoziativen Symptomen – Ein Selbstbeurteilungsverfahren zur syndromalen Diagnostik dissoziativer Phänomene (4., korr. u. akt. Aufl.). Bern: Hogrefe.

Essstörungen

Garner, D. (1991). Eating Disorder Inventory 2 – Manual. Odessa, Florida: Psychological Assessment Resources.

Paul, T. & Thiel, A. (2005). EDI 2 - Eating Disorder Inventory 2. Deutsche Version. Manual und Fragebogen. Göttingen: Hogrefe.

Hypochondrie

- Hiller, W. & Rief, W. (2004). Internationale Skalen für Hypochondrie. Deutschsprachige Adaptation des Whiteley-Index (WI) und der Illness Attitude Scales (IAS). Bern: Huber.
- Kellner, R. (1986). Somatization and Hypochondriasis. New York: Praeger – Greenwood Publishers.
- Pilowsky, I. (1997). Abnormal Illness Behaviour. Chichester: Wiley.

Posttraumatische Belastungsstörung

- Blevins, C. A., Weathers, F. W., Davis, M. T., Witte, T. K. & Domino, J. L. (2015). The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress*, 28, 489–498.
- Cloitre, M., Shevlin, M., Brewin, C.R., Bisson, J. I., Roberts, N.P., Maercker, A., Karatzias, T. & Hyland, P. (2018). The International Trauma Questionnaire: development of a self-report measure of ICD-11 PTSD and complex PTSD. *Acta Psychiatrica Scandinavica*, 138 (6), 536–546.
- Krüger-Gottschalk, A., Knaevelsrud, C., Rau, H., Dyer, A., Schäfer, I., Schellong, J. & Ehring, T. (2017). The German version of the Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): psychometric properties and diagnostic utility. *BMC Psychiatry*, 17, 379.
- Lueger-Schuster, B., Knefel, M., Maercker, A., 2018. International Trauma Questionnaire (Deutsche Übersetzung). Verfügbar unter: <https://www.traumameasuresglobal.com/itq>. Für Auswertungshinweise: <https://www.psychologie.uzh.ch/dam/jcr:ad66be44-4cd9-44c3-9911-488253de04cc/ITQ%20-%20Fragebogen.pdf> (26.07.2021).
- Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). PTSD Checklist for DSM-5 (PCL-5). Verfügbar unter: <https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp> (28.07.2021).
- Wortmann, J. H., Jordan, A. H., Weathers, F. W., Resick, P. A., Dondanville, K. A., Hall-Clark, B., Foa, E. B., Young-McCaughan, S., Yarvis, J., Hembree, E. A., Mintz, J., Peterson, A. L., & Litz, B. T. (2016). Psychometric analysis of the PTSD Checklist-5 (PCL-5) among treatment-seeking military service members. *Psychological Assessment*, 28, 1392–1403.

Psychosen

- Kay, S., Opler, L. & Fiszbein, A. (2000). Positive and Negative Syndrome Scale PANSS. Toronto: Multi-Health Systems.
- Moritz, S. (2021). Persönliche Mitteilung an Edgar Geissner. E-Mail vom 06.05.2021.
- Moritz, S., Krieger, E., Bohn, F. & Veckenstedt, R. (2017 II). MKT Plus – Individualisiertes Metakognitives Therapieprogramm für Menschen mit Psychose. Anhang Evaluationsinstrumente, PANSS (S. 149–152). Berlin: Springer.
- Opler, L., Kay, S., Lindenmayer, J.-P. & Fiszbein, A. (1999). SCI-PANSS. Structured Clinical Interview for Positive and Negative Syndrome Scale. Toronto: Multi-Health Systems.

Van der Gaag, M., Hoffman, T., Remijnsen, M. Hijman, R., de Haan, L., van Meijel, B. et al. (2006). The five-factor model of the Positive and Negative Syndrome Scale II. A ten-fold cross-validation of a revised model. *Schizophrenia Research*, 85, 280–287.

Schlafstörungen

Dieck, A., Morin, C.M. & Backhaus, J. (2018). A German version of the Insomnia Severity Index – Validation and identification of a cut-off to detect insomnia. *Somnologie*, 22, 27–35.

Gerber, M., Lang, C., Lemola, S., Colledge, F., Kalak, N., Holsboer-Trachsler, E. et al. (2017). Validation of the German version of the insomnia severity index in adolescents, young adults and adult workers: results from the cross-sectional studies. *BMC Psychiatry*, 16, 174.

Morin, C.M. (1993). *Insomnia: Psychological Assessment and Management*. New York: Guilford.

Riemann, D., Baum, E., Cohrs, S., Crönlein, T., Hajak, G., Hertenstein, E. et al. (2017). S3-Leitlinie Nicht erholsamer Schlaf/Schlafstörungen. *Somnologie*, 21, 2–44.

Schmerz

Geissner, E. (1996). *Die Schmerzempfindungs-Skala (SES)*. Handanweisung, Fragebogen. Göttingen: Hogrefe.

Geissner, E. (2001). *Fragebogen zur Erfassung der Schmerzverarbeitung (FESV) - Manual*. Göttingen: Hogrefe.

Martin, A. & Henningsen, P. (2021). Somatoforme Störungen, somatische Belastungsstörungen und Schmerzsyndrome. In W. Rief, E. Schramm & B. Strauß (Hrsg.), *Psychotherapie – Ein kompetenzorientiertes Lehrbuch* (S. 235–256). München: Elsevier Deutschland.

Nilges, P., Rief, W., Treede, R. & Zenz, M. (2021). Chronischer Schmerz ist weder eine psychische noch eine funktionelle Störung, sondern (endlich) eine eigenständige Diagnose. *Nervenarzt*, 92, 716–717.

Nobis, H.-G., Rolke, R. & Graf-Baumann, T. (Hrsg.). (2020). *Schmerz – eine Herausforderung*. Berlin: Springer.

Sexuelle Funktionsstörungen

Hoyndorf, S., Reinhold, M. & Christmann, F. (1995). *Behandlung sexueller Störungen*. Weinheim: Beltz.

Müller, M. (2007). Development and Retest Reliability of a German Version of the Sexual Behaviour Questionnaire SBQ-G. *Archives of Andrology*, 53, 67–69

Müller, M. (2016). *SBQ-G Sexual Behaviour Questionnaire / Fragebogen zur Sexualität FZS*. Leibniz Institut für Psychologie (vormals ZPID) – Elektronisches Testarchiv. Verfügbar unter: www.testarchiv.eu/de/test/SBQ-G (26.07.2021).

Reininger, K. & Briken, P. (2021). Sexualstörungen – Sexuelle Funktionsstörungen, Paraphilien, Geschlechtsidentitätsstörungen. In W. Rief, E. Schramm & B. Strauß (Hrsg.), *Psychotherapie – Ein kompetenzorientiertes Lehrbuch* (S. 303–321). München: Elsevier Deutschland.

Strauß, B., Richter, D. & Brähler, E. (2013). Diagnostik sexueller Störungen. In P. Briken & M. Berner (Hrsg.), *Praxisbuch Sexuelle Störungen* (S. 92–99). Stuttgart: Thieme.

Somatoforme Störungen

Martin, A. & Henningsen, P. (2021). Somatoforme Störungen, somatische Belastungsstörungen und Schmerzsyndrome. In W. Rief, E. Schramm & B. Strauß (Hrsg.), *Psychotherapie – Ein kompetenzorientiertes Lehrbuch* (S. 235–256). München: Elsevier Deutschland.

Rief, W. & Hiller, W. (2019). SOMS – Screening für Somatoforme Störungen. Dritter Nachdruck der zweiten vollständig überarbeiteten und neu normierten Auflage (Fragebogen „SOMS – 7 T“). Göttingen: Hogrefe.

Soziale Phobie

Bandelow, B., Aden, I., Alpers, G. W., Benecke, C., Beutel, M. E., Deckert, J., Domschke, K. et al. (2021). Deutsche S3-Leitlinie Behandlung von Angststörungen, Version 2. Verfügbar unter: www.awmf.org/leitlinien/detail/II/051-028.html (27.07.2021).

Consbruch, K., Stangier, U. & Heidenreich, T. (2016). SOZAS – Skalen zur Sozialen Angststörung. Göttingen: Hogrefe.

Liebowitz, M. R. (1987). Social Phobia. *Modern problems of pharmacopsychiatry*, 22, 141–173.

Stangier, U. & Heidenreich, T. (2015). Liebowitz Social Anxiety Scale (LSAS). In *Collegium Internationale Psychiatriae Scalarum* (Hrsg.), *Internationale Skalen für Psychiatrie* (6., überarb. u. erw. Aufl., S. 307-315). Göttingen: Beltz Test.

Stangier, U. Schramm, E., Heidenreich, T., Berger, M. & Clark, D.M. (2011). Cognitive therapy vs. interpersonal psychotherapy in social anxiety disorder. *Archives of General Psychiatry*, 68 (7), 692–700.

Trauer

Lumbeck, G., Brandstätter, M. & Geissner, E. (2012). Erstvalidierung der deutschen Version des „Inventory of Complicated Grief“ (ICG-D). *Zeitschrift für Klinische Psychologie und Psychotherapie*, 41, 243–248.

Prigerson, H., Maciejewski, P., Reynolds, C., Bierhals, A. J., Newsom, J. T., Fasiczka, A. et al. (1995). Inventory of Complicated Grief: A scale to measure maladaptive symptoms of loss. *Psychiatry Research*, 59, 65–79.

Rosner, R., Pfoh, G., Rojas, R., Brandstätter, M., Rossi, R., Lumbeck, G., Kotoucova, M., Hagl, M. & Geissner, E. (2015). Anhaltende Trauerstörung. *Manuale für die Einzel- und Gruppentherapie*. Göttingen: Hogrefe.

Wagner, B. (2016). Wann ist Trauer eine psychische Erkrankung? – Trauer als diagnostisches Kriterium in der ICD-11 und im DSM-5. *Psychotherapeutenjournal*, 15 (3), 250–255.

Zwangsstörungen

Büttner-Westphal, H. & Hand, I. (1991). Yale-Brown Obsessive Compulsive Scale (Y-BOCS). *Verhaltenstherapie*, 1, 226–233.

Ertle, A. (2012). Zwangsstörung. In G. Meinlschmidt, S. Schneider & J. Margraf (Hrsg.), *Lehrbuch der Verhaltenstherapie*. Bd. 4: Materialien für die Psychotherapie (Kapitel 33). Berlin: Springer.

Federici, A., Summerfeldt, L. J., Harrington, J. L., McCabe, R. E., Purdon, C. L., Rowa, K. et al. (2010). Consistency between self-report and clinician-administred versions of the Yale-Brown Obsessive-Compulsive Scale. *Journal of Anxiety Disorders*, 24 (7), 729–733.

Hand, I. & Büttner-Westphal, H. (1991). Die Yale-Brown Obsessive Compulsive Scale (Y-BOCS): Ein halbstrukturiertes Interview zur Beurteilung des Schweregrades von Denk- und Handlungszwängen. *Verhaltenstherapie*, 1, 223–225.

Jacobsen, D., Kloss, M., Fricke, S., Hand, I. & Moritz, S. (2003). Reliabilität der deutschen Version der Yale- Brown Obsessive Compulsive Scale. *Verhaltenstherapie*, 13 (2), 111–113.

Kordon, A., Lotz-Rambaldi, W., Muche-Borowski, C. & Hohagen, F. (2013). S3-Leitlinie Zwangsstörungen (AWMF-Registernummer 038/017). Verfügbar unter: https://www.dgppn.de/_Resources/Persistent/29d7ea6c2c635d3626ddaf70eabf31386b9bbda8/S3-Leitlinie%20Zwangsstörungen%20Langversion%20Endversion%2014%2005%202013.pdf (17.06.2021).

Lakatos, A. & Reinecker, H. (1999). *Kognitive Verhaltenstherapie bei Zwangsstörungen. Ein Therapiemanual*. Göttingen: Hogrefe.