

Beschwerden- validierung in der Begutachtung, Klinik und Rehabilitation

Thomas Merten
(Hrsg.)

Beschwerdenvalidierung in der Begutachtung, Klinik und Rehabilitation

Thomas Merten
(Hrsg.)

Beschwerdenvalidierung in der Begutachtung, Klinik und Rehabilitation



Dr. Thomas Merten, geb. 1959. 1981–1986 Studium der Klinischen Psychologie in Berlin. 1991 Promotion. 2000 Habilitation. Seit 1992 Tätigkeit im Klinikum im Friedrichshain; seit 1993 umfassende Lehrtätigkeit, u. a. Humboldt-Universität zu Berlin, Universität Leipzig, Friedrich-Schiller-Universität Jena, Psychologische Hochschule Berlin; seit 1994 Gutachtentätigkeit. Forschungsschwerpunkte: Testentwicklung und angewandte Testdiagnostik, neuropsychologische Diagnostik, Beschwerdvalidierung, psychologische Begutachtung im Zivil-, Sozial- und Verwaltungsrecht.

Wichtiger Hinweis: Der Verlag hat gemeinsam mit den Autor:innen bzw. den Herausgeber:innen große Mühe darauf verwandt, dass alle in diesem Buch enthaltenen Informationen (Programme, Verfahren, Mengen, Dosierungen, Applikationen, Internetlinks etc.) entsprechend dem Wissensstand bei Fertigstellung des Werkes abgedruckt oder in digitaler Form wiedergegeben wurden. Trotz sorgfältiger Manuskripterstellung und Korrektur des Satzes und der digitalen Produkte können Fehler nicht ganz ausgeschlossen werden. Autor:innen bzw. Herausgeber:innen und Verlag übernehmen infolgedessen keine Verantwortung und keine daraus folgende oder sonstige Haftung, die auf irgendeine Art aus der Benutzung der in dem Werk enthaltenen Informationen oder Teilen davon entsteht. Geschützte Warennamen (Warenzeichen) werden nicht besonders kenntlich gemacht. Aus dem Fehlen eines solchen Hinweises kann also nicht geschlossen werden, dass es sich um einen freien Warennamen handelt.

Copyright-Hinweis:

Das E-Book einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlags unzulässig und strafbar.

Der Nutzer verpflichtet sich, die Urheberrechte anzuerkennen und einzuhalten.

Hogrefe Verlag GmbH & Co. KG
Merkelstraße 3
37085 Göttingen
Deutschland
Tel. +49 551 999 50 0
Fax +49 551 999 50 111
info@hogrefe.de
www.hogrefe.de

Satz: Sabine Rosenfeldt, Hogrefe Verlag GmbH & Co. KG, Göttingen
Format: PDF

1. Auflage 2023

© 2023 Hogrefe Verlag GmbH & Co. KG, Göttingen

(E-Book-ISBN [PDF] 978-3-8409-3180-2; E-Book-ISBN [EPUB] 978-3-8444-3180-3)

ISBN 978-3-8017-3180-9

<https://doi.org/10.1026/03180-000>

Nutzungsbedingungen:

Der Erwerber erhält ein einfaches und nicht übertragbares Nutzungsrecht, das ihn zum privaten Gebrauch des E-Books und all der dazugehörigen Dateien berechtigt.

Der Inhalt dieses E-Books darf von dem Kunden vorbehaltlich abweichender zwingender gesetzlicher Regeln weder inhaltlich noch redaktionell verändert werden. Insbesondere darf er Urheberrechtsvermerke, Markenzeichen, digitale Wasserzeichen und andere Rechtsvorbehalte im abgerufenen Inhalt nicht entfernen.

Der Nutzer ist nicht berechtigt, das E-Book – auch nicht auszugsweise – anderen Personen zugänglich zu machen, insbesondere es weiterzuleiten, zu verleihen oder zu vermieten.

Das entgeltliche oder unentgeltliche Einstellen des E-Books ins Internet oder in andere Netzwerke, der Weiterverkauf und/oder jede Art der Nutzung zu kommerziellen Zwecken sind nicht zulässig.

Das Anfertigen von Vervielfältigungen, das Ausdrucken oder Speichern auf anderen Wiedergabegeräten ist nur für den persönlichen Gebrauch gestattet. Dritten darf dadurch kein Zugang ermöglicht werden. Davon ausgenommen sind Materialien, die eindeutig als Vervielfältigungsvorlage vorgesehen sind (z. B. Fragebögen, Arbeitsmaterialien).

Die Übernahme des gesamten E-Books in eine eigene Print- und/oder Online-Publikation ist nicht gestattet. Die Inhalte des E-Books dürfen nur zu privaten Zwecken und nur auszugsweise kopiert werden.

Diese Bestimmungen gelten gegebenenfalls auch für zum E-Book gehörende Download-Materialien.

Inhaltsverzeichnis

Vorwort	9
Teil I Einführung in die Beschwerdvalidierung	
1 Beschwerdvalidierung in der psychologischen Begutachtung – Einführung <i>Thomas Merten</i>	13
2 Beschwerdvalidierung in der Medizin – Klinische Tests zur Beschwerdvalidierung <i>Bernhard Widder</i>	46
Teil II Methoden und Verfahren im Überblick	
3 Psychologische Methoden der Beschwerdvalidierung – Überblick und Systematik <i>Thomas Merten</i>	69
4 Beschwerdvalidierung mittels psychometrischer Selbstbeurteilungsverfahren <i>Ralf Dohrenbusch</i>	83
5 Kognitive Beschwerdvalidierungstests <i>Matthias Henry</i>	99
6 Eingebettete kognitive Validitätskennwerte <i>Thomas Merten</i>	115
7 Was können Realkennzeichen (nicht)? – Zur Anwendbarkeit der Methodik von Glaubhaftigkeitsbegutachtungen auf die Beschwerdvalidierung <i>Jonas Schemmel und Renate Volbert</i>	121
Teil III Testinformation für ausgewählte Einzelverfahren	
8 Aggravations- und Simulationstest (AST 4.0)	137
9 Amsterdamer Kurzzeitgedächtnistest (AKGT)	140
10 Beschwerdvalidierungstest (BEVA)	144

11	Fifteen Item Test (FIT), Rey Memory Test (RMT)	146
12	Groninger Effort Test (GET)	149
13	Inventory of Problems–29 (IOP-29)	152
14	Minnesota Multiphasic Personality Inventory (MMPI-2, MMPI-2-RF)	156
15	Morel Emotional Numbing Test for Posttraumatic Stress Disorder (MENT) ..	159
16	Self-Report Symptom Inventory (SRSI)	162
17	Structured Interview of Reported Symptoms – 2 (SIRS-2)	166
18	Strukturierter Fragebogen Simulierter Symptome (SFSS).....	170
19	Test of Memory Malingering (TOMM)	173
20	Verhaltens- und Erlebensinventar (VEI)	176
21	Word Memory Test (WMT), Medical Symptom Validity Test (MSVT), Non-Verbal Medical Symptom Validity Test (NV-MSVT)	179

Teil IV Spezielle Zielgruppen und Untersuchungskontexte

22	Beschwerdengültigkeit in der medizinischen Rehabilitation <i>Wolfgang Pfeiffer, Stève Buchli und Andreas Stevens</i>	187
23	Beschwerdengültigkeit in der klinischen Untersuchung von Erwachsenen mit ADHS <i>Anselm Fuermaier, Miriam Becke, Oliver Tucha und Lara Tucha</i>	200
24	Beschwerdengültigkeit im geriatrischen Bereich – Besonderheiten der Beschwerdengültigkeit bei Patienten mit Demenz und im höheren Lebensalter <i>Matthias Henry und Thomas Merten</i>	216
25	Begutachtung von geltend gemachten tatbezogenen Amnesien <i>Marko Jelacic und Peter Giger</i>	233
26	Diagnostik vorgetäuschter Suizidalität bei Strafgefangenen <i>Joseph H. Obegi</i>	247
27	Beschwerdengültigkeit im interkulturellen Kontext <i>Thomas Merten</i>	264

Teil V Beschwerdengültigkeit in der interdisziplinären Begutachtung

28	Beschwerdengültigkeit: Pro und Kontra <i>Andreas Stevens</i>	285
29	Modell der Zusammenarbeit zwischen psychiatrischen, neuropsychologischen und weiteren Disziplinen im Rahmen von komplexen versicherungsmedizinischen Begutachtungen <i>Gerhard Ebner und Stève Buchli</i>	298

Teil VI Beschwerdenvalidierung in den deutschsprachigen Ländern und in den Niederlanden	
30	Beschwerdenvalidierung in Deutschland <i>Gordon Krahl, Alexandra Ernsthaus und Christina Ackermann</i> 307
31	Beschwerdenvalidierung in Österreich <i>Johann Lehrner, Martina Dienstl, Renate Fanninger, Wolfgang Neuwirth und Thomas Bodner</i> 316
32	Beschwerdenvalidierung in der Schweiz <i>Giuseppe Di Stefano und Adrian Frei</i> 327
33	Beschwerdenvalidierung in den Niederlanden <i>Brechje Dandachi-FitzGerald, Anselm Fuermaier und Harald Merckelbach</i> 334
	Die Autorinnen und Autoren dieses Bandes 344
	Sachregister 348

Vorwort

Die mit diesem Band vorgelegte dritte deutschsprachige Buchveröffentlichung zur Beschwerdendvalidierung baut auf zwei vorangegangene aus den Jahren 2009 und 2014 auf und aktualisiert den Stand des Wissens, das sich in den vergangenen zehn Jahren auf diesem Sachgebiet sehr stark vermehrt hat. Die neuen Texte sollen sie nicht ersetzen, ebenso wie sich die ausführlichen Testrezensionen aus dem *Handbuch neuropsychologischer Testverfahren* von 2019 nicht wiederholen sollen. Durch die Zusammenstellung der hier aufgenommenen Themen und die Art ihrer Darstellung soll zum einen eine aktualisierte Diskussion des Entwicklungsstandes erfolgen, zum anderen sollen einige bislang im deutschen Sprachraum kaum oder nur wenig behandelte Themen im Umfeld der Diagnostik problematischer Beschwerdendarstellungen vorgestellt werden. Der frühere Fokus der Beschwerdendvalidierung auf ihre Einbettung in die gutachtliche Tätigkeit von Psychologen lässt sich in dieser Weise nicht mehr erkennen; entsprechend ist eine Reihe von Kapiteln auch im klinischen sowie im Rehabilitationskontext angesiedelt und einzelne Kapitel beziehen diese Anwendungsgebiete mit in ihre Darstellung ein. Weiteres zum Inhalt ist dem ersten Abschnitt des Einleitungskapitels zu entnehmen; eine Zusammenstellung der weiterführenden deutsch- und englischsprachigen Literatur, die über die behandelten Themen hinausreicht (welche teils schon aus Platzgründen nicht mehr berücksichtigt werden konnten), findet sich im letzten Abschnitt des Einleitungskapitels.

Nahezu alle kontaktierten angedachten Autoren haben positiv auf die Anfrage reagiert. Dies betrifft ebenso die Mehrzahl der kontaktierten Testautorinnen, die, anders als bei Testrezensionen üblich, gebeten wurden, kritisch ihre eigenen Verfahren dem potenziellen Anwender darzustellen. Allen an der Verfassung der Texte beteiligten Psychologinnen, Psychologen und Ärzten gilt mein großer Dank, ohne ihre engagierten Beiträge wäre dieses Buchprojekt natürlich nie zu verwirklichen gewesen. Ein Blick auf die Autorenliste verrät rasch, dass ein ganz wesentlicher Teil der in Sachen Beschwerdendvalidierung kompetenten Autoren und Autorinnen aus dem deutschen Sprachraum, neben einigen Beitragenden aus dem nicht deutschsprachigen Ausland, für das Projekt gewonnen werden konnten.

Dem Hogrefe Verlag und namentlich Frau Hindermann, Frau Kielhorn sowie, in hervorgehobener Weise, Frau Rothauge ist besonderer Dank geschuldet. Nicht

nur wurde die Konzeption für diesen Band rasch und unkompliziert, sondern auch mit höchster Professionalität ausgeführt, dem sehr ehrgeizigen initialen Zeitplan folgend, sodass eine hohe Aktualität der Darstellung zum Zeitpunkt der Veröffentlichung erreicht werden konnte, die der vieler psychologischer Fachzeitschriften nicht nachstehen sollte.

In den einzelnen Texten wird zum Teil abwechselnd die männliche und die weibliche Form benutzt, zum Teil auch zur einfacheren Lesbarkeit nur eine Geschlechter-Form von Personen. Wenn nicht ausdrücklich auf Einzelpersonen bezogen, sind bei Verwendung des weiblichen Geschlechts natürlich Personen jeglichen Geschlechts gemeint, wie es bei Verwendung des männlichen Geschlechts ganz genau so intendiert ist. Auf schwer lesbare Alternativen wurde bewusst verzichtet. Ebenfalls zur einfacheren Lesbarkeit wird zur Bezeichnung einer begutachteten Person vorwiegend der Begriff der *Probandin* oder des *Probanden* verwendet, der in der deutschen gutachtlichen Literatur häufig herangezogen wird. Je nach rechtlichem Kontext kann es sich dabei spezifischer um einen Kläger, eine Beklagte, einen Zeugen, das Mitglied eines Versorgungswerks, einen Beamten, eine Antragstellerin, eine versicherte oder eine geschädigte Person u.Ä.m. handeln. Von *Patienten* und *Patientinnen* wird ausschließlich dann gesprochen werden, wenn Bezug auf einen klinischen oder rehabilitativen Kontext genommen wird.

Die Leserin, die einen Kochbuchansatz für die Durchführung einer sachkundigen Beschwerdengültigkeit in diesem Buch zu finden erhofft, wird zwangsläufig enttäuscht werden. Kurz gesagt, ist die Materie mittlerweile viel zu komplex, als dass ein solcher auch nur im Ansatz vermittelt werden könnte. Die Aneignung von Kompetenzen auf diesem wichtigen Spezialgebiet, insbesondere für gutachtlich arbeitende Psychologen, ist kompliziert und aufwändig; dies zu unterschätzen, kann folgenreiche Auswirkungen zeitigen. Wer sich auf dem Weg befindet, so ist zu hoffen, wird in den hier vorgelegten Texten Antworten auf viele Fragen und zahlreiche Anregungen finden. Für Rückmeldungen und Vorschläge für künftige Darstellungen werden der Verlag, der Herausgeber und die Autorinnen der Einzelbeiträge jederzeit dankbar sein.

Berlin, im Herbst 2022

Thomas Merten

Teil I

Einführung in die Beschwerdvalidierung

Kapitel 1

Beschwerdenvalidierung in der psychologischen Begutachtung – Einführung

Thomas Merten

1.1 Hintergrund und Ziel der vorliegenden Publikation

In den vergangenen zehn Jahren hat es auch im deutschen Sprachraum eine Anzahl von Veröffentlichungen zum Thema der Beschwerdenvalidierung gegeben, sodass es kaum möglich ist, ohne erhebliche Redundanzen hier eine thematische Einführung zu liefern. Für eingehende überblicksmäßige Darstellungen kann beispielhaft auf die jüngeren Arbeiten von Keppler et al. (2017), Merten (2019, 2020) oder Stevens und Pfeiffer (2020) verwiesen werden. Ohne dass hier alle in den genannten Arbeiten besprochenen Einzelheiten wiederholt werden sollen, erscheint dennoch zur Einführung in diesen Sammelband eine einleitende Darstellung sinnvoll, die das Verständnis der folgenden Kapitel auch ohne Konsultation weiterer Quellen und ohne einschlägige Vorkenntnisse in wesentlichen Zügen gewährleisten soll.

Nachdem schon früh von prominenter Seite eine deutschsprachige Einführung in die Technik der Beschwerdenvalidierung (die zu jener Zeit in den USA und Kanada mit der Stärkung der Rolle neuropsychologischer Sachverständiger für Fragen des Verhältnisses zwischen Gehirn/Hirnverletzungen und Funktionen/Funktionsstörungen ein zunehmendes Interesse erfuhr) veröffentlicht wurde (Pankratz & Paar, 1988), war für die weiteren Entwicklungen zunächst nicht nur in Deutschland, sondern in ganz Europa eine große Zurückhaltung zu bemerken. Damit war spätestens zum Ende der 1990er-Jahre eine Entwicklungsverzögerung von etwa zehn Jahren zu konstatieren. Bedeutsame Forschungsaktivitäten auf diesem Gebiet ließen sich ohne große Mühe überschauen. Für Deutschland war es dann eine Art Meilenstein, als im Jahre 2007 ein Themenheft der *Praxis der Rechtspsychologie* eingeladen, konzipiert und realisiert wurde, das, der damaligen Diktion entsprechend, thematisch noch als *Simulationsdiagnostik* betitelt wurde. Dieses The-

menheft, rasch vergriffen, war erkennbar ein ebenso großer Erfolg wie der darauf basierende erste, um wichtige Themen erweiterte deutschsprachige Sammelband zur Beschwerdenuvalidierung (Merten & Dettenborn, 2009). In beiden Bänden wurden erstmals im deutschsprachigen Raum wichtige Themen publiziert, wie etwa eine eingehende Analyse von Verfahren zur Aufdeckung manipulativer Antwortstile aus bioethischer Perspektive (Bush, 2009) oder eine juristische Diskussion von Simulation und Aggravation in der Begutachtung (Brockmeyer, 2009). An beiden Bänden beteiligte sich einige Autoren, die auch für eine Mitarbeit am vorliegenden Sammelband gewonnen werden konnten.

Zu erwähnen ist schließlich die Reihe *Fortschritte der Neuropsychologie*, in der als Band 14 eine überschaubare Monografie zur Beschwerdenuvalidierung erschien (Merten, 2014), welche den Sammelband von 2009 ergänzte. Die jetzt vorgelegte Publikation soll an diese Monografie anknüpfen und sie keineswegs ersetzen. Erneut in Form eines Sammelbandes, wie bereits 2009, sollen vielfältige Facetten der Thematik der Beschwerdenuvalidierung dargestellt und diskutiert werden, und zwar aus der Perspektive von Autoren mit unterschiedlicher vertiefender Spezialisierung. Dies lässt erwarten, dass nicht nur verschiedene Blickwinkel zur Sprache kommen, sondern viele Einzelthemen vertiefender und kompetenter dargestellt und diskutiert werden können, als dies durch einen Einzelautor hätte realisiert werden können.

Zudem sollte nach der ursprünglichen Planung des Bandes vertiefend insbesondere auf die fortgesetzte Debatte um den Einsatz und die Bedeutsamkeit von standardisierten Verfahren zur Beschwerdenuvalidierung in der psychiatrischen Begutachtung eingegangen werden, eine Debatte, die noch lange nicht abgeschlossen ist und die weiterhin immer wieder in realen Gutachtenfällen und gehäuft auch vor Gericht ausgetragen wird. Dies führt zu dem letztlich unbefriedigenden Ergebnis, dass Richter über den diagnostischen Wert von Instrumenten und methodischen Ansätzen und die Überzeugungskraft ihnen vorgetragener Argumente entscheiden, und zwar nicht im Grundsatz, sondern jeweils fallbezogen, mit Bezug auf die (mehr oder weniger profunde und sachkundige) Argumentation eines jeweiligen Sachverständigen. Eine große Heterogenität veröffentlichter Urteile aus Deutschland und Schweiz gibt ein beredtes Zeugnis darüber ab. Zum großen Bedauern des Herausgebers konnten Autoren für eine Gegenposition, die also die Sinnhaftigkeit des Einsatzes von evidenzbasierten Verfahren zur Beschwerdenuvalidierung bei Patienten und Probandinnen mit psychischen Störungen anführt und die in einem Teil der psychiatrischen Literatur immer wieder und mit wenig Variation vorgetragen wird, nicht gewonnen werden.

Um Missverständnissen vorzubeugen, bleibt noch einmal zu betonen, dass Redundanzen insbesondere zu den oben erwähnten Büchern von 2009 und 2014 nicht gänzlich zu vermeiden waren, sie aber auf ein unbedingt notwendiges Maß zu reduzieren versucht wurde. Auch dieses Einleitungskapitel wird keine umfassende

Darstellung im Sinne eines Überblicksartikels liefern, sondern soll nur soviel Grundlagen vermitteln, als für das Verständnis der folgenden Texte unabdingbar erschien. Der deutschsprachigen Leserin, die nicht oder nur eingeschränkt auf die vielfältigen im englischsprachigen Raum vorliegenden Quellen zurückgreifen kann oder möchte, wäre aus der Sicht des Herausgebers zu empfehlen, zunächst den überschaubaren und in sich geschlossenen Text von Merten (2014) zu lesen, um anschließend bei Bedarf zur Aktualisierung oder Vertiefung die entsprechenden Kapitel dieses Bandes heranzuziehen.

Auch die Form der Darstellung der wichtigsten im deutschsprachigen Raum verfügbaren Verfahren zur Beschwerdenvalidierung folgte dem Ansatz, eine unnötige Redundanz zu vermeiden. Für vertiefende Testbesprechungen soll ausdrücklich auf das Testkompendium von Schellig, Heinemann, Schächtele und Sturm (2019) verwiesen werden; in Abgrenzung zu dort publizierten ausführlichen Rezensionen wurden jetzt nur kurzgefasste Testinformationen erbeten, die dem Leser eine rasche Orientierung über gebräuchliche Verfahren gestatten sollen. Aus diesem Grund und in Abgrenzung zu den erwähnten ausführlichen Besprechungen wurden sie, wo dies möglich war, bevorzugt von den Testautoren erbeten.

1.2 Verwendete Grundbegriffe

Für das Verständnis der folgenden Texte unumgänglich ist eine kurzgefasste Erläuterung von Begriffen, wie sie (möglichst einheitlich) in diesem Band verwendet werden sollen.

Unter dem für diese Publikation zentralen Begriff der *Beschwerdenvalidierung* soll in Anlehnung an die im renommierten Dorsch-Wörterbuch der Psychologie (Wirtz, 2020) gelieferte Definition verstanden werden:

Die Glaubhaftigkeit oder der Grad an Vertrauen, den ein Untersucher der Aufrichtigkeit einer Beschwerdenschilderung und der Gültigkeit der Symptombildung durch eine untersuchte Person im klinischen oder gutachtlichen Kontext entgegenbringen kann. In psychologischen Testuntersuchungen wird damit die Gültigkeit des ermittelten Testprofils in dem Sinne verstanden, als dieses die tatsächlichen Leistungsvoraussetzungen der Testperson adäquat widerspiegelt oder etwa als Resultat einer suboptimalen Testmotivation als ungültig zurückzuweisen ist. Der Prozess, der zu Aussagen über die Beschwerdenvalidität führt, wird als *Beschwerdenvalidierung* bezeichnet.

Vorschlägen von Widder (vgl. dessen Beitrag, Kapitel 2, in diesem Band) folgend, wurde in den letzten Jahren im deutschen Sprachraum eine Begriffsausweitung gebräuchlich. Danach kann der Begriff der Beschwerdenvalidierung im weiteren Sinne auf alle medizinischen und psychologischen Untersuchungsansätze angewandt werden, die eine Prüfung der Authentizität geltend gemachter Beschwer-

den und in einer Untersuchung dargebotener Symptome zum Ziel haben. Damit diene „letztlich jede klinische Untersuchung der Beschwerdenuvalidierung ... – sei es ‚lediglich‘ zur Abgrenzung organischer von nicht organischen Befunden ..., sei es im Rahmen der gutachtlichen Untersuchung zur Abgrenzung einer Vortäuschung oder Aggravation von Beschwerden“ (Widder, 2017, S. 741).

Unter dem Begriff der *Beschwerdenuvalidierungstests* (BVT) lassen sich alle standardisierten Verfahren zusammenfassen, die der Beschwerdenuvalidierung dienen, die also Aussagen zur Authentizität dargestellter Symptome sowie zur Glaubhaftigkeit subjektiv gelieferter Angaben mit Bezug auf Gesundheitsstörungen und ihrer Auswirkungen liefern. In der Testdiagnostik werden Aussagen zur Gültigkeit (Validität) der ermittelten Testergebnisse und – breiter gefasst – damit letztlich auch der ermittelten individuellen Testprofile insgesamt erhalten.

Ursprünglich, bis Ende der 1980er-Jahre, war der Begriff zur Bezeichnung einzel-fallexperimenteller Anordnungen zur Prüfung vorgetäuschter Symptomdarstellungen entwickelt worden. Solche Einzelfallexperimente stellen nach wie vor eine sehr potente Methode für die Untersuchung der Gültigkeit bestimmter Symptomkonstellationen dar und sind insbesondere dort einsetzbar, wo standardisierte Verfahren versagen müssen. Doch scheint ihre Verbreitung in der klinischen und forensischen Praxis bedauerlicherweise sehr limitiert zu sein (vgl. aber, für einen neueren Fallbericht, Jelcic, 2018). Eine ausführliche Erörterung soll mit Bezug auf vorangegangene Darstellungen, so auch bei Merten (2014), nicht erneut erfolgen (vgl. Denney, 1996; Giger & Merten, 2009; Merten & Merckelbach, 2013a), jedoch wird in einem anwendungsbezogenen Kapitel (Jelcic & Giger, Kapitel 25, in diesem Band) auf diese Technik im Rahmen geltend gemachter tatbezogener Amnesien eingegangen werden.

Die BVT werden heute üblicherweise in zwei Hauptgruppen eingeteilt, nämlich die *kognitiven Beschwerdenuvalidierungstests* (seit Larrabee, 2012, im Englischen *performance validity tests*, PVT, genannt) und die *fragebogenbasierten Beschwerdenuvalidierungstests* (im Englischen seit 2012 nunmehr meist allein als *symptom validity tests*, SVT, bezeichnet). Vgl. zur Einteilung der Methoden Kapitel 3 (in diesem Band) und insbesondere die Systematik in Abbildung 3.1.

Kognitive BVT sind Verfahren, die dem Anschein nach kognitive Funktionen erfassen, in Wahrheit jedoch dazu konstruiert wurden, die Validität von individuellen Ergebnissen in kognitiven Leistungstests zu prüfen. Sie können wiederum in eigenständige BVT (*stand-alone* oder besser *freestanding PVTs*) und eingebettete kognitive BVT (*embedded PVTs*) unterschieden werden. Bei Letzteren handelt es sich um Untertests oder spezifisch berechnete Kennwerte oder Indizes, die in neuropsychologische Verfahren oder kognitive Leistungstests integriert wurden, nicht jedoch, um, wie letztere, Leistungen zu erfassen, sondern gleichzeitig und ohne oder mit nur geringem zusätzlichen Untersuchungsaufwand die Validität genau dieser ermittelten Leistungen zu beurteilen.

Auch die fragebogenbasierten BVT lassen sich in eigenständige Verfahren und solche Validitätsskalen unterscheiden, die in komplexere Fragebögen eingebettet sind, wie beispielsweise in Persönlichkeitstests, Beschwerdenerfassungsbögen oder Symptomskalen.

Da kognitive BVT auf eine Prüfung der Validität von Testprofilen ausgerichtet sind, gestatten sie die Identifizierung eines möglichen *suboptimalen Leistungsverhaltens*. Von einem solchen wird dann gesprochen werden, wenn eine Person in Leistungstests aufgrund einer eingeschränkten *Anstrengungsbereitschaft* (Testmotivation, engl. *test effort*) deutlich unterhalb dessen abschneidet, was ihr bei optimaler Motivationslage möglich gewesen wäre (im Englischen wird seit einiger Zeit hier auch häufiger von *underperformance* gesprochen). Im Gegensatz dazu sind fragebogenbasierte BVT auf die Erfassung der Zuverlässigkeit oder Glaubhaftigkeit der Beschwerdenschilderung ausgerichtet. Mit ihrer Hilfe sind ausgeweitete, überhöhte, verschlimmernde Darstellungsformen (engl. *overreporting*) identifizierbar, je nach Verfahren z. T. aber nur ungültige Antwortstile insgesamt (also nicht auf überhöhte Darstellungen beschränkt).

Beide Formen ungültigen Antwortverhaltens im Sinne einer ausgeweiteten, verschlimmernden, überhöhenden Darstellung von Beschwerden und Funktionseinschränkungen können auch unter dem Begriff der *negativen Antwortverzerrungen* zusammengefasst werden, der sich zum Teil großer Beliebtheit erfreut. In Anlehnung an Bush, Ruff et al. (2006) ist hiermit der Sachverhalt gemeint, dass ein Proband versucht, den Untersucher durch ungenaue oder unvollständige Antworten oder durch eine unzureichende Leistungsanstrengung zu täuschen. Im Gegensatz dazu treten *positive Antwortverzerrungen* im Rahmen einer Bagatellisierung, Verniedlichung, Verleugnung von Beschwerden oder Dissimulation auf. *Antwortverzerrungen* beziehen sich also generell auf eine kontaminierte Datenlage, in der Testwerte oder Aussagen untersuchter Personen ein zugrundeliegendes Merkmal aufgrund von Störeinflüssen nicht zutreffend beschreiben.

Weitere Begriffe wie Simulation, Aggravation u.Ä. werden weiter unten, im Abschnitt 1.5 zum konzeptionellen Wandel, erläutert.

1.3 Konsistenz und Plausibilität

Kernstück einer jeden Begutachtung muss eine eingehende Prüfung der Informationsbasis auf Konsistenz und Plausibilität sein. Normalerweise wäre diese Anforderung natürlich für jeden klinischen Fall und für jeden Rehabilitationspatienten ebenso zu formulieren, denn eine korrekte Diagnosestellung hängt nicht unwesentlich von einer sorgfältigen Prüfung der Datenlage auf ihre logische Kohärenz und Vertrauenswürdigkeit ab. Wo eine solche kompromittiert ist, muss auch die Zuverlässigkeit der gezogenen Schlüsse bereits auf diagnostischer Ebene infrage

gestellt werden, von therapeutischen und anderen Konsequenzen ganz abgesehen. Dennoch steht eine Prüfung von Konsistenz und Plausibilität im Qualitätsmanagement von Klinik und Rehabilitation nicht an favorisierter Stelle. Wohl aber ist dies im Rahmen der rechtlich besonders relevanten Feststellungen von Sachverständigengutachten der Fall.

Die psychologische Beschwerdenvalidierung im engeren Sinne, also mit Rückgriff auf standardisierte Verfahren zur Validitätsprüfung, ordnet sich letztlich in eine solche umfassende Konsistenz- und Plausibilitätsprüfung ein. Sie stellen dabei ein spezifisches Bestimmungsstück mit spezifischen Eigenschaften dar, die es von anderen unterscheiden kann (insbesondere etwa bzgl. ihrer Standardisierung, von Entscheidungs- und Interpretationsregeln, ihrer Evidenzbasierung oder bekannten Angaben zur Sensitivität, Spezifität oder zum Vorhersagewert individueller Klassifikationsresultate).

Eine Konsistenz- und Plausibilitätsprüfung geht naturgemäß weit über die Anwendung standardisierter Testverfahren hinaus und umfasst den Abgleich aller in einem konkreten Einzelfall zur Verfügung stehenden Informationen (vgl. dazu im Detail Kapitel 3 im Methodenteil, in diesem Band, insbesondere Tabelle 3.1). Der Begriff der *Konsistenz* bezieht sich dabei zunächst auf eine logische Widerspruchsfreiheit der verfügbaren Information, während unter *Plausibilität* gefasst wird, ob eine Argumentation einleuchtend, verständlich und überzeugend erscheint.

Beide Begriffe lassen sich noch weiter untersetzen (Abbildung 1.1). Für eine eingehende Diskussion dazu ist auf Merten und Dohrenbusch (2012) zu verweisen. In der Regel wird in der gutachtlichen Literatur zusammenfassend von einer Prüfung auf Konsistenz und Plausibilität gesprochen.

Der Begriff der *Konsistenz* wurde aus dem Latein entlehnt (*consistens* bzw. *consistentia*) und meint zum einen *Festigkeit, Dauerhaftigkeit* (hier den Begriff der *Konstanz* berührend); für den diagnostischen Urteilsprozess zentral ist die weitere Bedeutung als *Widerspruchsfreiheit* zwischen einzelnen Daten (oder, logisch gesprochen: zwischen den Argumenten). Mit dem Webster's (Gove et al., 1993) sind Daten dann logisch konsistent, wenn die folgende Eigenschaft auf sie zutrifft: „*jointly assertable so as to be true or not contradictory*“ (S. 484) – die Daten in ihrer Gemeinsamkeit können als wahr und widerspruchsfrei akzeptiert werden.

Im Gegensatz dazu bezieht sich *Plausibilität* auf die Dimension der Überzeugungskraft, die eine Argumentation entfalten kann. Das Wort wurde im 17. Jahrhundert aus dem Französischen übernommen und bedeutet im Ursprung „gefallend, von den Leuten gebilligt“, auf das lateinische *plausibilis* (Beifall verdienen, auf Beifall berechnet) zurückgehend (Autorenkollektiv, 1989). Plausibel bedeutet keineswegs „wahr, zutreffend, korrekt“. Eine plausible (oder plausibel erscheinende Erklärung) kann grundfalsch sein.

In diesem Sinne und mit dem Anschein einer gewissen Paradoxie können sowohl Plausibilität als auch Konsistenz, wiewohl beide Kriterien von zentraler und herausragender Bedeutung für die Informationsbewertung sind, nur von untergeordnetem Range sein, also Hilfskriterien. Dies begründet sich daraus, dass sie nur dann greifen, wenn stärkere Beweismittel nicht zur Verfügung stehen. Eine diagnostische (gutachtliche) Urteilsbildung, die eine Gesundheitsstörung nur deshalb als nachgewiesen ausweist, weil sich die Informationsbasis als plausibel und konsistent darstellt, kann von ihrer Wertigkeit her wohl in etwa mit einem Indizienbeweis verglichen werden.

Warum nicht nur im Gutachtenkontext, sondern natürlich auch im realen Leben Experten und Nichtexperten häufig allzu leicht eklatantesten Fantasiegeschichten aufsitzen, wird durch den *Truth bias* erklärt, einer fehleraffinen Neigung zur Gutgläubigkeit. Zwei prominente Beispiele dafür, unter vielen anderen, sind der Fall Wilkomirski, eines gefeierten und geehrten Überlebenden zweier Konzentrationslager, der in Wahrheit nie in einem Lager inhaftiert war (z. B. Merckelbach, 2002), und der ebenfalls lange Zeit unaufgedeckt gebliebene Fall einer Überlebenden des Anschlags vom 11. September 2001, die gar nicht am Ort des Geschehens war ... (Fisher & Guglielmo, 2012). Angesichts solcher prominenten Geschichten möchte man fast geneigt sein, mit den Römern zu sagen: *Mundus vult decipi, ergo decipiatur* (die Welt will betrogen werden, also soll sie betrogen werden).

Anekdotisch interessant im Kontext von Begutachtung und Beschwerdenvvalidierung erscheint schließlich ein Wort Goethes, das auch Aufnahme in das Grimmsche Deutsche Wörterbuch fand: „Wer die Menschen betrügen will, muß vor allen Dingen das Absurde plausibel machen“ (Goethe, 1972, S. 667).

Eine der größten Schwächen der Konsistenz- und Plausibilitätsprüfung ist ihre fehlende Standardisierung. Es gibt kein ausgearbeitetes und allgemein akzeptiertes

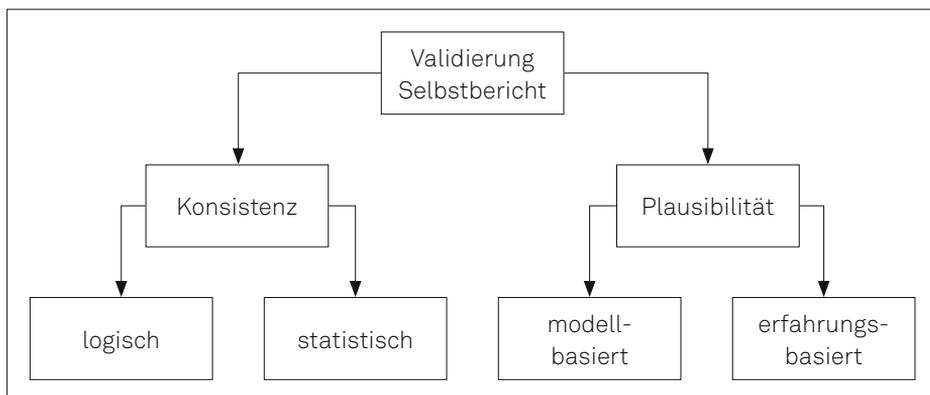


Abbildung 1.1: Taxonomie zur Validierung interview- und fragebogenbasierter Angaben, hier spezifisch auf Selbstberichts-Daten bezogen (Merten & Dohrenbusch, 2012)

formales Regelwerk, das zur Bewertung einer solchen Prüfung angelegt werden könnte. Sie erfolgt, falls sie überhaupt erkennbar vorgenommen wird, vielmehr individuell, unsystematisch, uneinheitlich, häufig nicht logisch stringent, also unter Missachtung der Denkgesetze, in vielen Fällen willkürlich und nicht ausreichend nachvollziehbar. Zuvor vom Sachverständigen selbst erwähnte widersprüchliche Informationen werden häufig unbeachtet gelassen. Nicht nur die Methodik der Prüfung, sondern auch die Folgen, die sich aus Verstößen ergeben, werden in der tatsächlichen Gutachtenpraxis mit erkennbarer Beliebigkeit gehandhabt. Hier ist ein großer Entwicklungsbedarf für eine unbestimmte Zukunft zu erkennen.

Die Aussagefähigkeit oder der Beweiswert einer Plausibilitäts- und Konsistenzprüfung wird zudem von Gutachtern häufig unberücksichtigt gelassen. Dabei gilt es zum einen zu beachten (und dies sollte in geeigneter Form auch dem Empfänger eines Gutachtens mitgeteilt werden), dass einer *bestätigten* Konsistenz und Plausibilität der Datenlage keineswegs der Rang eines objektiven Wahrheitskriteriums zukommt (also zum Beispiel eine Gesundheitsstörung nicht bereits deshalb als bewiesen gelten kann und darf, weil sich die dem Gutachter zur Verfügung stehende Informationsbasis als einleuchtend und in sich widerspruchsfrei darstellt). Zum anderen wird häufig nicht ausreichend klar und überzeugend herausgearbeitet, welche Konsequenzen es für die Überzeugungsbildung haben muss, wenn bedeutsame Inkonsistenzen und Unplausibilitäten nicht in zufriedenstellender, überzeugender Weise aufgelöst werden können.

Eine abschließende Bemerkung zu diesem Abschnitt gilt fremdanamnestischen Angaben. Diese sind, wenn sie bereits vorliegen oder wenn sie, sofern rechtlich zulässig, im Rahmen einer gutachtlichen oder klinischen Untersuchung erhoben werden, keineswegs stillschweigend und unbesehen als wahr und valide anzusehen, insbesondere dann nicht, wenn sie von *interessierten* Dritten geliefert werden. Das sind Personen, die selbst ein mittelbares oder unmittelbares Interesse etwa am Ausgang eines Rechtsstreits haben (z. B. in Form eines tertiären [Krankheits-]Gewinns). Dann können deren Angaben ähnlichen intentionalen Verzerrungen unterliegen, wie dies von Selbstaussagen einer untersuchten Person bekannt ist. Diese Problematik hat bislang in der Literatur kaum Beachtung gefunden; sehr interessant ist eine neue empirische Untersuchung von Webber, Sullivan-Baca, Modiano, Taiwo und Grabyan (2022), auf die für Weiteres verwiesen werden kann.

1.4 Ziele und Kontexte der Beschwerdenuvalidierung

Übergeordnetes Ziel einer eingehenden Beschwerdenuvalidierung ist eine Differenzierung zwischen vertrauenswürdiger und nicht oder weniger vertrauenswürdiger Information zu einem konkreten klinischen oder gutachtlichen Fall. Damit

werden Aussagen zur Authentizität der Symptompräsentation und zur Glaubhaftigkeit geltend gemachter Beschwerden getroffen. Gleichzeitig geben die Ergebnisse einer eingehenden und qualifizierten Beschwerdenvalidierung Auskunft über den Grad an Vertrauen, den der Diagnostiker der Diagnosestellung und der Beurteilung möglicher Funktionseinschränkungen entgegenbringen kann.

Ursprünglich war der Begriff der Validität in der Testtheorie als Eigenschaft von psychologischen Tests definiert (z.B. Borsboom, Mellenbergh & van Heerden, 2004; Cronbach & Meehl, 1955); er wurde dann ausgeweitet auf andere Methoden und Verfahren der empirischen Datengewinnung angewandt. Mittlerweile kann von einem fachgeschichtlichen Wandel gesprochen werden, den der Begriff erfahren hat, und in diesem Sinne kann Validität heute als Qualitätskriterium im Sinne der „Zuverlässigkeit von Testinterpretationen“ aufgefasst werden (Hartig, Frey & Jude, 2020).

Aus den Ergebnissen der Beschwerdenvalidierung ergibt sich nicht allein Information zur Zuverlässigkeit getroffener diagnostischer Aussagen, sondern darüber hinaus beispielsweise auch zur Sicherheit, mit der prognostische Vorhersagen getroffen werden können, und zu den Erwartungen, die einer Behandlungscompliance und dem Behandlungserfolg realistischere entgegengebracht werden sollten. In diesem Gesamtkontext können etwa die einschlägigen Studien zu Erwartungen sekundären Störungsgewinns bei psychiatrischen Patientinnen (van Egmond & Kummeling, 2002; van Egmond, Kummeling & Balkom, 2005) sowie Untersuchungen bei psychosomatischen Rehabilitationspatienten (Kaminski, Merten & Kobelt-Pönicke, 2020; Merten, Kaminski & Pfeiffer, 2020; Pfeiffer, Gutermann & Bassler, 2017) aufschlussreich sein. Die Tatsache, dass es sich um einen klinischen Patienten handelt und nicht um eine Gutachtenprobandin, ist allein keineswegs eine Garantie dafür, dass seine Beschwerdendarstellung glaubhaft und valide erfolgt; unter bestimmten Bedingungen ist vielmehr auch bei klinischen Patienten mit deutlichen Prävalenzraten nicht glaubhaften Darstellungen zu rechnen (z. B. zuletzt Schroeder, Clark & Martin, 2022).

In Tabelle 1.1 sind die unterschiedlichen Ziele einer Beschwerdenvalidierung in verschiedenen Kontexten dargestellt. Ursprünglich vor allem für gutachtliche Fragestellungen herausgearbeitet, hat sich die internationale Forschung im letzten Jahrzehnt zunehmend Fragen im klinischen und rehabilitativen Kontext zugewandt, darunter auch der wichtigen Frage, wie mit auffälligen Ergebnissen bei klinischen Patienten umzugehen ist (vgl. Carone & Bush, 2018; Dandachi-FitzGerald & Martin, 2022; Martin & Schroeder, 2021). Für bestimmte korrekte Diagnosestellungen erscheint eine eingehende Beschwerdenvalidierung in jedem Untersuchungskontext unabdingbar zu sein. Dies ist insbesondere für die artifiziellen Störungen (engl. *factitious disorder*) der Fall (z.B. Bass & Halligan, 2014; Merten & Merckelbach, 2020), ebenso aber auch (unter weiteren problematischen Diagnosen) für Konversionsstörungen, dissoziative und somatoforme Störungen (Mer-

ten & Merckelbach, 2013b). Diese teilen mit vorgetäuschten oder grob überhöht dargestellten Gesundheitsstörungen die auffallende Diskrepanz zwischen subjektiven Geltungsmachungen und objektiven Befunden und sind phänomenologisch nicht oder kaum voneinander unterscheidbar.

Allgemein gesprochen sollte eine Beschwerdvalidierung entsprechend moderner konzeptioneller Entwicklungen keineswegs mehr mit einer *Simulationsdiagnostik* gleichgesetzt werden, wie dies etwa vor 20 Jahren fast routinemäßig erfolgte. Wie durch Merten und Merckelbach (2013b) eingeschätzt wurde, hat sich unter den meisten Experten auf der Welt vor etwa zehn Jahren ein weitgehender Konsens dahingehend eingestellt, dass

1. Beschwerdvalidierungstests wichtig dafür sind, Fälle unklarer Beschwerdekongstellationen aufzuklären;
2. zur Beschwerdvalidierung sowohl Selbstberichtsverfahren (Fragebögen) zur Identifizierung einer ungültigen, ausgeweiteten, überhöhenden Beschwerdeshildung als auch kognitive BVT herangezogen werden, die Aussagen zur Validität ermittelter Testprofile gestatten;
3. eine überhöhte, nicht glaubhafte Beschwerdeshildung und eine unzureichend entfaltete Testmotivation in der Untersuchung als zwei Erscheinungsformen negativer Antwortverzerrungen zutage treten, die teils gemeinsam auftreten, teils aber auch nicht;
4. Simulation und Aggravation lediglich zwei Erscheinungsformen ungültigen Antwortverhaltens durch Probanden darstellen, unter einer Reihe weiterer Kontextbedingungen (siehe Abschnitt 1.5).

Die Ergebnisse einer sorgfältig und der jeweiligen Fallkonstellation angemessen durchgeführten Beschwerdvalidierung liefern also Aussagen über den Grad an Vertrauen, das den ermittelten Befunden und Daten entgegengebracht werden kann.

Tabelle 1.1: Zielrichtung und Aussagefähigkeit von Ergebnissen einer eingehenden Beschwerdvalidierung in unterschiedlichen Untersuchungskontexten

	Klinische Kontexte	Rehabilitation	Begutachtung
Beurteilung der Glaubhaftigkeit oder Vertrauenswürdigkeit der erfassten Befunde	X	X	X
Korrekte Diagnosestellung/ Zuverlässigkeit der Diagnosestellung	X	X	X
Beurteilung von Behandlungsbedarf und Behandlungs-Compliance	X	X	(X)
Verbesserung der Prognosesicherheit	X	X	(X)

Tabelle 1.1: Fortsetzung

	Klinische Kontexte	Rehabilitation	Begutachtung
Beeinflussung möglicher Probleme bei der Mitwirkung (Intervention)	X	X	
Sozialmedizinische Beurteilung		X	X
Befähigung des Auftraggebers (Gericht, [Sozial-]Versicherung, Behörde) zu einer sachlich korrekten Entscheidung		(X)	X

1.5 Konzeptioneller Wandel, Bedingungen und Kontexte auffälliger Ergebnisse, Simulation und Aggravation

Die Gleichsetzung der Beschwerdenvalidierung mit einer *Simulationsdiagnostik* kann nur noch fach- und begriffsgeschichtlich von Interesse sein und muss spätestens mit Beginn der 2010er-Jahre als überholt eingestuft werden (vgl. dazu auch jüngst Rogers, 2021). Die ursprüngliche Ausrichtung von BVT, Aggravation und Simulation zu erfassen (d.h. also die Auffassung: BVT=Simulationstest) ist damit ebenso als desolat zu begreifen. Testnamen wie *Test of Memory Malingering*, *Aggravations- und Simulationstest* oder *Structured Inventory of Malingered Symptomatology* zeugen noch davon; inhaltlich lohnt sich etwa die Lektüre der Interpretationsleitlinien des letztgenannten Verfahrens (Widows & Smith, 2005), für den Fall, dass eine interessierte Leserin dieser früheren Konzeptionalisierung nachspüren möchte.

Der internationale Trend in der Bezeichnung auffälliger Ergebnisse in der Beschwerdenvalidierung scheint dahin zu gehen, die betroffenen Daten als ungültig (engl. *invalid, nonvalid*) oder – zunehmend – als nicht glaubhaft (engl. *noncredible*) zu beschreiben, und zwar ungeachtet des Ursachen- oder Bedingungsgefüges für ein solches Resultat. Ein Punktwert in einem Validitätstest kann isoliert, außerhalb der Gesamtbewertung der Ergebnisse, als *im gültigen Bereich, unbestimmt* oder *im ungültigen Bereich* beschrieben werden (engl. *within the valid range, indeterminate range, within the invalid range*), so ein aktueller Vorschlag von Guilmette et al. (2020).

Dem aufmerksamen Leser wird nicht entgangen sein, dass in Tabelle 1.1 als Stichworte weder Simulation noch Aggravation auftauchen; sie erscheinen dort auch vollständig verzichtbar. Falls in einem Einzelfall bedeutsame simulative Tendenzen zu vermuten oder nachzuweisen sind, so wird dies jedoch zwangsläufig die

Diagnosestellung, die Diagnosesicherheit und, im gutachtlichen Kontext, die Beantwortung gestellter (Beweis-)Fragen beeinflussen.

Das Bedingungsgefüge, in dem eine eingehende Beschwerdenuvalidierung auffällig ausfallen kann und auf das Vorliegen ungültiger oder nicht ausreichend vertrauenswürdiger Befunde zu schließen ist, findet sich in Abbildung 1.2 illustriert. Besondere Sorgfalt ist dahingehend geboten, einzelne Bestimmungsstücke in diesem Bedingungsgefüge nicht stillschweigend und unkritisch im Sinne von Kausalitäten oder Wirkursachen aufzufassen. So muss etwa eine Persönlichkeitsstörung keineswegs kausal dafür sein (im Sinne einer *conditio sine qua non*, einer *notwendigen* Bedingung), dass in einer Testuntersuchung eine so geringe Anstrengungsbereitschaft entfaltet wird und dass alle eingesetzten BVT korrekt positiv die Ungültigkeit des ermittelten Testprofils anzeigen. Ebenso kann von einem Patienten oder einer Probandin mit der mitgeteilten Diagnose einer somatoformen Störung in aller Regel zumutbar erwartet werden, dass er oder sie in einer psychologischen Untersuchung den individuellen Möglichkeiten entsprechend mitwirkt. Die rechtliche Würdigung einer fehlenden oder eingeschränkten Mitwirkung obliegt dem Entscheidungsträger (wie etwa dem Juristen). In der Verantwortung des Diagnostikers liegt es, eine begrenzte Mitwirkung, wo sie vorliegt, zielsicher und korrekt zu erkennen und dem Entscheidungsträger zu kommunizieren. Damit wird der Entscheidungsträger befähigt, seinerseits abzuleiten, wie sich eine zweifelhafte, ungültige oder unvollständige Befundlage auf die Beantwortung der Rechtsfragen auswirkt (Brockmeyer, 2009). Für eine neuere ausführliche Diskussion der juristischen Implikationen, die sich aus verzerrten oder verfälschten Probandenangaben ergeben, ist in besonderer Weise auf die Darstellung von Neuhaus (2021; vgl. auch die Synopsis bei Ostendorf, 2022) hinzuweisen, der die Problematik für geltend gemachte psychische Störungen im Rahmen privater Berufsunfähigkeitsversicherungen analysiert hat.

In diesem Zusammenhang ist es häufig (und möglicherweise sogar im Regelfall) rechtlich unerheblich, ob eine unzureichende Mitwirkung, wenn sie denn nachgewiesen werden kann und als Beweisführungshindernis in Erscheinung tritt, im Sinne einer Simulation oder Aggravation zu bewerten ist oder nicht. Daraus folgt, dass, anders als dies vor noch zehn Jahren meist als Empfehlung vorgetragen wurde, diese Begriffe häufig in einem Gutachten völlig entbehrlich sind und die Sachverhalte anders („sanfter“) kommuniziert werden können, ohne sie jedoch in der Sache irgend zu beschönigen, zu verniedlichen oder wegzudiskutieren.

Von dieser Regel gibt es natürlich Ausnahmen, etwa dann wenn eine gezielte Simulation von Gesundheitsstörungen strafrechtliche Konsequenzen haben kann (z. B. bei Versicherungsbetrug) oder wenn differenzialdiagnostisch eine artifizielle Störung zu untersuchen ist. Auch wenn in einer Beweisfrage direkt danach gefragt wird, ob bei der untersuchten Person eine Simulation oder Aggravation festgestellt wurde, sollte direkt und nicht verschleiern darauf geantwortet werden.

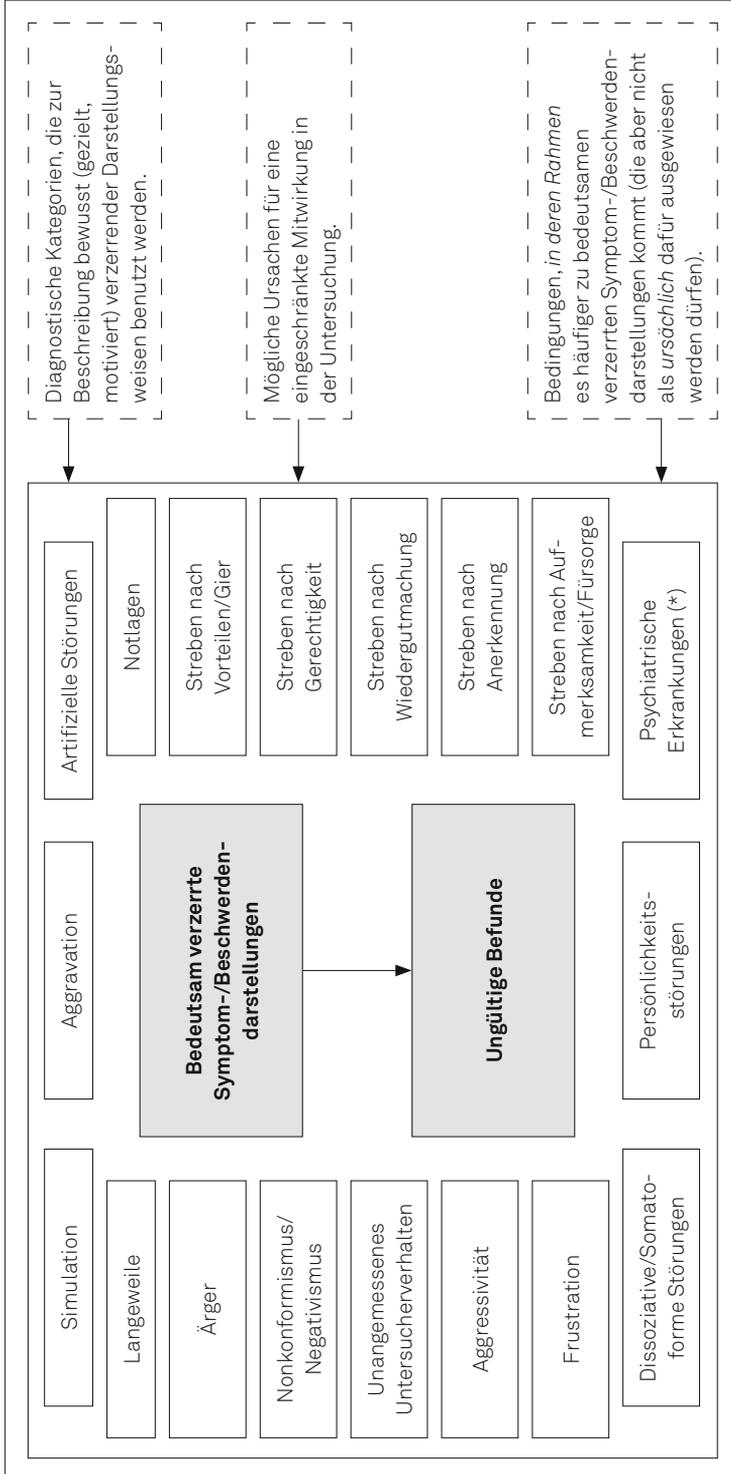


Abbildung 1.2: Bedingungsgefüge für das Auftreten auffälliger Ergebnisse in der Beschwerdvalidierung (korrekt-positive Klassifikation einer ungültigen Befundlage). Den psychiatrischen Erkrankungen (*) kommt in der Abbildung eine Sonderstellung zu, weil sie zum einen Kontextbedingung für das Auftreten von nicht-authentischen Darstellungsformen von Beschwerden und Symptomen sind, zum anderen in speziellen Fällen auch ursächlich für eine eingeschränkte Mitwirkung sein können. Eine etwaige ursächliche Zuschreibung stellt allerdings strenge Maßstäbe an die Beweisführung; die bloße Präsenz einer authentischen psychischen Störung ist hierfür in keiner Weise hinreichend. Aus: Merten (2020). © Deutscher Psychologischer Verlag, Abdruck mit freundlicher Genehmigung.

Die hier benutzten Begriffe sollen der Vollständigkeit halber und in Ergänzung zur Begriffsklärung von weiter oben noch erläutert werden:

Unter *Simulation* wird eine bewusste, gezielte Vortäuschung von Gesundheitsstörungen oder Funktionseinschränkungen verstanden, die einem externalen Ziel dient, auch sekundärer Krankheitsgewinn genannt. Dazu gehören finanzielle Vorteile, Arbeits- oder Kriegsdienstbefreiung, Medikamentenbeschaffung, strafrechtliche Konsequenzen (Schuld-, Straf- oder Verhandlungsunfähigkeit, Haftverschonung).

Aggravation meint hingegen die verschlimmernde Darstellung, Übertreibung, Überhöhung, Ausweitung oder zeitliche übertriebene Ausdehnung von Gesundheitsbeschwerden, d.h. ihr liegen im Kern authentische Beschwerden zugrunde, die aber ausgeweitet dargestellt werden, und zwar ebenfalls mit einem externalen Ziel.

Eine *minder schwere Verdeutlichung* gehört zur Aggravation und stellt nur ihre geringste Schweregradstufe dar. Im Englischen entspricht sie dem *Mild malingering* (Rogers, 2018). Eine solche geringgradige Aggravation sollte nur dann klassifiziert werden, wenn die Befunderhebung durch sie nicht substantiell erschwert wird. Standardisierte Verfahren zur Beschwerdenuvalidierung sollten im Regelfall noch unauffällig ausfallen. Für eine weiterreichende Diskussion zu dieser leichten Aggravation und der häufig irreführenden, verschleiernenden Bagatellisierung höherer Schweregrade durch manche Gutachter vgl. Merten (2020).

Der englische Begriff des *Malingering* könnte mit Wurzer (1992) im Deutschen als „simulative Tendenzen“ übersetzt werden. Er umfasst Aggravation und Simulation gleichermaßen, schließt aber auch falsche Ursachenzuschreibungen (*false imputations*) mit ein. So kann beispielsweise gegenüber der Haftpflichtversicherung als Folge eines Schleudertraumas ein Kopfschmerz geltend gemacht werden, der in Wahrheit schon seit Kindestagen die verunfallte Autofahrerin plagt. Der Kopfschmerz muss nicht übertrieben dargestellt werden, aber Gutachter und Haftpflichtversicherung werden hinsichtlich seiner Vorgeschichte, seines Ursachen- und Bedingungsgefüges getäuscht.

Anders als häufig geglaubt, in Gutachten und gelegentlich auch in Veröffentlichungen dargestellt, bedingen Faktoren wie Übermüdung, eine bedeutsame depressive Symptomatik, Schmerzen, der Einfluss psychotroper Substanzen u. Ä. m. keineswegs die Präsenz ungültiger Befunde, unter der Voraussetzung, dass die untersuchte Person während einer Testuntersuchung mit der in dieser Situation möglichen Anstrengungsbereitschaft mitwirkt und offen und aufrichtig eine Beschwerdenschilderung liefert. Werden dann beispielsweise leichte Aufmerksamkeits- oder Kurzzeitgedächtnisstörungen ermittelt, so würden diese den in Wahrheit wirkenden Einfluss der genannten Faktoren auf die Leistungsfähigkeit (zum Zeitpunkt der Untersuchung) ausweisen. Unter den hier unterstellten Bedingungen einer vollen Mitwirkung müssten auch die Ergebnisse einer eingehenden Beschwerdenuvalidierung unauffällig ausfallen. Dies wäre nur dann nicht der Fall,

wenn die eingesetzten Instrumente nicht resistent gegenüber den genannten Faktoren wären – in diesem Falle könnten falsch-positive Ergebnisse zustande kommen. Um in einem solchen Falle aber zu dem Schluss zu gelangen, dass es sich tatsächlich um Falsch-positive handelt, muss überzeugend die Anfälligkeit der entsprechenden Validierungsverfahren gegenüber den infrage stehenden Störbedingungen aufgezeigt werden. Sie kann nicht einfach per Verdikt deklariert werden. Andernfalls gewinnt die Beurteilung der Gutachterin leicht den Charakter eines Weg-Diskutierens.

Eine gänzlich andere Frage ist es jedoch (und dies wird gelegentlich durch Sachverständige nicht sorgfältig unterschieden), ob ein ermitteltes Testprofil, welches unter dem Einfluss widriger Faktoren zustande gekommen ist, generalisierbar auf den im Gutachtenauftrag umrissenen Sachverhalt ist. Wenn also etwa eine eingeschränkte kognitive Leistungsfähigkeit unter dem Einfluss geltend gemachter *ständiger, ununterbrochen täglich auftretender* Kopfschmerzen untersucht werden soll, so wäre die angemessene Situationsbedingung während der neuropsychologischen Untersuchung, dass solche Kopfschmerzen auch für den Untersuchungszeitraum bejaht werden. Wenn der Proband üblicherweise wach und ausgeruht ist und dennoch Aufmerksamkeits- und Konzentrationsstörungen beklagt, so würde eine Untersuchung an einem Tag, an dem er starke Müdigkeit bei fehlendem Schlaf in der Nacht zuvor angibt, für die Fragestellung völlig unbrauchbare Ergebnisse liefern müssen.

Vor einer Verwendung des mancherorts beliebten *Hilferufs* (engl. *cry for help*) als Erklärung für eine manipulative und interessegeleitete Beschwerdendarstellung ist dringend abzuraten, wie bereits von Iverson (2006) überzeugend dargestellt wurde (vgl. auch Merten & Merkelbach, 2013b). Ein solches metaphorisches, emotional geladenes und pseudoakustisches Konzept ist nicht geeignet, die Aufklärung einer ungültigen, ungläubhaften Datenlage positiv zu befördern oder sie im Einzelfall gar ungeschehen zu machen. Wenn bei der Aufklärung oder Diskussion des Motivationsrahmens für eine eingeschränkte Mitwirkung Faktoren wie eine finanzielle oder psychosoziale Notlage oder das Streben nach Aufmerksamkeit oder Fürsorge oder Anerkennung plausibel herausgearbeitet werden können, sollte dies auch besser so und nicht unangemessen affektiv geladen zum Ausdruck gebracht werden (vgl. Abbildung 1.2; für eine weiterreichende Diskussion Dandachi-FitzGerald, Merkelbach & Merten, 2022).

1.6 Die Interpretation von Ergebnissen der Beschwerdenvvalidierung

Wie bereits oben dargestellt, ist die Gleichsetzung der Beschwerdenvvalidierung mit einer Simulationsdiagnostik und von Beschwerdenvvalidierungstests mit Simulationstests nur noch historisch von Bedeutung. Auffällige Ergebnisse in einem

BVT zeigen eine ungültige oder nicht glaubhafte Befundlage an, ohne dass aus einem Testergebnis selbst ableitbar ist, worauf diese zurückzuführen ist. Eigentlich psychologisches Allgemeingut, dennoch immer wieder vergessen oder in der praktischen Testanwendung ignoriert, ist an ein Wort von Vanderploeg (1994) zu erinnern, das selbstverständlich auch für BVT und ihre Einbettung in eine eingehende Beschwerdvalidierung gilt:

In and of itself, testing is not capable of answering questions and requires minimal clinical expertise other than the correct administration and scoring of test instruments. Testing is a tool that may be utilized during a neuropsychological assessment as one source of information (S. 2). – *An und für sich* kann man mit Tests (allein) keine Fragen beantworten.

Selbst für Antworten signifikant unterhalb der reinen Ratewahrscheinlichkeit in Zwangs- oder Mehrfachwahl-Verfahren (Unter-Zufall-Antworten), die traditionell als Alarmsignal (*red flag*) für eine Simulation von Funktionsstörungen behandelt wurden, gilt dies. Was solche Antwortmuster anzeigen, ist die *gezielte Intention* falschen Antwortens, also die absichtsvolle, bewusste Manipulation des Antwortverhaltens durch die Probandin (auch wenn dies immer noch in manchen Gutachten oder sogar in einzelnen veröffentlichten Arbeiten wegzudiskutieren versucht wird). Genau in diesem Sinne nannten Pankratz und Erickson (1990) solche Antwortmuster *“the smoking gun of intent”*. Ob der Kontext einer solchen gezielten Manipulation jedoch der einer Simulation, einer Aggravation, einer artifiziiellen Störung oder ein gänzlich anderer ist, lässt sich anhand des Testergebnisses allein nicht bestimmen. – Interessant ist in diesem Zusammenhang übrigens, dass nach neueren Untersuchungen zu vermuten ist, dass in der tatsächlichen Testpraxis Unter-Zufall-Antworten als Falsch-positive noch sehr viel seltener sein dürften, als dies die theoretischen Wahrscheinlichkeiten ausweisen (Olsen, Schroeder & Martin, 2019).

Für Einzelheiten speziell zur Interpretation kognitiver BVT kann im Übrigen auf die Darstellung bei Merten (2014) verwiesen werden.

1.6.1 Interpretation auffälliger Ergebnisse

Auffällige (positive) Ergebnisse in einem BVT zeigen die Ungültigkeit eines ermittelten Testprofils oder die Ungültigkeit bzw. Unzuverlässigkeit der subjektiven Beschwerdendarstellung durch die Probandin an. Eine weiterreichende Beurteilung könnte etwa im Sinne einer eingeschränkten Anstrengungsbereitschaft in kognitiven Tests, einer suboptimalen Testmotivation, einer fehlenden Glaubhaftigkeit geltend gemachter Beschwerden oder einer unzureichenden Mitwirkung in der Untersuchung lauten. Das Testergebnis selbst gestattet keine Aussage dazu, aufgrund welcher Faktoren ein solches suboptimales Leistungsverhalten in der Testuntersuchung oder eine unglaubliche Beschwerdendarstellung durch den Probanden ver-

wirklicht wurde. Das kann schlichte Langeweile eines Probanden gewesen sein oder sein Ärger über die Versicherungsgesellschaft, die als ihm gegenüber feindlich eingestellt wahrgenommen wird und für deren Begutachtung er sich in keiner Weise zu engagieren bereit zeigt, das kann eine finanzielle Notsituation sein, in deren Zusammenhang eine in Aussicht stehende (aber in der Sache ungerechtfertigte) Entschädigungszahlung nach einem Bagatellunfall gerade recht zu kommen erscheinen mag. In den ersten beiden Fällen wäre es verfehlt, auf eine Aggravation oder Simulation zu schließen, und dennoch bleibt die Bewertung im Sinne einer zweifelhaften oder fehlenden Validität der Befunde festzustellen. Die rechtlichen Konsequenzen, die sich daraus ergeben, können in bestimmten Kontexten exakt die gleichen sein, unabhängig davon, ob eine eingeschränkte Mitwirkung gezielt manipulativ (*interessegeleitet*) oder aus lediglich *fehlendem Interesse* heraus erfolgte. Wenn die Probandin die objektive Beweislast trägt (und damit eine nicht gelungene Beweisführung juristisch zu ihren Ungunsten zu werten ist), kann es völlig unerheblich sein, warum exakt der Nachweis nicht gelingt. In einem solchen Kontext könnte also eine unzureichende Mitwirkung, eine unzuverlässige oder zweifelhafte Befundlage, gleich welcher Quelle, zu einem *Beweisführungshindernis* werden.

Dies gilt potenziell auch für Fälle von Aggravation, in denen ein realer Kern von Symptomen oder Beschwerden – im Gegensatz zur Simulation – unterstellt wird. Gestatten Art und Ausmaß an manipulativ verzerrten Daten keine ausreichend gesicherten Aussagen zu Art und Ausmaß von realen Symptomen und Beschwerden, so darf der Gutachter nicht einfach durch einen Pauschal-Abzug von der geltend gemachten Störungsschwere die Aggravation in Rechnung stellen. Ein solches Vorgehen, wie es bedauerlicherweise immer wieder praktiziert wird, entbehrt, soweit hier bekannt, jeglicher rechtlichen Grundlage. Leider wird in solchen Fällen allzu häufig von Sachverständigen übersehen, dass auch hier ein *Nachweis* von Gesundheitsstörungen oder Funktionseinschränkungen mit der jeweils rechtlich gebotenen Beweisstärke zu erfolgen hat – und keineswegs in pauschalisierend spekulativer Weise vorgenommen werden darf. Dies setzt, das sei unbenommen, u. U. höchste Anforderungen an die gutachtliche Qualifikation und erfordert häufig eine minuziöse Aufarbeitung der Datenlage und eine logisch stringente, klar herausgearbeitete Argumentation, die am Ende potenziell sogar für alle Beteiligten überzeugend sein sollte. – Das heißt also, noch einmal auf den Punkt gebracht: eine nachgewiesene Aggravation schließt die Notwendigkeit für den (explizit zu führenden) Nachweis eines realen Kerns an Funktions- oder Gesundheitsstörungen keineswegs aus, aber eine Aggravation kann einen solchen Nachweis auch komplett verhindern. Da es vor diesem Hintergrund in der Praxis häufig unmöglich ist, ausreichend sicher zwischen einer Aggravation und einer Simulation zu unterscheiden (die im englischen *malingering* ohnehin begrifflich zusammengefasst sind), könnte etwa, wie bereits weiter oben ausgeführt, die Bezeichnung *simulative Tendenzen* benutzt werden (Wurzer, 1992). Noch weniger als bewertend empfunden, in der Sache aber auch kaum missverständlich dürfte die Bezeichnung einer *interessegeleiteten Darstel-*

lung sein, die ein Proband zeigt, wenn er in manipulativer Weise, bei eingeschränkter Mitwirkung, ungültige, nicht glaubhafte Daten produziert.

1.6.2 Falsch-positive Ergebnisse

Die Ausführungen des vorangegangenen Abschnitts setzen voraus, dass die Ergebnisse der Beschwerdenvalidierung korrekt (richtig-positiv) die in Wahrheit vorliegenden Verhältnisse widerspiegeln, also zu Recht die Ungültigkeit oder eingeschränkte Gültigkeit der Befundlage ausweisen. Es zeigt einen kaum verzeihlichen Mangel an wissenschaftlicher Sachkenntnis an, wenn BVT gelegentlich vorgeworfen wird, sie erzeugten nicht in 100 Prozent der Fälle korrekte Ergebnisse. Ein solcher Vorwurf muss alle empirisch-wissenschaftlichen Methoden, alle Tests und Prüfprozeduren treffen. Absolute Fehlerfreiheit einer psychologischen Messmethode ist ein Maßstab, der lediglich Sachkenntnis verrät, wenn er angelegt wird; ein solcher Maßstab wird im Übrigen auch für viele Messinstrumente selbst in reinsten Naturwissenschaften verfehlt.

Für alle gängigen BVT liegen Schätzungen der Sensitivität und Spezifität vor, aus denen sich auch die Wahrscheinlichkeiten bestimmen lassen, mit denen in verschiedenen Untersuchungskontexten ein positives Testergebnis für oder gegen das Vorliegen des Sachverhalts (hier: ungültige Befunde) spricht (siehe dazu weiter unten, Abschnitt 1.7). Um ein krasses Beispiel aufzuführen, das auch die besondere Bedeutung der Prävalenz in einem spezifischen Untersuchungskontext deutlich macht, sei mit Chafetz (2022) das folgende Szenario skizziert: der Grenzwert eines Tests werde so festgelegt, dass er eine 100-prozentige Sensitivität, aber nur 90-prozentige Spezifität aufweise. In einem Kontext, in dem die Prävalenz der infrage stehenden Bedingung (z. B. Erkrankung) auf lediglich 1% zu schätzen ist, würde dies dazu führen, dass von 100 getesteten Personen 11 positiv abschnitten, von diesen aber lediglich eine korrekt positiv (9%) und 10 falsch-positiv (91%) wären. Dieses nur dem Anscheine nach paradoxe und ungünstige Ergebnis würde aber, kontextabhängig, u. U. einen exzellenten Test markieren.

Bereits in der Frühphase der Ära der modernen Beschwerdenvalidierung herrschte weitgehende Einigkeit in der wissenschaftlichen und gutachtlichen Literatur dahingehend, dass angesichts der „Kosten“ falsch-positiver und falsch-negativer Beurteilungen bzgl. der Beschwerdenvalidierung die Spezifität der Verfahren hoch anzusetzen sei. Das waren für kognitive BVT zunächst 95%. Dieses Kriterium, bei dem eine reduzierte Sensitivität in Kauf genommen wird, um eine größere Zahl falsch-positiver Ergebnisse zu vermeiden, findet sich inzwischen häufig aufgeweicht und auf 90%, also letztlich Screening-Niveau, reduziert. Aus der Perspektive des Unterzeichners sollte es für solide und gut konstruierte (eigenständige) kognitive BVT jedoch weiterhin angestrebt werden, während die schwieriger zu entwickelnden fragebogenbasierten Validitätsskalen auf Siebtestniveau durchaus

eine Spezifität von nur 90 % aufweisen können, für höhere Grade der Gewissheit von Beurteilungen jedoch ebenfalls 95 % in Ansatz gebracht werden können. Für die eingebetteten kognitiven BVT (siehe Kapitel 6, in diesem Band) wird sich das hohe 95-%-Niveau kaum realisieren lassen, weil für diese bei einer so hoch gesetzten Zielmarke in der Regel die Sensitivität drastisch abfallen würde. Der Trend heute geht ohnehin dahin, den Einsatz mehrerer Verfahren zur Beschwerdenvalidierung im Verlauf einer Untersuchung (insbesondere im gutachtlichen Kontext) zu fordern und in Abhängigkeit von deren Anzahl auch die Zahl positiver BVT zu bestimmen, die für die Feststellung einer ungünstigen Datenlage notwendig ist (vgl. dazu Henry, Kapitel 5, in diesem Band).

Mit Bezug auf die jeweilige Zielgruppe der Anwendung der Verfahren liefern die Daten zur Spezifität Aussagen darüber, wie hoch der Anteil von *zu erwartenden* falsch-positiven Klassifikationen ausfällt. Die Möglichkeit eines falsch-positiven Ergebnisses ist im Einzelfall stets sorgfältig zu prüfen. Diese Prüfung darf aber keineswegs pauschal oder plakativ erfolgen, sie sollte in einer so überzeugenden Weise geschehen, dass sich auf keinen Fall der Eindruck oder gar die Überzeugung vermitteln, unliebsame positive Ergebnisse würde wegdiskutiert, etwa durch die Proklamation einer Anpassungsstörung. Eine Anpassungsstörung ist jedoch in Wahrheit in keiner Weise geeignet, positive Ergebnisse kognitiver oder fragebogenbasierter BVT zu bedingen (vgl. dazu auch Tabelle 14 in Merten, 2014).

Zur Illustration falsch-positiver Ergebnisse in der Beschwerdenvalidierung kann etwa auf die ausführliche Kasuistik eines Gutachtenfalls bei Henry, Merten und Wallasch (2008) verwiesen werden. Es handelt sich um den Fall eines 55-jährigen Mediziners, der überzeugend und gänzlich widerspruchsfrei das klinische Bild einer (bis dahin nicht diagnostizierten) bereits fortgeschrittenen semantischen Demenz zeigte, und zwar mit rechts-temporal betontem Atrophiemuster in der Bildgebung. Im zweiten Durchgang des Test of Memory Malingering (Tombaugh, 1996) erreichte er lediglich 42, im dritten Durchgang 39 Punkte. Dieses positive Ergebnis konnte im Sinne konvergenter Beweislinien schlüssig als aus der Präsenz einer als unzweifelhaft authentisch zu bewertenden Störung der Objekterkennung und zusätzlich vorliegender Lern- und Behaltensstörungen resultierend interpretiert werden.

Ein höherer Anteil von falsch-positiven Ergebnissen ist naturgemäß dort zu finden, wo eine Anwendung außerhalb der Indikation für den Einsatz eines Verfahrens erfolgte. Dies ist beispielsweise dann der Fall, wenn der Amsterdamer Kurzzeitgedächtnistest (AKGT; Schmand & Lindeboom, 2005) bei Patienten mit klinisch offenkundigen kognitiven Störungen eingesetzt wird, für die er ausdrücklich nicht geeignet ist, oder wenn der deutschsprachige Strukturierte Fragebogen Simulierter Symptome (SFSS; Cima et al., 2003) regelmäßig bei Patientinnen mit Migrationshintergrund eingesetzt wird, ohne dass zuvor explizit geprüft wird, ob der Grad der deutschen Sprachbeherrschung für diesen Fragebogeneinsatz ausreichend ist. Zudem stellen insbesondere die eingebetteten kognitiven BVT ein besonderes Pro-

blem dar, das von Erdodi und Lichtenstein (2017) als Paradoxon gekennzeichnet wurde. Die Testverfahren, in die sie eingebettet sind, wurden zur Diagnostik kognitiver Störungen entwickelt und sollen deren Präsenz gegenüber sensitiv sein; für viele eingebettete Validitätskennwerte resultiert daraus eine erhöhte Gefahr, dass bei Vorliegen veritabler bedeutsamer kognitiver Störungen auch diese fälschlich positiv ausfallen.

1.6.3 Interpretation unauffälliger Ergebnisse

Unbedingt zu beachten gilt, dass unauffällige (negative) Ergebnisse einer eingehenden, sachgerecht durchgeführten Beschwerdvalidierung keineswegs mit einem geführten Nachweis der Validität der Befunde gleichgesetzt werden dürfen. Anders ausgedrückt, eine ungültige, zum Beispiel in manipulativer Weise verzerrt zustande gekommene Datenbasis oder simulative Tendenzen lassen sich durch negative Ergebnisse in der Beschwerdvalidierung *nicht ausschließen*. Darauf hat jüngst auch Chafetz (2022) aufmerksam gemacht. Eine fehlerhafte Darstellung der Bedeutung negativer Ergebnisse in der Beschwerdvalidierung dürfte zu den aktuell häufigsten Fehlern in der Gutachtenpraxis gehören, soweit dies die in diesem Buch behandelte Problematik betrifft. Dem liegt eine fehlerhafte Anwendung der Denkgesetze zugrunde.¹

Ebenso, wie ein Testergebnis falsch-positiv ausfallen kann, sind auch falsch-negative Ergebnisse zu erwarten, und zwar aufgrund des besonderen Augenmerks auf eine hohe Spezifität (zuungunsten der Sensitivität) bei Beschwerdvalidierungstests mit deutlich höherer Auftretenshäufigkeit. Dies spiegelt sich auch in den positiven und negativen prädiktiven Werten wider (siehe dazu den folgenden Abschnitt, Tabelle 1.3).

Mit anderen Worten: ebenso, wie sich lediglich aus der Tatsache einer Konsistenz und Plausibilität in der Datenlage im logischen Sinne noch kein Beweis dafür ergibt, dass eine behauptete Symptomatik in Wahrheit vorliegt, kann dies auch durch die Unauffälligkeit von Beschwerdvalidierungstests nicht bewiesen werden. Dies ist bei der Formulierung zu berücksichtigen, die für die Ergebnisdarstellung und -interpretation gewählt wird.

Trotz dieser Einschränkung für die Generalisierbarkeit eines unauffälligen Ergebnisses in der formalen Beschwerdvalidierung und namentlich bei einschlägigen, gut entwickelten BVT sind solche Resultate von herausragender Bedeutung. Sie belegen, mit welcher Sorgfalt entsprechend des aktuellen Wissens- und Me-

1 Bizarr und ethisch sehr fragwürdig wird eine solche fehlerhafte Anwendung logischer Gesetze etwa dann, wenn, wie immer wieder beobachtbar, als einziger BVT der Fifteen Item Test (FIT) von Rey (1958) eingesetzt wird und etwa aus einem dort ermittelten Punktwert von 10 dem Auftraggeber suggeriert wird: „Eine Simulation oder Aggravation kann demzufolge ausgeschlossen werden.“

thodikstandes die Untersucherin die Validität der Datenlage geprüft hat. Wenn diese sich in sich als vollständig widerspruchsfrei darstellt und standardisierte Verfahren zur Beschwerdenvalidierung komplett unauffällig ausfallen, kann auch kein Grund dafür erkannt werden, die Validität der Datenlage zu bezweifeln. Dies ist auch dann so, wenn objektiv dieses Ergebnis falsch ist, was aber in einem solchen Fall nicht zu Lasten der Sachverständigen oder Behandlerin gehen kann, denn diese kann nur innerhalb der gutachtlichen oder klinischen Erkenntnismöglichkeiten agieren. Angesichts der (voll offengelegten) Datenbasis, wenn diese soweit wie im gegebenen Kontext notwendig gesichtet oder erhoben wurde, hätte auch jede andere vernünftig und sorgfältig urteilende sachkundige Person zu der gleichen Einschätzung gelangen müssen (selbst wenn, dies sei noch einmal hervorgehoben, diese Beurteilung objektiv falsch sein sollte).

1.6.4 Falsch-negative Ergebnisse

Gründe für falsch-negative Ergebnisse in standardisierten BVT können vielfältig sein. Einige seien hier aufgeführt. Eine von ihnen seien hier aufgeführt:

1. Das eingesetzte Verfahren oder der umgesetzte Untersuchungsansatz ist für die Fragestellung inhaltlich nicht geeignet (z. B. liefert das Ergebnis eines kognitiven BVT keine Grundlage dafür, den Anspruch eines Patienten bei der Hausärztin, er leide seit dem Vortag unter kaum tolerierbaren Bauchschmerzen, hinsichtlich seiner Glaubhaftigkeit zu überprüfen).
2. Die Gutachterin hat in ihrer Auswahl auf ungeeignete Verfahren zurückgegriffen, also zum Beispiel den Fifteen Item Test (Rey, 1958) als einzigen BVT am Ende der gutachtlichen Untersuchung eines 30-jährigen Versicherten nach Arbeitsunfall mit leichtem Schädel-Hirn-Trauma durchgeführt. Hier schimmert durch, dass bereits die Auswahl des Tests nur mit Alibifunktion erfolgte. Eine Zeitlang war es in einem Gutachteninstitut im deutschsprachigen Raum üblich, routinemäßig eine einzelfallexperimentelle Anordnung von Merten und Puhlmann (2004) zur „Abdeckung“ der Beschwerdenvalidierung durchzuführen, ohne zu erkennen, dass die Anwendungsvoraussetzungen für die gutachtliche Routinediagnostik gar nicht gegeben waren. Jener „Rechentest“ war vielmehr *für einen Einzelfall* mit unplausibel geltend gemachter Akalkulie konstruiert worden. Seine wiederholte Verwendung wäre höchstens noch in anderen Fällen mit geltend gemachter, aber unplausibel erscheinender kompletter Rechenunfähigkeit zu rechtfertigen gewesen, nicht aber in jeder gutachtlichen Untersuchung.
3. Fälschlich wird bei Einsatz eines Tests mit Alternativwahlformat lediglich ein Antwortmuster statistisch unterhalb des Zufallsbereichs als Kriterium für eine auffällige Leistung gewählt. Damit wird die Mehrzahl in Wahrheit nicht glaubhafter, ungültiger Testprofile nicht erkannt und fälschlich als valide ausgewiesen.
4. Eine Probandin kann sich soweit auf die Untersuchung vorbereitet haben oder ist durch einen kommerziell agierenden „Rentenberater“ oder einen unethisch