

Leitlinie Validierung psychologischer Befunde



Deutsche Gesellschaft für
Psychologische Begutachtung

2025

Inhaltsverzeichnis

PRÄAMBEL/ZWECK DER LEITLINIE	3
1. STELLENWERT DER VALIDIERUNG IM DIAGNOSTISCHEN PROZESS.....	4
2. MINDESTSTANDARDS DER VALIDIERUNG	5
3. VALIDIERUNGSDIAGNOSTISCHE AUSGANGSBEDINGUNGEN.....	6
4. VALIDIERUNGSDIAGNOSTISCHE BEGRIFFE	8
5. EMPFEHLUNGEN ZUR HYPOTHESENBUILDUNG	11
6. VALIDIERUNGSDIAGNOSTISCHE UNTERSUCHUNGSPLANUNG	14
7. DURCHFÜHRUNG DER UNTERSUCHUNG.....	18
8. EMPFEHLUNGEN ZUR VALIDIERUNGSDIAGNOSTISCHEN URTEILSBILDUNG	20
9. EINBEZUG VON VALIDIERUNGSBEFUNDEN IN DIE INHALTLICHE ERGEBNISINTERPRETATION	26
10. BEANTWORTUNG DER RECHTLICHEN FRAGESTELLUNGEN IM GUTACHTEN	28
LITERATUR.....	31

Präambel/Zweck der Leitlinie

Die vorliegende Leitlinie dient v. a. der wissenschaftlich begründeten und praxisorientierten Aufstellung und Prüfung validierungsdiagnostischer Hypothesen im Rahmen einer psychologischen Begutachtung von ...

- kognitiven, psychomotorischen oder emotional-motivationalen *Beschwerden*, die im Zusammenhang mit neurologischen oder anderen körperlichen Erkrankungen oder psychischen Störungen beklagt werden;
- Angaben oder Verhaltensweisen, die auf krankheits- oder störungsbedingte kognitive, motorische, psychische oder soziale *Funktions- und Leistungseinschränkungen* hinweisen oder hinweisen können.

Die auf gutachtliche bzw. entscheidungsorientierte Tätigkeit ausgerichtete Leitlinie grenzt sich ab von behandlungsorientierten medizinischen und psychotherapeutischen Leitlinien. Sie soll u. a. die Qualitätsstandards für psychologische Gutachten der Föderation deutscher Psychologenvereinigungen, die Leitlinie psychologische Begutachtung SGB XIV, in Vorbereitung zur Herausgabe durch die DGPpB, sowie die „Leitlinie Neuropsychologische Gutachten“ (Neumann-Zielke et al., 2014) zu Fragen der Sicherung psychologischer Befunde um Spezifika der Validierungsdiagnostik und des validierungsdiagnostischen Vorgehens erweitern.

1. Stellenwert der Validierung im diagnostischen Prozess

Wird ein (neuro-)psychologisches Gutachten oder Zusatzgutachten zur Beurteilung der allgemeinen oder beruflichen Funktions- oder Leistungsfähigkeit von einem Gericht in Auftrag gegeben, können körperliche Erkrankungen und/oder psychische Störungen bei der zu begutachtenden Person vorliegen. Die Schilderung und Demonstration gesundheitlicher Beschwerden sowie darauf bezogener Funktions- und Leistungseinschränkungen sind insofern zentrale Bestandteile (neuro-)psychologischer Diagnostik im Rahmen einer Begutachtung.

Wenn eine Person um die rechtliche Anerkennung gesundheitlicher Schädigungen, Krankheiten, psychischer Störungen oder Schädigungs- und Krankheitsfolgen streitet, liegt der Fokus der betroffenen Person, ihrer Angehörigen und ggf. auch der behandelnden Sachverständigen überwiegend auf den berichteten oder beobachteten Funktions- oder Leistungsdefiziten. Häufig wird die rechtliche Anerkennung gesundheitsbedingter Funktions- oder Leistungseinschränkungen auch als psychische Entlastung und Bestätigung des subjektiv empfundenen Leidens angesehen. Vor diesem sozial-kommunikativen Hintergrund sind willentlich und interessengeleitet akzentuierte Beschreibungen von psychischen, körperlichen oder kognitiven Beschwerden und Funktionseinschränkungen in der Begutachtung relativ häufig.

Studien zur Prävalenz von interessengeleiteten Antwortverzerrungen im Rahmen der Begutachtung gesundheitlicher Schädigungen sprechen dafür, dass etwa 30 bis 50 % der Begutachtungsfälle von ergebnisorientiert verzerrten Angaben betroffen sind (Stevens et al., 2008). Auf dieser Grundlage sollten z. B. Ergebnisse von Leistungstests konsequent mit Bezug auf die motivationalen Bedingungen interpretiert werden, unter denen sie erzielt wurden.

Auch Selbstauskünfte sollten validiert werden, z. B. indem Antwortmuster oder Antwortinhalte zu Informationen in Beziehung gesetzt werden, die eine Beurteilung der Gültigkeit und Verwertbarkeit der Angaben erlauben. Dabei ist grundsätzlich zwischen negativen und positiven Antwortverzerrungen zu unterscheiden. Die Prävalenzen beider Tendenzen können kontextabhängig variieren, da sie von diversen Einflüssen abhängig sind. So sind z. B. bei nachgewiesenen hirnorganischen Funktionsstörungen im Rahmen einer unmittelbar auf eine Operation folgenden Anschlussheilbehandlung nur wenige nicht-authentische Beschwerdenschilderungen zu erwarten, während die Prävalenz bei späteren Begutachtungen auf etwa 20 % ansteigt (Bahlo, 2014). Differenziertere Angaben zur Häufigkeit und zu den Bedingungen *negativer Antwortverzerrungen* liefern z. B. Larrabee (2012), Martin und Schröder (2022), Merten (2014; 2023), Dettenborn und Merten (2009), Rogers und Bender (2018) und Young (2014).

Die Prävalenz *positiver Antwortverzerrungen* variiert ebenfalls in Abhängigkeit von Untersuchungsbedingungen (z. B. Eignungsdiagnostik, Sorgerechtsstreitigkeiten, Kriminalprognose, strafrechtliche Bewertungen), personalen Faktoren und zugrunde gelegten Schwellenkriterien. Mit Bezug auf die wissenschaftliche Literatur kann erwartet werden, dass in einem Setting mit erhöhtem Risiko positiv verzerrter Angaben ca. 20 – 40 % der Probanden auffällig ergebnisverfälschende positive Antwortverzerrungen aufweisen (Baer & Miller, 2002; Dohrenbusch & Brockhaus, 2025; Grossman & Wasyliv, 1988; Heilbrun et al, 1990; Tippelt, 2015).

A. Die Forderung nach rechtlicher Anerkennung gesundheitlicher Schädigungen, Krankheiten oder Krankheits- bzw. Schädigungsfolgen lenkt im Verlauf des Untersuchungs- und Anerkennungsprozesses den Fokus häufig auf vorhandene Defizite. So steigt das Risiko, dass Beschwerden oder Beeinträchtigungen gegenüber Sachverständigen und Entscheidungsträgern interessengeleitet verzerrt dargestellt werden. Die Forschungslage spricht für eine Basisrate von ca. 40 % interessengeleitet negativ verzerrt dargestellter Sachverhalte im rechtlichen Begutachtungskontexten.

2. *Mindeststandards der Validierung*

Internationale Forschung belegt die exponierte Bedeutung validierungsdiagnostischer Probleme und Lösungsansätze sowie validitätssichernder Verfahrensweisen im Kontext klinischer und rechtspsychologischer Begutachtung (z. B. Boone, 2021; Boskovic et al, 2017; Fokas & Brovko, 2020, Groth-Marnath & Wright, 2016; Hall & Poirier, 2020; Horton & Reynolds, 2021; Larrabee, 2012; 2015; Martin & Schroeder, 2022; Picard et al., 2023; Rohling et al., 2011; Schroeder, 2022; Thies, 2012; Wygant et al. 2010).

Die in Deutschland veröffentlichten diagnostischen Qualitätsstandards für psychologische Gutachten der Föderation deutscher Psychologenvereinigungen sehen eine Validierung individueller Untersuchungsergebnisse im Rahmen der psychologischen Befunderhebung vor¹. Diese betrifft bei psychometrischer Messung v. a. die Zuverlässigkeit (Reliabilität, Objektivität) und Gültigkeit (Validität) von Mess- und Testergebnissen, darüber hinaus auch die logische Konsistenz und fachliche Plausibilität frei erhobener Informationen zu Beschwerden, Krankheitsverläufen und Fähigkeitseinschränkungen. Validierungsdiagnostische Themen und Verfahrensweisen im engeren Sinne (zu Fragen der Unverfälschtheit bzw. Authentizität individueller Testergebnisse, Beobachtungen und Selbstauskünfte) werden in den Standards jedoch nur vergleichsweise schwach behandelt verglichen mit internationalen Standards (z. B. Ferrara et al, 2016; Sherman, 2020; Sweet et al, 2021) und Best-Practice-Empfehlungen (Chafez et al., 2025).

Der validierungsdiagnostische Fachdiskurs im deutschsprachigen Raum wird aktuell durch verschiedene Faktoren limitiert:

- So wird die Terminologie zur Validierungsdiagnostik nicht einheitlich verwendet, wodurch konzeptionelle Missverständnisse entstehen können. Zu fordern ist ein einheitlicher Gebrauch validierungsdiagnostischer Begriffe.
- Uneinheitlich sind auch die veranschlagte Bedeutung klinischer Konzepte und Störungsmodelle für die Validierung psychologischer und psychiatrischer Befunde, die Grenzziehungen zwischen krankheits- bzw. störungsbedingten und nicht krankheitsbedingten Ergebnisverzerrungen sowie der Umgang mit standardisierten und normierten Validierungstests.
- Zur Frage geeigneter Beurteilungskriterien sind nach Merten (2023) die so genannten Slick-Kriterien (Slick, Sherman & Iverson, 1999) im Bereich der neuropsychologischen Begutachtung mit Schwerpunkten auf der Validierung funktions- und leistungsdiagnostischer Untersuchungsergebnisse handlungsleitend, die auch im amerikanischen Rechts- und Begutachtungssystem übernommen wurden. Der neuropsychologische Validierungsschwerpunkt ist jedoch nicht für alle psychologischen Anwendungsbereiche gleichermaßen geeignet.
- Es existiert nur eine begrenzte Zahl deutscher bzw. an deutsche Verhältnisse angepasster validierungsdiagnostischer Mess- und Testverfahren. Die in Deutschland verwendeten Verfahren wurden überwiegend aus dem anglo-amerikanischen Sprachraum übernommen, übersetzt und neu normiert bzw. validiert.

Bislang fehlt in Deutschland u. a. eine wissenschaftlich begründete Leitlinie dazu, wie spezifische standardisierte Validierungsverfahren in die validierungsdiagnostische Urteilsbildung eingebunden werden

¹ Testkuratorium der Föderation deutscher Psychologenvereinigungen. Qualitätsstandards für psychologische Gutachten (2017). Online Publikation https://www.dgps.de/fileadmin/user_upload/PDF/Empfehlungen/GA_Standards_DTK_10_Sep_2017_Final.pdf

können.

Die Arbeitsgruppe zur Validierungsdiagnostik der DGPsB sieht daher *validierungsdiagnostische Mindeststandards* als notwendig an, um die *Validierung individueller Untersuchungsergebnisse im Sinne einer kontrollierten Praxis* auf eine wissenschaftliche Grundlage zu stellen und der Validierung psychologischer Untersuchungsergebnisse einen angemessenen Stellenwert zukommen zu lassen. Die vorliegenden Mindeststandards umfassen:

- Angaben zu validierungsdiagnostischen Ausgangsbedingungen
- Definitionen validierungsdiagnostischer Begriffe
- Angaben zur Hypothesenbildung
- Angaben zur validierungsdiagnostischen Methodenauswahl und Untersuchungsplanung
- Angaben zur validierungsdiagnostischen Ergebnisinterpretation und Urteilsbildung
- Angaben zur Integration von Validierungsbefunden in die Stellungnahme

B. Die aktuellen Standards und Empfehlungen zur Beurteilung der Validität (Gültigkeit) gesundheits- und funktionsbezogener Aussagen und Testleistungen in einem rechtlichen Untersuchungskontext sind teilweise unzureichend. Validierungsdiagnostische Mindeststandards sollten Angaben zur Berücksichtigung der individuellen objektiven und subjektiven Ausgangs- bzw. Anreizbedingungen, zur Festlegung von Untersuchungshypothesen, zum diagnostischen Vorgehen und zur gutachterlichen Urteilsbildung enthalten.

3. *Validierungsdiagnostische Ausgangsbedingungen*

Als validierungsdiagnostische Ausgangsbedingungen werden die anreizbezogenen, kontextuellen und personellen Bedingungen bezeichnet, unter denen validierungsdiagnostische Informationen erhoben und entsprechende Beurteilungen durchgeführt werden. Sie sind u. a. durch Basisraten und Risiken für Antwortverzerrungen bestimmt. Folgende Ausgangsbedingungen können unterschieden werden:

- a) *Äußere (objektive) Anreizbedingungen* für willentlich verzerrte Angaben oder Leistungsmotive. Sie können danach differenziert werden, welchem potentiell ergebnisverzerrenden Zweck die Untersuchung und das Untersuchungsergebnis dient (z. B. wissenschaftlicher Erkenntnisgewinn, Erlangung individueller Hilfe- oder Kompensationsleistungen durch Nachweis gesundheitlicher Beeinträchtigungen oder Unterstützungsbedarf, Erlangung von Vorteilen durch Nachweis positiver Eigenschaften wie z. B. psychischer Gesundheit, hoher Leistungsfähigkeit oder Ausschluss von Gefährdungsrisiken.)
- b) *Von der untersuchten Person subjektiv erlebte Anreizbedingungen und Ergebnisinteressen.* Sie sind bestimmt durch die mentale Repräsentation möglicher Anreizbedingungen oder wahrgenommener Vorteile, die sich aus dem Untersuchungsergebnis für die untersuchte Person ergeben oder ergeben können. Objektive Anreizbedingung und die mentale Repräsentation von Anreizen können übereinstimmen oder voneinander abweichen (z. B. „objektiver Anreiz“: Vorteile durch Nachweis psychischer Gesundheit, „subjektiver Anreiz“: Vorteile durch forcierte Anpassung an die Untersuchungsbedingungen). Daher ist die Erfassung der subjektiven Ergebnisinteressen auch bei vermeintlich eindeutigen äußeren Anreizbedingungen in aller Regel sinnvoll.

- c) *Art und Umfang eingesetzter validierungsdiagnostischer Mess- und Testverfahren.* Ob interessengeleitet verzerrte oder verfälschte Angaben oder Testleistungen in einer Untersuchung aufgedeckt und in die Ergebnisinterpretation einbezogen werden, hängt wesentlich vom Einsatz geeigneter validierungsdiagnostischer Verfahren und Diagnosesysteme ab. Der Verzicht auf evaluierte Validierungsverfahren und Kontrollskalen erhöht das Risiko unzureichender oder fehlerhafter Identifikation motivationaler Ergebnisverzerrungen und fehlerhafter Merkmalszuschreibungen.
- d) *Bereitschaft und Fähigkeit der Sachverständigen, wissenschaftlich evaluierte Validitätskriterien in die Beurteilung der Glaubhaftigkeit von Selbstauskünften oder Leistungstestergebnissen einzubeziehen.* Die Fähigkeit und Bereitschaft des Untersuchers/der Untersucherin kann sich auf die Wahrscheinlichkeit auswirken, mit der vorhandene Verzerrungen oder Verfälschungstendenzen zuverlässig aufgedeckt werden. Auf die tatsächliche Häufigkeit (Basisrate) interessengeleitet verzerrter oder fehlerhafter Angaben wirkt sich die Bereitschaft nicht aus.
- e) *Art und Gerichtetheit der validierungsdiagnostischen Hypothesen.* Hypothesen können ihren Ausgang in den Vorgaben des Auftraggebers haben (z. B. dem Erfordernis der forensischen Beweiswürdigung, die Hypothese so zu formulieren, dass falsche Angaben als Ausgangsbedingung unterstellt werden). Sie können sich auch auf Aspekte der psychologischen Befundsicherung stützen und von der Praxis der rechtlichen Beweiswürdigung abweichen. Bei der Befundsicherung können die von äußeren Anreizbedingungen ausgehenden gerichteten Hypothesen einen Beitrag dazu leisten, mit welcher Wahrscheinlichkeit Antwortverzerrungen aufgedeckt werden.
- f) *Das praktische Vorgehen bei der Hypothesenprüfung.* Ein leitlinienorientiertes Vorgehen, gemessen u. a. an einer einheitlichen fachlichen Begriffswahl, hypothesengeleitetem Vorgehen, der Auswahl geeigneter Messverfahren, einem multimethodalen Abgleich von Validierungsinformation sowie kriterienorientierter Bewertung trägt mutmaßlich dazu bei, die Aufdeckung verzerrter oder verfälschter Ergebnisse zu verbessern.

C. Die vorliegende Leitlinie geht von der Dimensionalität validierungsdiagnostischer Eigenschaften wie „Anstrengungsbereitschaft“ (bei Leistungsmessungen) und „Antworttendenz“ (bei Selbstauskünften) und der multifaktoriellen Bedingtheit dieser Merkmale aus. Ob verzerrte oder verfälschte Angaben, Darstellungen und/oder Mess- bzw. Testergebnisse in einer Untersuchung zutreffend identifiziert werden, kann von äußeren Rahmenbedingungen, objektiven und subjektiven Anreizbedingungen und der Qualität des diagnostisch-gutachterlichen Vorgehens abhängen. Die psychologische Urteilsbildung zu Art und Ausprägung mutmaßlich willentlich verzerrter Reaktionen stützt sich primär auf wahrscheinlichkeitsbasierte Vergleiche und probabilistische Erklärungs- oder Vorhersagemodelle. Kategoriale Zuschreibungen (z. B. „glaubhaft – nicht glaubhaft“) ohne Einbezug der Ergebnisse dimensionaler und wahrscheinlichkeitsbasierter validierungsdiagnostischer Verfahren sind in aller Regel unzureichend.

4. Validierungsdiagnostische Begriffe

Die DGPsB schlägt im Sinne einer fachsprachlichen Vereinheitlichung folgende Definitionen bzw. Umschreibungen für validierungsdiagnostische Begriffe vor:

Aggravation: Bewusste Übertreibung oder Ausweitung von Klagen über gesundheitliche Beschwerden oder Beeinträchtigungen, bei denen angenommen wird, dass ein tatsächlicher authentischer Kern an Beschwerden vorhanden ist, der aber willentlich übertrieben wird, entweder in seiner Schwere oder durch Ausweitung auf andere Beschwerdenbereiche.

Antworttendenz: Beschreibung der Neigung einer Person, Testwerte oder Aussagen über ein zugrunde liegendes Merkmal, ein Ereignis oder eine Eigenschaft nicht zutreffend zu beschreiben. Antworttendenzen können unterschiedlich eng (z. B. syndrom- oder themenspezifisch) oder weit gefasst (generalisiert) wirksam sein. Sie können sowohl aufgrund unwillkürlicher bzw. unbewusster Einflüsse wie z. B. Stimmung oder Müdigkeit, als auch aufgrund willentlicher Einflüsse zustande kommen. Bei Selbstauskünften wird von inhaltlichen Antworttendenzen gesprochen, wenn die befragte Person aktiv versucht, die Darstellungen an bestimmten Ergebnisinteressen auszurichten.

Anstrengungsbereitschaft: Die in der Leistungsdiagnostik geforderte Bereitschaft, das bestmögliche Leistungsniveau abzurufen und ein Testverfahren instruktionsgemäß zu bearbeiten. Diese motivationale Voraussetzung für den psychodiagnostischen ist Teil der informierten Einwilligung.

Authentizität: Authentizität ist ein Orientierungsbegriff, der die Möglichkeit beschreibt, dass die Äußerungen einer Person tatsächliche psychische Sachverhalte oder Erlebnisse (z. B. psychische Eigenschaften, Fähigkeiten, Erlebnisqualitäten) zutreffend abbilden. Von Authentizität beklagter oder geltend gemachter körperlicher, psychischer und/oder kognitiver Beschwerden kann umso eher gesprochen werden, je eindeutiger diese dem gesundheitlichen Zustand oder dem tatsächlichen Funktionsniveau der Person entsprechen. Authentizität kann auch als Art und Ausmaß gültiger, d. h. unverzerrter Äußerungen bestimmt werden. Ob Selbstbeschreibungen einer Person gänzlich authentisch sind, kann durch den Ausschluss verfälschender Antworttendenzen nicht vollständig gesichert werden, weil auch nach Ausschluss einzelner potentiell verfälschender Antworttendenzen immer noch die Möglichkeit besteht, dass die Authentizität der Selbstbeschreibung von nicht geprüften Störeinflüssen oder Verzerrungen beeinflusst ist. Eine dichotome Unterscheidung, ob eine Person generell als authentisch oder nicht authentisch zu bewerten ist, sollte vermieden werden.

Beschwerdenuvalidierung. Bezeichnung für die Gesamtheit der diagnostischen Maßnahmen, die zu Aussagen über die Gültigkeit gesundheitlicher Beschwerden oder daraus resultierender Funktionsbeeinträchtigungen führen.

Dissimulation. Tendenz, eine Erkrankung zu negieren oder zu bagatellisieren oder sich als psychisch unbelastet darzustellen, obwohl gesundheitliche Belastungen oder Beeinträchtigungen in Wirklichkeit vorliegen. Dissimulation kann darin zum Ausdruck kommen, dass vorhandene Beschwerden oder Beeinträchtigungen in ihrer Qualität, ihrer Schwere oder der Breite ihrer Auswirkungen verharmlost oder geleugnet werden oder die eigene Leistungsfähigkeit übertrieben positiv dargestellt wird. So wie bei der Aggravation wird auch bei der Dissimulation angenommen, dass die verzerrte Darstellung vorwiegend durch willentliche, interessengeleitete Einflüsse bedingt ist. Im Untersuchungskontext wird Dissimulation typischerweise mit positiven Antwortverzerrungen in Verbindung gebracht.

Glaubhaftigkeit. Bezeichnung für das Ausmaß, in dem verbale Äußerungen oder das gezeigte Verhalten einer Person einen bestimmten Sachverhalt (z. B. ein Ereignis, eine Eigenschaft oder Eigenschaftsausprägung, eine Fähigkeit oder Fähigkeitsausprägung) in der Untersuchungssituation zutreffend beschreiben. Die Glaubhaftigkeit einer Zeugenaussage kann umso höher veranschlagt werden, je wahrscheinlicher die Aussage der Person auf einem tatsächlichen Erlebnisbezug basiert. Dieser ist nicht gleichzusetzen mit dem faktischen Realitätsgehalt des beschriebenen Ereignisses. Die aussagepsychologische Beurteilung der Glaubhaftigkeit von Aussagen erfordert ein mehrstufiges Vorgehen. Dabei wird zunächst geprüft, ob die Aussage die Voraussetzungen der aussagenden Person für gerichtsverwertbare Aussagen verwendet werden kann. Sofern sich der Ursprung und die Entwicklung einer Aussage ohne wesentliche Verfälschungen darstellt (insbesondere Ausschluss von Suggestion) wird schließlich geprüft, inwiefern die Aussage hinreichende Merkmale aufweist, die in erlebnisbasierten, nicht aber in frei erfundenen Schilderungen zu erwarten sind.

Glaubwürdigkeit. Glaubwürdigkeit als Rechtsbegriff bezeichnet die Vertrauenswürdigkeit eines Zeugen im Rahmen der richterlichen Beurteilung einer Zeugenaussage. In der psychologischen Nomenklatur Bezeichnung für einen eigenschaftsorientierten Ansatz, der davon ausgeht, dass die Glaubwürdigkeitswertigkeit (Beurteilung des Zutreffens von Aussagen) von Zeugenaussagen aus ihrer Zuordnung zu Persönlichkeitsmerkmalen der aussagenden Person möglich ist.

Konsistenz. Bezeichnung für das Ausmaß, in dem wiederholte Messungen oder Beobachtungen des gleichen Merkmals zu übereinstimmenden Ergebnissen führen. Es ist insofern ein Maß für „Beständigkeit“. In der Testpsychologie wird innere Konsistenz als Indikator für die Messgenauigkeit (Reliabilität) eines Tests interpretiert, die durch die mittlere Höhe der Interkorrelationen der Einzelbeobachtungen (Items) und die Höhe der Korrelationen der Items (Testteile) mit dem Gesamtestwert bestimmt ist. Bei freier Beobachtung liegt Verhaltenskonsistenz in dem Maße vor, in dem das beobachtete oder gemessene Verhalten über verschiedene Situationen stabil bleibt. In freier Exploration werden inhaltliche Äußerungen als konsistent bezeichnet, wenn sie sich logisch nicht widersprechen.

Objektivität. Objektivität im testpsychologischen Sinne bezeichnet das Ausmaß, in dem mehrere Beobachter bzw. Untersucher zu übereinstimmenden Bewertungen eines bestimmten (gemessenen oder beobachteten) Sachverhalts gelangen. Die gutachterliche Objektivität kann eingeschränkt sein durch individuelle Schwerpunkte oder Verfahrensweisen bei der Auswahl und Erhebung von Informationen sowie durch die individuelle Gewichtung der erhobenen Informationen und damit gegebenenfalls verbundene Urteilsfehler.

Performanzvalidierung. Bezeichnung für die Sammlung, Gewichtung und Integration von Informationen zu den motivationalen Bedingungen, unter denen eine leistungsdiagnostisch untersuchte Person Testleistungen erzielt hat. Entsprechend wird die Bezeichnung Performanzvalidierungstest für alle psychometrischen Tests, Kennwerte und Indikatoren verwendet, die ungültige Leistungen bei der Prüfung der kognitiven oder motorischen Leistungsfähigkeit hinreichend zuverlässig erfassen. Meist sind Performanzvalidierungstests dem Anschein nach zur Leistungserfassung konzipiert, tatsächlich erfassen sie aber die Testmotivation eines Probanden in der Testuntersuchung. Die Tests können auf unterschiedlichen Prinzipien der Beschwerdenvalidierung basieren. Eingesetzt werden sie vor allem bei erhöhtem Risiko negativer Antwortverzerrungen.

Plausibilität. Bezeichnung für ein fachliches Bewertungskriterium, nach dem erhobene Informationen zu psychologischen Sachverhalten bewertet und z. B. in Bezug auf Glaubhaftigkeit vs. Unglaubhaftigkeit oder Gültigkeit vs. Ungültigkeit unterschieden werden können. Psychologische Plausibilitätsbewer-

tungen sind typischerweise wahrscheinlichkeitsbasiert und auf psychologische Heuristiken oder Erklärungsmodelle gestützt. Ein Ergebnis bzw. psychologischer Befund ist mit einer bestimmten Wahrscheinlichkeit plausibel, wenn die Ergebnisse oder Beobachtungen in einen theoretischen (wissenschaftlichen) oder begrifflichen Kontext weitgehend widerspruchsfrei und stimmig integriert werden können.

Psychologischer Befund. Ein psychologischer Befund im Kontext einer Begutachtung ist die fachliche Analyse und Bewertung der zu einem bestimmten Zweck erhobenen psychologischen Informationen. Ein einzelnes Testergebnis, eine einzelne Äußerung oder eine einzelne Beobachtung wird in der Regel nicht als „Befund“ bezeichnet. Typischerweise werden die Ergebnisse mehrerer Datenquellen und mehrerer Untersuchungsebenen in einen psychologischen Befund integriert. Dazu werden die erhobenen Informationen systematisch und mit geeigneten Validierungsmethoden danach bewertet, wie zuverlässig und valide sie die zu beschreibenden psychischen Sachverhalte (Ereignisse, Erfahrungen, Eigenschaften, Fähigkeiten usw.) abbilden und wie ausgeprägt die Darstellungen oder Verhaltensweisen mutmaßlich von den Untersuchungsbedingungen (einschließlich Antwortmotiven der Untersuchten) beeinflusst waren. Die Beschränkung der Darstellungen und Bewertungen auf die Probandenperspektive ist für einen psychologischen Befund nicht ausreichend.

Reliabilität. Im testpsychologischen Sinne die Zuverlässigkeit oder Genauigkeit, mit der ein psychologischer Test misst. Die Reliabilität kann umso höher veranschlagt werden, je häufiger wiederholte Messungen des gleichen Merkmals oder der gleichen Eigenschaft zu übereinstimmenden Ergebnissen kommen. Bei der Beurteilung frei formulierter Aussagen kann die Konsistenz bzw. Übereinstimmung mehrerer Aussagen zum gleichen Sachverhalt als Indikator für die Erfassung derjenigen Varianzanteile angesehen werden, die in testpsychologischen Verfahren durch die Reliabilität kontrolliert werden.

Simulation. Bezeichnung für das absichtliche Vortäuschen tatsächlich nicht vorhandener Beschwerden oder Beeinträchtigungen. Häufig wird als weiteres Kriterium für Simulation der Anreiz zum Täuschen durch äußere Bedingungen genannt, der das Täuschungsverhalten erklärt. Bei Performanz-Validierungsverfahren nach dem Forced-Choice-Prinzip gilt ein signifikant unterhalb der Ratewahrscheinlichkeit angesiedeltes Ergebnis als Anhaltspunkt für bewusste Antwortverfälschungen im Sinne der Simulation.

Unverfälschbarkeit. Nebentestgütekriterium. Ein Test erfüllt das Gütekriterium der Unverfälschbarkeit, wenn die getestete Person ihr Testergebnis nicht oder nur unwesentlich nach eigenem Belieben beeinflussen kann.

Validität. Testgütekriterium. Unter Validität eines Tests ist zu verstehen, dass er tatsächlich jenes Merkmal (z. B. Zustand, Fähigkeit, Persönlichkeitseigenschaft) misst, welches er zu messen vorgibt. Unterschieden werden die inhaltliche Gültigkeit (logische Validität), die Konstruktvalidität und die Kriteriumsvalidität. Unter einem hypothetischen Konstrukt versteht man „Merkmale, Zustände oder Instanzen, die nicht direkt beobachtbar sind, sondern aufgrund von Verhaltensbeobachtungen erschlossen werden.“

Validierung. Bezeichnung für unterschiedliche Maßnahmen zur Sicherung der Aussagekraft, der Zuverlässigkeit, Gültigkeit, teilweise auch der Vorhersagekraft erhobener Merkmale oder Eigenschaften für das Erleben oder Verhalten der untersuchten Person außerhalb der Untersuchungssituation. Die *Validierung individueller Messergebnisse* erfolgt in psychologischer Begutachtung u. a. über die Prüfung der Konsistenz bzw. Widerspruchsfreiheit von Aussagen oder Testergebnissen, psychologische Heuristiken zur Beurteilung der Plausibilität von Äußerungen oder Beobachtungen, den Gebrauch spezifischer Kontrollskalen zur Erfassung von Antwortverzerrungen sowie über Testverfahren zur Überprüfung der Testmotivation und Anstrengungsbereitschaft in Leistungssituationen.

Verdeutlichungstendenz. Bezeichnung für eine leichte Übertreibung oder Überhöhung der Beschwerdendarstellung, von der angenommen wird, dass sie überwiegend aus der unbewussten Motivation heraus geschieht, den Untersucher von der Schwere der genannten Probleme oder Beeinträchtigungen zu überzeugen. In der Praxis wird sie abgegrenzt von der willentlichen Übertreibung oder Verfälschung bestehender Beschwerden oder Beeinträchtigungen, um so bestimmte Ziele zu erreichen (Aggravation).

D. Ein einheitlicher Gebrauch testtheoretischer und validierungsdiagnostischer Begriffe schafft konzeptionelle Voraussetzungen für eine sachgerechte Erfassung und Beurteilung mutmaßlich verzerrter Angaben und Untersuchungsergebnisse.

5. *Empfehlungen zur Hypothesenbildung*

Übersetzung rechtlicher in psychologische Fragestellungen

Fragen zur Gültigkeit bzw. Validität eigener Untersuchungsergebnisse bzw. selbst erhobener Informationen sind Bestandteil jedes Gutachtens mit eigener Untersuchung. Fragen oder Hypothesen zur Gültigkeit von Selbstauskünften und zur Gültigkeit individueller Testergebnisse sollen auch dann formuliert werden, wenn sie nicht explizit vom Auftraggeber gestellt werden.

Priorisierung hypothesengeleiteter Diagnostik

Ein hypothesengeleitetes validierungsdiagnostisches Vorgehen zur Beurteilung der Authentizität von Selbstberichten oder Testergebnissen ist einem nicht hypothesengeleiteten Vorgehen in der Regel überlegen.

Validierungsdiagnostische Hypothesen sollten so spezifisch formuliert werden, wie die Erkenntnisse über Rahmenbedingungen, Anreize, situationsbezogene Motive und spezifische diagnostische Risiken dies zulassen. Ungerichtete validierungsdiagnostische Fragen (z. B.: Sind die Angaben zutreffend?) lassen bei der Beantwortung in der Regel mehr Interpretationsspielräume zu als gerichtete, auf spezifische Erwartungen gestützte hypothesengeleitete Fragestellungen (z. B.: Kann die Person mit mindestens 95%-iger Wahrscheinlichkeit der Gruppe aggravierender oder authentisch antwortender Personen zugeordnet werden?).

Abstimmung der Ausgangshypothese auf Anreizbedingungen für Antwortverzerrungen

Die Validierungshypothese sollte auf die Ergebnisinteressen der zu untersuchenden Person abgestimmt werden, sofern diese durch den Untersuchungskontext und das Untersuchungsanliegen bestimmt sind. Wenn die zu untersuchende Person z. B. den Nachweis gesundheitlicher Beeinträchtigungen und damit einhergehender Vorteile anstrebt, sollten sich die Validierungshypothesen auf Aspekte der möglichen Übertreibung gesundheitlicher Beeinträchtigungen beziehen. Strebt die zu untersuchende Person den Nachweis von Gesundheit oder fehlender gesundheitlicher Beeinträchtigungen an, sollten sich die Validierungshypothesen auf Aspekte der möglichen Untertreibung gesundheitlicher Beschwerden oder Beeinträchtigungen beziehen.

Wenn das Untersuchungsanliegen im Einzelfall eindeutig bestimmt ist, ist eine gerichtete Hypothese in der Regel angemessen. Wenn die Ergebnisinteressen der zu untersuchenden Person im Einzelfall nicht eindeutig bestimmt sind, wenn also z. B. nicht klar ist, ob die Person im eigenen Interesse zu negativen oder zu positiven Antwortverzerrungen neigt, ist eine ungerichtete Hypothese angemessen. In diesem

Fall lautet die Fragestellung, ob Antwortverzerrungen oder Hinweise auf eine fehlende Gültigkeit von Aussagen oder Testergebnissen ein kritisches Maß übersteigen.

Bei ereignis- bzw. tatbezogenen Aussagen im Strafverfahren wird im Rahmen der richterlichen Beweiswürdigung von der Nullhypothese ausgegangen, dass die gemachten Angaben *nicht zutreffen*. Im Sinne deduktiver Hypothesenprüfung ist dann zu prüfen, ob die validierungsdiagnostischen Ergebnisse für die Zurückweisung der Nullhypothese sprechen. In der Regel wird die Ausgangshypothese der strafrechtlichen Beweiswürdigung auch von Sachverständigen angewendet.

Validierungsdiagnostische Hypothesen in anderen Rechtsbereichen sollten bei rechtlichen Fragen zum Vorliegen gesundheitlicher Störungen oder Beeinträchtigungen das Risiko gutachterlicher Attributionsfehler in Richtung „(unzutreffender) Bestätigung des subjektiven (i. d. R. interessen geleiteten) Beschwerdenvortrags“ minimieren. Dies kann dadurch erreicht werden, dass ...

1. bezüglich der kognitiven und psychomotorischen Leistungsfähigkeit die Nullhypothese formuliert wird, dass krankheitsbedingte Leistungseinschränkungen – sofern nicht bereits zuvor eine dauerhaft einschränkende Hirnleistungsstörung bei u. a. hirnorganischer Erkrankung nachgewiesen ist – vorübergehend und zum Untersuchungszeitpunkt restituiert beziehungsweise remittiert sind;
2. bezüglich psychischer Beschwerden und gesundheitlicher Beeinträchtigungen die Nullhypothese formuliert wird, dass die Beschwerden und Beeinträchtigungen in der Untersuchungssituation zuverlässig und überwiegend unverzerrt/valide dargestellt werden können, auch wenn die Person im Alltag bzw. in anderen Situationen diese Beeinträchtigungen aufweist bzw. darunter leidet.

Die validierungsdiagnostische Hypothese bei eignungsdiagnostischen Fragen mit erhöhtem Risiko positiv verzerrter Selbstzuschreibungen sollte das Risiko der Bestätigung forciert positiver Selbstzuschreibungen minimieren. Das kann durch die Nullhypothese erreicht werden, dass unter der Voraussetzung hinreichend konsistenter (logisch widerspruchsfreier) Antwortmuster die Neigung zu positiv verzerrten Selbstbeschreibungen (z. B. sozial erwünschtem Antworten, Dissimulation von Beschwerden, Impression-Management-Tendenzen) unauffällig ist bzw. bestimmte validierungsdiagnostische Norm- oder Trennwerte nicht überschreitet.

Abstimmung der Hypothese auf mögliche Konsequenzen einer Fehlentscheidung

Die Festlegung der H0 und der H1 („Wunschhypothese“) hat Auswirkungen auf Aspekte der Überprüfung, der wahrscheinlichkeitsbasierten Bewertung und der integrierenden gutachterlichen Urteilsbildung. Je nach Festlegung kann sich das Risiko der fälschlichen Bestätigung oder fälschlichen Zurückweisung der Nullhypothese bzw. der fälschlichen Akzeptanz oder Zurückweisung der Alternativhypothese verschieben.

Grundsätzlich kann in der Validierungsdiagnostik der positive Nachweis eines vollständig „wahren“, d. h. von Störeinflüssen und unwillentlichen und willentlichen Verzerrungen gänzlich befreiten Untersuchungsergebnisses nicht erreicht werden. So wie z. B. Erinnerungen auch bei stärkstem Bemühen um die Richtigkeit und Authentizität der mentalen Rekonstruktion immer auch ungewollt oder gewollt selektiv, zufällig und individuell assoziativ gebunden sind, so unterliegen auch andere Untersuchungsergebnisse stets multiplen kognitiven und motivationalen Einflüssen, die nicht vollständig kontrolliert werden können. Durch die Kontrolle einzelner Facetten von Antwortverzerrungen oder ausgewählter motivationaler Bedingungen bei der Leistungstestung kann daher immer nur ein Teil potentieller Stör-

faktoren ausgeschlossen (oder belegt) werden. Aus dem Ausschluss einer begrenzten Zahl von Verzerrungen kann aber nicht auf die uneingeschränkte Gültigkeit der erhobenen Informationen gefolgert werden. Die aussagepsychologisch geprägte forensische Praxis hat sich dementsprechend darauf festgelegt, keine Hypothese zum positiven Nachweis der Wahrheit/Gültigkeit (einer Aussage) zu formulieren, sondern *die H0 so zu bestimmen, dass ihre Zurückweisung das Fortbestehen weiterer Zweifel an der Richtigkeit der erhobenen Angaben erlaubt*. Hypothesengeleitet geprüft wird, ob und in welchem Umfang die Untersuchungsergebnisse *dagegensprechen, dass die untersuchte Person die Unwahrheit sagt*. Unter dieser Bedingung führt auch der Ausschluss diverser umschriebener Verzerrungen nicht zu der Schlussfolgerung, eine Aussage als „vollständig wahr/gültig“ zu bewerten. Restzweifel können bei Gericht aber durch den Grundsatz „in dubio pro reo“ so verarbeitet werden, dass eine Entscheidung möglich ist.

Auch in anderen Rechtsgebieten, z. B. dem Sozial- oder Zivilrecht, sind gerichtliche Entscheidungen bei nicht auszuräumenden Zweifeln (z. B. darüber, ob eine bestimmte gesundheitliche Beeinträchtigung besteht oder nicht) erforderlich. Und auch hier ergeben sich hierdurch Konsequenzen für die Hypothesenbildung und -prüfung des psychologischen Sachverständigen, der beispielsweise entsprechend von einer H0 ausgeht, dass keine leistungsmindernden Gesundheitsstörungen vorliegen, und die H1 annimmt, wenn befundgestützt begründbar ist, dass eine leistungsmindernde Gesundheitsstörung vorliegt.

Die Erläuterung und Quantifizierung der dann noch verbleibenden Irrtumswahrscheinlichkeit ("Alphafehler") gegenüber dem Gericht mit dem Ziel, dass dieses hieraus Schlüsse und Erkenntnisse für die rechtliche Bewertung der verbleibenden Zweifel ziehen kann, ist somit rechtsgebietsübergreifend ein zentraler Bestandteil der probabilistisch fundierten gutachtlichen Äußerung.

Ausrichtung der Hypothesenprüfung an statistischen Prognosewerten

Geht man von Prävalenzzahlen für Antwortverzerrungen im Begutachtungskontext aus, kann sowohl die H0 begründet werden, dass Antwortverzerrungen im Einzelfall wahrscheinlich vorliegen, als auch, dass sie im Einzelfall wahrscheinlich nicht vorliegen. Im ersten Fall sollten die zur Überprüfung eingesetzten Validierungsverfahren eine hohe Spezifität aufweisen, um die Wahrscheinlichkeit zu erhöhen, dass Personen, die authentisch/unverzerrt antworten oder mit ausreichender Anstrengungsbereitschaft Leistungstests bearbeiten, auch tatsächlich als valide/authentisch antwortend identifiziert werden (so dass die H0 zurückgewiesen werden kann). Im zweiten Fall sollten die zur Überprüfung der H0 eingesetzten Validierungsverfahren eher eine hohe Sensitivität aufweisen, um die Wahrscheinlichkeit zu erhöhen, dass Personen, die Antwortverzerrungen aufweisen, auch tatsächlich als fraglich valide bzw. fraglich authentisch antwortend identifiziert werden.

Grundsätzlich ist auch eine validierungsdiagnostische Hypothese möglich, die unterstellt, dass das Antwort- oder Reaktionsverhalten in der Untersuchung valide, also unverzerrt ist. In diesem Fall kann sich die Frage stellen, ob aus der Zurückweisung der H0 eine Bestätigung der H1 abgeleitet werden kann. Zu prüfen wäre dann, ob der positive Nachweis gültiger (unverzerrter) Ergebnisse *in einem für das praktische gutachterliche Vorgehen und die konkrete gutachterliche Urteilsbildung ausreichenden Umfang* gelingen kann. So könnte die o. g. Einschätzung, dass der positive Nachweis gültiger, d. h. vollständig unverzerrter Angaben oder Ergebnisse nicht möglich ist, an praktische Erfordernisse und Entscheidungsbedingungen angepasst werden.

Entscheidbarkeit der Hypothese(n) sichern

Hypothesen sollten so formuliert werden, dass sie anhand empirisch ermittelter Kriterien gestützt oder widerlegt werden können. Bei standardisierten validierungsdiagnostischen Verfahren ist zu diesem Zweck eine *Spezifizierung prüfbarer Facetten generalisierter Antworttendenzen* sinnvoll. Hypothesen

sollten die Facette verzerrenden Antwortverhaltens möglichst konkret benennen. Prüfbare Facetten invaliden Antwortverhaltens können z. B. sein: Tendenz zu inkonsistentem Antworten, Tendenz zu Extremantworten, Tendenz zur Bejahung unwahrscheinlicher Sachverhalte, Tendenz zur Leugnung bzw. Verneinung wahrscheinlich zutreffender Sachverhalte, Tendenz zur Ausrichtung an Stereotypen, Tendenz zur Ausrichtung an vermuteten Erwartungen usw.

Die Ausgestaltung der Hypothesen sollte *an die Qualität des Datenniveaus angepasst* werden. In der qualitativ orientierten Aussagepsychologie fehlen die Voraussetzungen für eine psychometrische Bewertung des Antwortmusters. Quantitative psychometrische Diagnostik bietet hingegen Voraussetzungen, um z. B. intraindividuelle Merkmalsdifferenzen gegen Zufallsschwankungen oder andere Störeinflüsse im Einzelfall abzugrenzen und so Inkonsistenzen oder mangelnde psychologische Plausibilität mit Bezug auf statistische Kriterien (Effektgrößen, Signifikanzniveau) abzusichern.

E. Validierungsdiagnostische Fragestellungen und Hypothesen sollten aus den rechtlichen Fragestellungen abgeleitet werden, wenn eine Übersetzung der rechtlichen in psychologische Fragen erforderlich ist. Die Wahl der H0 und der H1 bei Validierungsentscheidungen und die Festlegung von Maßnahmen zu ihrer Bestätigung bzw. Widerlegung kann von verschiedenen Bedingungen, darunter auch den möglichen Konsequenzen einer Fehlentscheidung, abhängig gemacht werden. So können rechtliche Rahmenbedingungen (z. B. Schutzzweck des Sozialrechts), aber auch Prävalenzen für invalide Ergebnisse und Verfälschungsrisiken oder individuelle Faktoren (z. B. Schutzbedürftigkeit bzw. Hilfebedarf des Einzelnen) maßgeblich für die Festlegung der validierungsdiagnostischen Hypothesen sein.

6. Validierungsdiagnostische Untersuchungsplanung

Die folgenden Empfehlungen betreffen die validierungsdiagnostische Untersuchungsplanung einschließlich der darauf abgestimmten Methodenauswahl.

Validierungsorientierte Aktenanalyse

- Vorinformationen, die für die Validierung eigener Befunde oder für die Interpretation selbst erhobener oder auch fremder Befunde genutzt werden können, sollten den vorliegenden Quellen entnommen werden. Dies können z. B. psychometrische Werte klinischer Fragebögen oder anderer psychologischer Testverfahren sein, die zu den selbst erhobenen Messwerten zum Konsistenzabgleich in Beziehung gesetzt werden. Konsistenzabgleiche von Informationen zu gleichen Sachverhalten sind ebenso möglich für bereits vorliegende Aussagen über relevante Ereignisse, Zustände oder Erfahrungen, sofern sich deren Darstellung (Beurteilung) allein durch die Zeit zwischen den Erhebungen bzw. durch die unterschiedlichen Erhebungskontexte nicht verändert haben sollte. In diesem Fall sind konsistente, also übereinstimmende Angaben zu erwarten.
- Erhoben werden sollten ebenfalls Informationen zu den Ergebnissen validierungsdiagnostischer Verfahren (Performanz-Validierungstests, Kontrollskalen) aus früheren diagnostischen Untersuchungen sowie zusammenfassende Bewertungen oder befundliche Sicherungen, die validierungsdiagnostische Aspekte berücksichtigen (z. B. psychischer Befund im psychiatrischen oder psychotherapeutischen Kontext, Bewertungen der Glaubhaftigkeit von Aussagen oder der Glaubwürdigkeit).
- Bei Nennung validierungsdiagnostischer Bewertungen bzw. Befunde sollte die Person und der

Kontext mit aufgenommen werden, in dem die Bewertung vorgenommen wurde (z. B. Behandlungskontext, gutachterlicher Kontext).

- Schließlich sollte insbesondere Vorgutachten entnommen werden, ob und inwiefern Angaben zur gezielten Erhebung und Bewertung validierungsdiagnostischer Informationen dort enthalten sind (z. B. eindrucksbasierte oder testbasierte Validierung, kriterienbasierte Beurteilung).
- Bei Fragen zum *Krankheits- oder Behandlungsverlauf* sollten der Akte nach Möglichkeit die folgenden Informationen entnommen und in Bezug auf die Zuverlässigkeit der Beschwerdenschilderung unter motivationalen und sozial-kommunikativen Aspekten ausgewertet werden:
 - mögliche Ursachen der beklagten Beschwerden,
 - früher bestehende bzw. nachgewiesene psychische Beschwerden oder Erkrankungen/Störungen,
 - zwischenzeitlich erlangte psychische oder kognitive Restitution bzw. erhöhte Leistungsfähigkeit,
 - motivationale Aspekte im Behandlungsverlauf.

Übersetzung rechtlicher in psychologische und validierungsdiagnostische Fragen /Hypothesen

Die Untersuchungsplanung baut insofern auf der Hypothesenbildung auf, als die Erhebungsmethoden (Exploration, Verhaltensbeobachtung, testpsychologische Diagnostik, Validierungsdiagnostik u. A.) nach Möglichkeit spezifisch auf die Überprüfung der ungerichteten oder gerichteten Validierungshypothesen auszurichten sind.

Die validierungsdiagnostische Hypothese kann u. a. folgende Sachverhalte betreffen:

- Aussagefähigkeit (bezogen auf Sprachverständnis, Gedächtnisfähigkeit, Urteilsfähigkeit)
- Glaubhaftigkeit frei explorierter ereignisbezogener Aussagen (Abgrenzung gegen mangelnde Aussagentüchtigkeit und Erlebnisnähe)
- Glaubhaftigkeit frei und standardisiert erhobener Informationen zu stabilen Eigenschaften oder Merkmalszuschreibungen (Konsistenz- und Plausibilitätsprüfungen)
- Gültigkeit von Testleistungen und Fähigkeitszuschreibungen
- Ausprägung formaler Antworttendenzen, Neigung zu zufälligem Antworten
- Ausprägung spezifischer inhaltlicher Antworttendenzen (spezifischer Antwortmotive)

Festlegung der zu validierenden Sachverhalte oder Eigenschaften

Nicht für alle Merkmale oder Eigenschaften kann und sollte ein gleichermaßen hoher Validierungsaufwand betrieben werden. Der Validierungsaufwand kann umso höher veranschlagt werden, je wichtiger ein Sachverhalt, ein Merkmal oder eine Eigenschaft für die Beantwortung der gutachterlichen Fragestellung ist.

Dies können bei der *Beurteilung von Ereignissen* (z. B. bei Straftaten, traumatischen Erfahrungen, spezifischen Anforderungen und darauf bezogenem Bewältigungsverhalten) Merkmale des Ereignisses selbst, der äußeren und inneren Bedingungen vor und während des Ereignisses und der nachfolgenden Ereignisverarbeitung sein.

Bei der *Beurteilung relativ stabiler Eigenschaften* und Fähigkeiten kann die Ausprägung der Eigenschaft, ihre zeitliche Konstanz und ihre Abhängigkeit von äußeren und inneren Bedingungen bevorzugter Gegenstand der Validierung sein. Bei gesundheitlichen Problemen können das Erscheinungsbild und die Schwere der Erkrankung oder psychischen Störung oder Merkmale des individuellen Bewältigungsstils (subjektive Krankheitstheorie, Einstellung und Compliance in der Rehabilitation oder Psychotherapie, fördernde und hemmende intrapsychische Wirkfaktoren im Gesundungsprozess oder ggf. bei der Chronifizierung von Beschwerden) im Fokus der Validierung stehen.

Auswahl geeigneter (nicht explizit validierungsdiagnostischer) Mess- und Testverfahren

Psychometrische Mess- und Testverfahren sollten sowohl nach allgemeinen Testgütekriterien ausgewählt werden, als auch nach ihrem Wert für die Beantwortung der gutachterlichen Fragestellungen. Die Auswahl sollte auch danach vorgenommen werden, dass eine einseitig auf die Bestätigung oder auf die Widerlegung der Untersuchungshypothese ausgerichtete Datenerhebung vermieden wird (z. B. einseitige Ausrichtung der Datenerhebung auf gesundheitliche Probleme). Insbesondere sollten Verfahren mit erhöhtem Risiko für messreaktive Verzerrungen vermieden werden (z. B. reine Symptomlisten bei Fragen zu gesundheitlichen Einschränkungen).

Eine *messreaktive Verzerrung* liegt vor, wenn das Untersuchungsergebnis vorwiegend als das Ergebnis der spezifischen Wechselwirkung der untersuchten Person mit den Besonderheiten der Erhebungsmethode zu interpretieren ist. Je stärker diese Wechselwirkung ist, umso weniger eindeutig kann von dem Testergebnis auf die Ausprägung der gemessenen Eigenschaft in anderen (Test-)Situationen geschlossen werden. Entsprechend ist der explikative und prognostische Wert messreaktiv verzerrter Testergebnisse gering zu veranschlagen.

Kombination von Verhaltens- und Persönlichkeitsdiagnostik

Eine multimethodal ausgerichtete psychologische Diagnostik sollte auf die Erhebung und *Verknüpfung verhaltensnaher, d. h. durch unmittelbare Beobachtung oder freie Exploration erhobener Information einerseits und eigenschaftsbezogener Information mittels standardisierter Methoden andererseits ausgerichtet* sein. Das gilt auch, wenn eine bestimmte Informationsquelle (z. B. Exploration) bei der Datenerhebung überwiegt.

Der Vorteil verhaltensnah (durch Beobachtung oder Befragung) erhobener Informationen (Verhaltensdiagnostik) besteht in der Möglichkeit, Verhalten (einschließlich verbaler Äußerungen) in seinen spezifischen situativen Kontexten zu beschreiben und zu interpretieren.

Persönlichkeitsdiagnostik gibt Aufschluss über Verhaltensdispositionen und sollte in aller Regel auch mit standardisierten Messmethoden durchgeführt werden. Sie sollte so erfolgen, dass die erhobenen Eigenschaften Beiträge zur Interpretation des Untersuchungsverhaltens sowie zum Bezugssystem der untersuchten Person bei der Befragung liefern können.

Die Ergebnisse der Persönlichkeitsdiagnostik sollten anlassunabhängig dazu genutzt werden, den Verständnishintergrund und relevante Facetten der Handlungssteuerung der Person in der Untersuchungs- bzw. Befragungssituation angemessen einschätzen zu können.

Auswahl validierungsdiagnostischer Methoden

Validierungsdiagnostik sollte Informationen unterschiedlicher Erhebungsmethoden einschließen. Die Auswahl der Variablen sollte so erfolgen, dass *die verschiedenen Methoden wechselseitig zur Validierung der Ergebnisse genutzt werden können*.

Im Fokus der Exploration sollten v. a. Informationen stehen, die in anderer Form (z. B. durch standardisierte Befragung oder Checklisten oder Aktenanalyse) nicht gewonnen werden können. Explorationsinhalte können aber auch dadurch bestimmt sein, dass sie in Kombination mit standardisiert erhobenen Konstrukten zur Validierung von Eigenschaften oder Verhaltensweisen genutzt werden können.

Geeignete Validierungskennwerte können bei standardisierter Messung in Mess- oder Testverfahren integriert sein („eingebettete Validierungsparameter“) in Leistungstests, Kontrollskalen in Fragebögen). Möglich ist auch der Einsatz von Testverfahren, die ausschließlich zur Validierungsdiagnostik genutzt werden.

Für die Auswahl validierungsgerechter Instrumente gelten die folgenden Empfehlungen:

- Wenn *nur Konsistenzaspekte zu prüfen sind*, können dazu die freie Exploration sowie standardisierte Verfahren ohne expliziten Validierungszweck verwendet werden wie z. B. klinische

Verfahren, Persönlichkeitstests ohne Kontrollskalen oder psychologische Leistungstests, die geeignet sind, Antwortverhalten oder Reaktionsmuster gegen Zufälle, Messfehler (Reliabilitätseinschränkungen) und inhaltliche Mehrdeutigkeiten (Validitätseinschränkungen) zu sichern. Strukturierte Interviews können in Ergänzung zu freier Exploration ebenso dazu verwendet werden, die Konsistenz von Äußerungen zu prüfen und situative Einflüsse bei der Datenerhebung und Datenauswertung (Objektivitätseinschränkungen) zu kontrollieren.

- Wenn *Verhaltensbeobachtungen in die Validierungsdiagnostik einbezogen werden*, sollte ein strukturiertes Vorgehen mit festgelegten Beobachtungskategorien nach Möglichkeit mit freier Verhaltensbeobachtung kombiniert werden.
- Wenn die *Datenerhebung auf freie Exploration beschränkt* ist, sollte die Exploration ausgezeichnet werden, um eine nachträgliche sorgfältige Analyse des Gesprächsverhaltens zu sichern. Die Analyse sollte sich bei ereignisbezogenen Aussagen an aussagepsychologischen Kriterien orientieren.
- Wenn die Datenerhebung eine *standardisierte Befragung mittels Fragebogen* vorsieht, sollten standardisierte und normierte Fragebögen mit Skalen zur Erfassung generalisierter und spezifischer inhaltlicher Antworttendenzen verwendet werden. In der Regel sind dazu mehrdimensionale Persönlichkeitsinventare mit mehreren Validierungsskalen geeignet.
- Wenn die Datenerhebung eine *standardisierte Leistungsmessung* mit psychologischen Leistungstests vorsieht, sollten mehrere geeignete Parameter zur Erfassung der Anstrengungsbereitschaft bzw. der motivationalen Voraussetzungen für die sachgerechte Interpretation der Testleistung (Performanz-Validierungstest, „eingebettete Validierungsparameter“ oder testimmanente Kontrollparameter) verwendet werden.
- Grundsätzlich sollte sich die *Auswahl der Methoden an den situativ gegebenen Verfälschungs- oder Verzerrungsrisiken orientieren*. Diese sind bestimmt durch die Interessen der untersuchten Person am Untersuchungsergebnis und die Wahrscheinlichkeit, mit der die Person ihr Untersuchungsverhalten auf das gewünschte Untersuchungsergebnis abstimmt.
- Bei *erhöhtem Risiko für positive Antwortverzerrungen bei Selbstauskünften* (z. B. sozial erwünschtes Antworten, Leugnung bzw. Dissimulation von Beschwerden, „Impression Management“) sollten validierungsdiagnostische Kontrollskalen oder Testverfahren so ausgewählt werden, dass die Wirkung positiver Antwortverzerrungen auf die zu beurteilenden Sachverhalte zuverlässig beurteilt werden kann.
- Bei *erhöhtem Risiko für negative Antwortverzerrungen bei Selbstauskünften* (z. B. von der zu beurteilenden Person gewünschter Nachweis über gesundheitliche Probleme oder Funktionseinschränkungen) sollten validierungsdiagnostische Kontrollskalen so ausgewählt werden, dass die Wirkung negativer Verzerrungen zuverlässig beurteilt werden kann. Voraussetzung dafür sind in der Regel Fragebogen-Kennwerte zu verschiedenen Facetten negativer Antwortverzerrungen.

Die folgenden *mess- und testdiagnostischen Verfahrensweisen* können für Validierungszwecke geeignet sein:

- Profilinterpretation multifaktorieller Testverfahren zur Abschätzung der Plausibilität intraindividuellere Antwortmuster;
- Zufalls- und valenzkritische Prüfung intraindividuellere Merkmalsdifferenzen, bezogen auf Merkmalsabweichungen konvergent valider Kennwerte;
- Abgleich verschiedener Datenquellen bezogen auf bestimmte Zielgrößen (Selbstbericht, psychophysiologische Daten, Verhaltensbeobachtung, Leistungstest);

- Standardisierte Erhebungsmethoden mit mehrdimensionalen integrierten Kontrollskalen zu Antwortverzerrungen (z. B. MMPI-2 (Hathaway & McKinley; deutsche Version von Engel, 2000), MMPI-2-RF (Ben-Porath & Tellegen; deutsche Version von Engel, 2019), Verhaltens- und Erlebensinventar (VEI; Engel & Groves, 2012).
- Strukturierte Interviews oder standardisierte Fragebogen zur Beschwerdenuvalidierung (z. B. deutsche Version des SIRS 2 von Schmidt, Watzke, Lanquillon & Stieglitz, 2019). Andere Verfahren können im Sinne eines Screenings störungsbildübergreifend sein, bspw. Deutsche Version des Self-Report-Symptom-Inventory (SRSI; Merten, Giger, Merckelbach & Stevens, 2019; Beschwerdenuvalidierungstest (BEVA; Walter, Petermann & Kobelt, 2016). Dazu zählen auch störungsspezifische Validierungsinstrumente.
- Performanzvalidierungstests, die dem Anschein nach zur Leistungserfassung konzipiert sind, tatsächlich aber die Testmotivation oder Anstrengungsbereitschaft eines Probanden in einer Leistungstestung untersuchen.

F. Die validierungsdiagnostische Untersuchungsplanung geht von einer differenzierten Aufnahme und Reflexion der bereits vorliegenden Informationen (v. a. Aktenanalyse) mit darauf abgestimmter Festlegung psychologischer Fragestellungen aus und leitet daraus die Planung einer multimethodalen Datenerhebung und Auswahl validierungsdiagnostischer und anderer (inhaltlicher) Erhebungsmethoden ab.

7. Durchführung der Untersuchung

Die Untersuchung sollte so durchgeführt werden, dass die Gültigkeit der für die Beantwortung der gutachterlichen Fragestellungen zentralen Angaben, Eigenschaften oder Fähigkeiten mehrfach diagnostisch gesichert wird (z. B. durch Exploration, Verhaltensbeobachtung, Selbstbericht, Kontrollskalen, Fremdbbericht, physiologische Parameter, psychologische Tests, Performanzvalidierungstest). Dazu gehört auch der Abgleich verschiedener Datenebenen (z. B. physiologisch-motorisch-kognitiv-emotional-motivational-sozial).

Für die psychologische Validierung einzelfallbezogener Befunde ist der *Einbezug testpsychologischer Erhebungsmethoden* in die Durchführung der Untersuchung/Datenerhebung *obligatorisch*.

Vorschläge für das praktische gutachterliche Vorgehen

- **Gestaltung des Untersuchungsraums.** Die Untersuchungsbedingungen sollten so gestaltet sein, dass sie den Untersuchten keine Hinweise auf Art und Zeitpunkt validierungsdiagnostischer Aktivitäten im Untersuchungsverlauf liefern.
- **Sicherung der Identität.** Zu Beginn der Untersuchung sollten die Identität der zu untersuchenden Person und weitere relevante Personenvariablen erhoben werden.
- **Informierte Einwilligung**
Es sollte eine schriftliche informierte Einwilligung des Probanden in die Teilnahme an der Untersuchung und die damit verbundenen Anforderungen eingeholt werden. Die Einwilligung informiert auch darüber, dass Testverfahren in die Validierung der Ergebnisse einbezogen werden, ohne aber die Verfahren oder deren Eigenschaften konkret und im Detail zu benennen. Sie

enthält auch Informationen dazu, wie bei Hinweisen auf eingeschränkte Kooperativität verfahren wird. Gleiches gilt bei gesetzlich betreuten Personen in Bezug auf ihre Bezugsperson. Die Regeln können einen vorzeitigen Abbruch bei belastbaren Hinweisen auf fehlende Einsicht in die Ziele der Untersuchung und/oder fehlende Anstrengungsbereitschaft der zu Begutachtenden, bei stark oppositionellem Untersuchungsverhalten oder bei Coaching durch einen Rechtsbeistand beinhalten.

- **Rechtzeitige Sicherung validierungsdiagnostischer Daten**

Validierungsinformationen (z. B. zur „Aussagetüchtigkeit“ mit Abgrenzung von Antworten gegen den Zufall und unsystematische Störeinflüsse, zur Identifikation eines generalisierten oder themenspezifischen Antwort- oder Reaktionsmusters) sollten im gesamten Untersuchungsverlauf und nicht erst gegen Ende der Untersuchung erhoben werden, um einen Verlust von Validierungsinformation durch vorzeitigen Abbruch oder ermüdungsbedingte Verzerrungen zu vermeiden.

- **Abstimmung der Informationserhebung auf die zu beantwortenden Fragestellungen.**

Hypothesenkonforme und hypothesendiskrepante Fragestellungen oder Testanforderungen sollten in ausgewogenem Verhältnis zur Beantwortung bzw. Bearbeitung vorgegeben werden (z. B. Fragen zu erlebnisbasierten vs. nicht erlebnisbasierten Erinnerungen oder Erfahrungen, zu intakten vs. gestörten Funktionen, zu Krankheit und Gesundheit usw.).

- **Ergebnisneutrale Erhebung von Selbstberichten**

Die Informationserhebung sollte weitgehend hypothesenunabhängig so gestaltet sein, dass die Erhebung von Einschränkungen, Defiziten, Misserfolgen, Mängeln oder „negativen Aspekten“ in angemessenem Verhältnis zur Erhebung von normal ausgeprägten oder intakten Eigenschaften und Funktionen, von Erfolgen, erhaltenen Fähigkeiten oder „positiven Aspekten“ steht. Bei Untersuchungen zum gewünschten Nachweis krankheitsbedingter Funktions- und Leistungseinschränkungen sollten z. B. neben gezielten Fragen zu Beschwerden und Beeinträchtigungen auch ergebnisoffene, neutrale Formulierungen verwendet werden sowie gezielte Fragen zu Ressourcen, Therapieerfolgen oder vorhandenen Kompetenzen.

- **Empfehlungen zur Verhaltensbeobachtung**

Die Untersuchungsplanung sollte Verfahrensweisen für eine kontinuierliche freie und – für umschriebene Situationen oder Situationstypen (z. B. während standardisierter Testungen) - nach Möglichkeit auch strukturierte Beobachtung des Untersuchungsverhaltens vorsehen. Bei strukturierter Beobachtung sollten Beobachtungskategorien vorab festgelegt und geeignete Dokumentationssysteme bereitgestellt werden. Verhaltensbeobachtungen mit direktem Validierungsbezug sollten kontextbezogen dokumentiert werden, z. B. Merkmale der Situation – beobachtetes Verhalten – erhobene Vergleiche zur Beurteilung der Verhaltenskonsistenz – erhobene Vergleich zur Beurteilung der Plausibilität.

- **Erhebung von Informationen zu Mindestvoraussetzungen des Antwortverhaltens und zu willentlich verzerrten Angaben**

Die validierungsdiagnostische Datenerhebung sollte sowohl Aspekte der „Berichts-Tüchtigkeit“ bzw. der kognitiven Mindestvoraussetzungen für die Ergebnisinterpretation, als auch Merkmale willentlicher bzw. kontrollierter Antwortverzerrungen umfassen. Das heißt, dass Validierungskennwerte und zur Validierung verwendete Beobachtungen/Informationen möglichst spezifisch auf die validierungsdiagnostische Hypothese zur Aussagefähigkeit bzw. Berichts-

tüchtigkeit, Reliabilität und zum Vorliegen oder Ausschluss generalisierter inhaltlicher (positiver oder negativer) Antwortverzerrungen abgestimmt werden sollten. Informationen zur Erfassung von Zufallsverhalten oder mangelnder Informationsverarbeitung (mangelnde Reliabilität, „Aussagetüchtigkeit“) können u. a. über spezifische medizinische Befunde, Verhaltensbeobachtung und Kennwerte zu formalen Antworttendenzen erhoben werden. Informationen zu willentlichen Antwortverzerrungen oder mangelnder Anstrengungsbereitschaft in Funktions- oder Leistungstests sollten durch multimethodalen Abgleich unter Einbezug freier und standardisierter validierungsdiagnostischer Methoden gesichert werden.

G. Die Durchführung der Untersuchung sollte auf die Umsetzung validierungsdiagnostischer Maßnahmen und den Erhalt ausreichender, multimethodal erhobener und geeigneter validierungsdiagnostischer Informationen ausgerichtet sein. Dabei ist zwischen Informationen zu Mindestvoraussetzungen für die Ergebnisinterpretation (z. B. im Sinne von Aussagetüchtigkeit, logischer Konsistenz, formaler Antworttendenz) und Informationen zu inhaltlichen interessengeleiteten Verzerrungen zu unterscheiden.

8. Empfehlungen zur validierungsdiagnostischen Urteilsbildung

Validierung vor inhaltlicher Interpretation

Die Validierung der erhobenen Informationen (Angaben, Testergebnisse) geht der weiteren inhaltlichen Interpretation von verbalen Angaben, Verhaltensbeobachtungen oder psychologischen Untersuchungsergebnissen voraus.

Relevanz der zu validierenden Informationen und Validierungsaufwand

Die validierungsdiagnostische Befundsicherung sollte umso differenzierter und fundierter sein, je aussagekräftiger die zu validierenden Informationen für die inhaltliche Beantwortung der Fragestellung bzw. für die abschließende gutachterliche Stellungnahme sind.

Zweistufiges validierungsdiagnostisches Vorgehen bei Selbstberichten

Verbale Berichte über erlebte Ereignisse, Erfahrungen, Wahrnehmungen, Erinnerungen, Überzeugungen, Zuschreibungen oder andere kognizierte Sachverhalte können von mindestens zwei Komponenten verzerrt bzw. verfälscht sein:

- a) Der *Fähigkeit* der Person, den erfragten Sachverhalt zutreffend mental zu verarbeiten und entsprechend zu beschreiben („Kognitive Mindestvoraussetzungen“). Gemeint ist eine für die Untersuchung ausreichende Fähigkeit zur Informationsverarbeitung und angemessenen Reaktion, gemessen an Merkmalen der Wahrnehmung, der Orientierung und Aufmerksamkeit, des Arbeitsgedächtnisses und sonstiger Gedächtnisfunktionen, des logischen Denkens sowie sprachlicher Funktionen und des Ausdrucksverhaltens.
- b) Der *Bereitschaft bzw. Neigung* der Person, den erfragten Sachverhalt zutreffend beschreiben oder darstellen zu wollen („Motivationale Bedingungen“). Der Nachweis der Bereitschaft erfordert den Ausschluss willentlich ergebnisverzerrender Antworttendenzen oder den positiven Nachweis der psychodiagnostischen Güte (Zuverlässigkeit, Gültigkeit) des Antwort- oder Reaktionsverhaltens.

Sowohl die Fähigkeit, erfragte Sachverhalte zutreffend zu beschreiben, als auch die Bereitschaft, dieses unter den gegebenen äußeren und inneren Anreizbedingungen auch zu tun, können unkontrollierten bzw. unwillkürlichen ebenso wie kontrollierten bzw. willkürlichen Einflüssen unterliegen.

Dazu im Einzelnen:

zu a) Beurteilung der *kognitiven Mindestvoraussetzungen für die Ergebnisinterpretation*

Folgende Indikatoren des Untersuchungsverhaltens können auf das **Fehlen der Mindestvoraussetzungen für die Ergebnisinterpretation** (unzureichende „Aussagetüchtigkeit“ in der Exploration) hinweisen:

- Verhaltensbeobachtung: mangelnde Orientierung zu Zeit, Ort, und zur eigenen Person; Hinweise auf mangelnde Wahrnehmung, gestörte Aufmerksamkeit, gestörte Gedächtnisfunktionen, unzureichende Fähigkeit zu logischem Denken, mangelnde sprachliche Fähigkeiten oder instruktionswidriges Verhalten.
- Exploration: Hinweise auf zufälliges oder thematisch ungeordnetes oder logisch inkonsistentes Antworten.
- Standardisierte Fragebogendiagnostik: Hinweise auf überdurchschnittlich/auffällig hohe Inkonsistenz in Kontrollskalen zur logischen Antwortkonsistenz.
- Performanz-Validierungstest: Hinweise auf zufälliges Antwort-/Reaktionsmuster.

Folgende **klinischen Befunde mit Erklärungswert** können auf unzureichende Mindestvoraussetzungen der Ergebnisinterpretation hinweisen:

- Nachweise über hirnorganische Schädigungen oder Erkrankungen, die typischerweise zu Einschränkungen der Aufmerksamkeits- und Konzentrationsleistung, der Wahrnehmungs- und Gedächtnisleistung oder der Exekutivleistung führen.
- Nachweise über andere organische Schädigungen oder Erkrankungen, die dauerhaft und situationsübergreifend die kognitive Leistungsfähigkeit oder Facetten der Verarbeitungsfähigkeit einschränken (z. B. erhebliche Schmerzen).
- Nachweise über bestehende Intelligenzminderungen oder Entwicklungsverzögerungen.

Die genannten Befunde können sich auf folgende Quellen der Prüfung stützen:

- Anhaltspunkte aus der Aktenlage zu o. g. gesundheitlichen Schädigungen, Erkrankungen oder psychischen Störungen unter Berücksichtigung der dortigen Validierung, ggf. Anforderung von Berichten und Testprotokollen, wenn Vorberichte nicht ausreichend transparent dargelegt sind.
- Anhaltspunkte aus einer aktuellen Einschätzung der allgemeinen und sprachlichen Intelligenz bzw. der Einsichtsfähigkeit, aus der Verhaltensbeobachtung.
- Eigene testpsychologische Untersuchung, wenn keine ausreichende Informationsgrundlage gegeben ist.
- Einschätzung der Persönlichkeit der klagenden Person und möglicher allgemeiner, inhaltsunabhängiger Antworttendenzen.

Für die **Urteilsbildung** gelten die folgenden Empfehlungen:

Begrenzter Vorhersagewert klinischer Diagnosen für Mindestvoraussetzungen

Das Vorliegen bzw. der gesicherte Vorbefund einer psychischen Störung begründet in der Regel keinen

Ausschluss der Mindestvoraussetzungen für die Interpretation von Selbstauskünften. Auch schwer psychisch gestörte Personen (z. B. wahnhafte Symptomatik, schwere Zwangsstörung oder depressive Störung, Drogenabhängigkeit, Borderline-Störung) sind nicht automatisch in ihrem Sprachverständnis, ihrer Aufmerksamkeit, ihrer Gedächtnisfähigkeit und ihrer Denkfähigkeit so stark beeinträchtigt, dass sie nicht in der Lage wären, in einer zeitlich begrenzten und auf ihre Eigenschaften und Fähigkeiten abgestimmten Situation angemessen zu reagieren.

zu b) Beurteilung der *motivationalen Bedingungen des Antwortverhaltens*

Für die **validierungsdiagnostische Urteilsbildung** gelten die folgenden Empfehlungen:

Validierung als Ergebnis multimethodaler Analyse

Die Entscheidung, ob und in welchem Ausmaß eine verzerrte Darstellung von Beschwerden, Funktionen bzw. Funktionsbeeinträchtigungen sowie beschwerden- oder funktionsmodulierenden Eigenschaften vorliegt, soll bei standardisierten Verfahren mit Bezug auf Normierungsstichproben, interindividuelle Vergleiche statistischer Kennwerte, Ausprägung und/oder Vorhersagevalidität von Validierungskennwerten sowie durch den systematischen Abgleich von Informationen verschiedener nichtstandardisierter Erhebungsmethoden (Vorberichte, Exploration und Erstellung des psychologischen Befunds, Verhaltensbeobachtung, ggf. physiologischer Messungen) erfolgen.

Hinweise auf willentlich kontrollierte Antwortverzerrungen

Folgende Merkmale können Anhaltspunkte oder ergänzende Hinweise auf willentliche Verzerrungen (unzureichende Gültigkeit/Validität der Angaben oder des Untersuchungsverhaltens) liefern:

- Verhaltensbeobachtung: auffällig stark an den Befragungskontext oder die Befragungsbedingungen angepasstes und dadurch inkonsistentes Reaktionsverhalten.
- Exploration: Hinweise auf auffällig stark an die Ergebnisinteressen der untersuchten Person und dadurch inkonsistentes oder psychologisch nicht plausibles Antwortverhalten.
- Standardisierte Fragebogendiagnostik: Zufalls- und valenzkritische intraindividuelle Differenzen zwischen Kennwerten konvergent valider Eigenschaften, sofern diese einen Bezug zu den Ergebnisinteressen der untersuchten Person aufweisen.
- Standardisierte Validierungsdiagnostik/Kontrollskalen zur Erfassung generalisierter oder spezifischer inhaltlicher Antworttendenzen, z. B. einer forcierten ...
 - Neigung zu positiver / idealisierter / sozial erwünschter Selbstdarstellung
 - Neigung zu unwahrscheinlichen Angaben
 - Neigung zur Leugnung/Verneinung zutreffender Angaben
 - Neigung zur Über- oder Untertreibung störungsdisponierender Eigenschaften
 - Neigung zur Über- oder Untertreibung aktueller Belastungsfaktoren
 - Neigung zur Angabe bizarrer/seltsamer Beschwerden oder Symptomkombinationen
 - Neigung zur verzerrten Darstellung des prämorbidem Funktionsniveaus
 - Neigung zur Ausrichtung an störungs-bzw. krankheitsbezogenen Stereotypen
 - Neigung zur Ausrichtung des Verhaltens am Untersucher
 - Neigung zur Ausrichtung des Antwortverhaltens an Merkmalen der Krankenrolle
- Performanz-Validierungstest: Hinweise auf Antwort-/Reaktionsmuster, die für willentlich kontrollierte mangelnde Anstrengungsbereitschaft in der Testsituation sprechen.
- Methodenintegrierende Hinweise: Erhebliche Differenzen zwischen Selbstberichten, Verhaltensbeobachtung, Fragebogendiagnostik, Leistungstest und ggf. psychophysiologischer Messung.

Mindestvoraussetzungen vor willentlichen Antwortverzerrungen

Die Beurteilung der Mindestvoraussetzungen für verwertbare Mess- oder Testergebnisse („Aussagetüchtigkeit“) geht der Beurteilung willentlicher Antwortverzerrungen voraus. Bei deutlichen Hinweisen auf eine unzureichende kognitive Verarbeitungskapazität oder mangelnde Fähigkeit, den Untersuchungsanforderungen zu entsprechen (erkennbar an ausgeprägten formalen Antworttendenzen oder an zufälligem Antworten), sollte nicht auf willentliche Antwortverzerrungen (Aggravation, Simulation, Dissimulation) geschlossen werden.

Auch Hinweise auf zielgerichtete, willentlich kontrollierte Antwortverzerrungen sollten bei Hinweisen auf oder Belegen für unzureichende Mindestvoraussetzungen zurückhaltend oder gar nicht als solche (gültige Belege für willentliche Verzerrungen) interpretiert werden.

Begrenzter Vorhersagewert klinischer Diagnosen für willentlich verzerrte Angaben

Der gesicherte Vorbefund einer psychischen Störung begründet in der Regel keinen Ausschluss willentlicher Antwort- oder Reaktionsverzerrungen in der Untersuchungssituation. Umgekehrt begründet die fehlende befundliche Sicherung einer psychischen Störung bei Klagen über gesundheitliche Beschwerden oder Beeinträchtigungen nicht, dass diese durch willentliche Antwort- oder Reaktionsverzerrungen erklärt werden können oder müssen.

Abgrenzung authentischer und willentlich verzerrter Selbstberichte

Wenn gesundheitliche Beeinträchtigungen des Erlebens und Verhaltens, die der willentlichen Steuerung entzogen sind (psychische Störungen), befundlich gesichert sind, liegt der Fokus der Beurteilung auf der Abgrenzung mutmaßlich authentischer und mutmaßlich nicht authentischer Beschwerden oder Beeinträchtigungen. Anhaltspunkte für die Abgrenzung können sein:

- Art und Ausmaß der Vereinbarkeit bzw. Unvereinbarkeit der Klagen über psychische Störungen mit klinischen Syndrom-Merkmalen
- Art und Ausmaß der Ausrichtung der Klagen über psychische Störungen an stereotypen Krankheitsmerkmalen
- Art und Ausmaß der Ausrichtung der Klagen über Merkmale der Krankheitsverarbeitung an stereotypen Merkmalen der Krankenrolle
- Art und Ausmaß der Diskrepanzen zwischen Selbst- und Fremdberichten über Merkmale von Gesundheit, Krankheit und Leistungsfähigkeit
- Art und Ausmaß an Diskrepanzen zwischen der Schwere krankheitsbedingter Beeinträchtigungen und Merkmalen der Therapie bzw. Krankheitsverarbeitung
- Nennung von Gesundheits- oder Krankheitsmerkmalen oder von Merkmalen der Krankheitsverarbeitung, die den Ergebnisinteressen der untersuchten Person widersprechen
- Art und Ausmaß der Angabe unwahrscheinlicher, bizarrer oder psychologisch nicht plausibler Sachverhalte
- Art und Ausmaß der Verneinung oder Leugnung wahrscheinlicher oder psychologisch plausibler Sachverhalte.

Das für die Abgrenzung authentischer und willentlich verzerrter Angaben relevante Risiko für willentlich verzerrte Selbstberichte kann anhand der Schwere der befundlich gesicherten psychischen Störung oder Erkrankung veranschlagt werden.

Bei Personen mit leichter Gesundheitsstörung (leichte bis mittelgradige psychische Störung, leichte körperliche Erkrankung oder Schädigung) kann das Risiko für negative Antwortverzerrungen oder unzureichende Anstrengung in Leistungstests im Vergleich zu Personen mit schwerer Gesundheitsstörung erhöht sein, wenn die Person am Nachweis ihrer Beeinträchtigungen interessiert ist.

Bei Personen mit schwerer Gesundheitsstörung (schwere psychische Störung oder körperliche Erkrankung/Schädigung) kann das Risiko für positive Antwortverzerrungen im Vergleich zu leichter Gesundheitsstörung erhöht sein, wenn die Person am Nachweis ihrer Gesundheit oder intakter Funktionen interessiert ist.

Ausgewogene Gewichtung von Hinweisen auf willkürliche und unwillkürliche Verzerrungen

In der gutachterlichen Argumentation sollten Hinweise auf krankheits- oder störungsbedingte Beschwerden und Beeinträchtigungen, d. h. Hinweise auf Merkmale des Erlebens und Verhaltens, die der willentlichen Steuerung mutmaßlich entzogen sind, nicht prinzipiell stärker gewichtet werden als Hinweise auf Merkmale des Erlebens und Verhaltens, die der willentlichen Steuerung unterliegen. Hinweise auf mutmaßlich verzerrtes Antwortverhalten sind nicht automatisch als Indikatoren für willentlich verfälschte Angaben zu werten, wenn sie im Einklang mit den Anreizbedingungen interpretiert werden können.

Integration von Persönlichkeitsdiagnostik und freier Exploration

Zur Prüfung der psychologischen Plausibilität von Aussagen zu Ereignissen, Erfahrungen, Erinnerungen, Verhaltensweisen und deren Interpretation sollten Ergebnisse zu standardisiert gemessenen Verhaltensdispositionen (Persönlichkeitseigenschaften) und frei erhobene Informationen (explorierte Angaben oder freie Beobachtungen) in die Beurteilung integriert werden. Zweck der Persönlichkeitsdiagnostik ist dabei nicht die Bewertung der Glaubwürdigkeit der zu beurteilenden Person, sondern eine möglichst *differenzierte Beschreibung der kognitiven und handlungsbezogenen Referenzsysteme, an und in denen sich die zu beurteilende Person bei ihren Darstellungen/Selbstauskünften und Beschreibungen orientiert*.

Persönlichkeitsdiagnostik kann Angaben zu Denk- und Verhaltensstilen, überdauernden emotionalen Zuständen, Motiven, Überzeugungen, Funktionen, Fähigkeiten, Determinanten der Handlungssteuerung usw. liefern. Ihre Kenntnis kann dazu beitragen, den subjektiven Bedeutungsgehalt von Äußerungen oder Verhaltensweisen einer Person besser zu verstehen oder die Funktionalität situativer Aussagen oder Verhaltensweisen für die zu beurteilende Person zutreffend zu beurteilen. Insofern kann Persönlichkeitsdiagnostik auch Beiträge zur Bewertung der *psychologischen Plausibilität* und Glaubhaftigkeit frei erhobener Selbstauskünfte leisten. Es wird empfohlen, in der Validierungsdiagnostik die Glaubhaftigkeitsanalyse frei explorierter Informationen nicht nur an aussagepsychologischen Kriterien auszurichten, sondern persönlichkeitsdiagnostische Befunde in die Bewertung von Aussagen einzubeziehen.

Abgrenzung authentischer und willentlich verzerrter Leistungstestergebnisse

Psychische Störungen (z. B. depressive Störungen) sind in der Regel nicht so stark ausgeprägt, dass die betreffende Person nicht in der Lage wäre, die geringen motivationalen und kognitiven Mindestvoraussetzungen zu erfüllen, um psychodiagnostische Verfahren zu bearbeiten. Daher können auffällige Kennwerte in Performanz-Validierungstests auch bei gesicherter klinischer Symptomatik in der Regel als Hinweise auf eine unzureichende Testmotivation interpretiert werden, die der willentlichen Steuerung unterliegt.

Sofern Performanz-Validierungstests getrennte Normwerte für klinische Gruppen und für Normalpersonen ohne klinische Symptomatik vorsehen, eignet sich der Normwert der klinischen Vergleichsgruppe für die Unterscheidung klinisch begründeter und klinisch nicht begründeter Antwortverzerrungen.

Zeitlicher Bezug der Validierungsergebnisse

Das Ergebnis einer Kontrollskala oder eines Performanz-Validierungstests sollte nicht ohne weitere Prüfung zur Interpretation sämtlicher in einer Untersuchung erhobenen Information genutzt werden. Wegen

der Situationsspezifität der Validierungsdiagnostik kann ein größerer zeitlicher Abstand zwischen Validierungstest und inhaltlicher Diagnostik zu eingeschränkter Generalisierbarkeit der Validierungsergebnisse auf die Interpretation der inhaltlichen Ergebnisse führen. In Leistungstests eingebettete Validierungsindikatoren gelten streng genommen nur für die Validität des Verfahrens, in das sie eingebettet sind. Gleiches gilt für Kontrollskalen zur Erfassung generalisierter Antworttendenzen.

Erheblicher (mehrständiger) zeitlicher Abstand zwischen der Beantwortung mehrerer Kontrollskalen zur Erfassung gleicher Antwortverzerrungen kann abweichende Befunde ggf. erklären.

Auffällige validierungsdiagnostische Ergebnisse/Befunde in früheren Untersuchungen oder zu anderen Anlässen haben in aller Regel keinen Erklärungs- oder Vorhersagewert für die Interpretation validierungsdiagnostischer Ergebnisse in einer späteren Untersuchung. Entsprechend kann auch nicht von den Validierungsbefunden einer psychologischen (Zusatz-)Begutachtung auf die Gültigkeit oder Ungültigkeit der Ergebnisse in einer zeitnah zur gleichen Fragestellung durchgeführten medizinischen Untersuchung geschlossen werden.

Interpretation von Facetten generalisierter Antworttendenzen

Von einer auffälligen spezifischen Antworttendenz sollte nicht direkt auf eine oder mehrere andere Antworttendenzen generalisiert werden. Wenn sich die Hypothese zur Validierung z. B. auf die Tendenz zu stereotypen, mutmaßlich nicht erlebnisbasierten Antworten bezieht, dann sollte von dem Ergebnis zur Ausprägung stereotyper Antworttendenzen nicht auf andere Antworttendenzen (z. B. die generelle Tendenz, Unwahrscheinliches fälschlich zu behaupten) generalisiert werden. Die Interpretation eng umschriebener Antwortverzerrungen sollte den gegebenen inhaltlichen Rahmen einhalten.

Integration persönlichkeits- und verhaltensdiagnostischer Information

In die Interpretation frei explorierter Informationen sollten Ergebnisse und Erkenntnisse zu dispositionellen Eigenschaften der untersuchten Person integriert werden, um insbesondere intrapsychische dispositionelle Referenzgrößen (z. B. zu Wahrnehmungsstilen, Denkgewohnheiten, Emotionen, Motiven, Einstellungen, Überzeugungen) zuverlässiger und valider beurteilen zu können. Persönlichkeitsdiagnostik dient dabei nicht der Erfassung der Glaubwürdigkeit, sondern der Abschätzung multipler Referenzgrößen im mentalen Repräsentationssystem der befragten Person. Persönlichkeitsmerkmale können auch genutzt werden, um Aspekte der Normalität, der Vergleichbarkeit individueller Äußerungen mit Normstichproben, die Differenz von Außen- und Innenperspektive, aber auch Aspekte der Vereinbarkeit spezifischer Äußerungen oder Verhaltensweisen mit individuellen Denk- und Verhaltensgewohnheiten in die validierungsdiagnostische Beurteilung freier Selbstauskünfte einzubeziehen.

Einbezug von Informationen zum Krankheitsverlauf in Validierungsbefunde

Die Beurteilung von Validierungsinformationen erfordert bei gesundheitlichen Beschwerden und Beeinträchtigungen meist einen sorgfältigen Bezug zu den Kontextbedingungen, unter denen Beschwerden oder Beeinträchtigungen aufgetreten sind. Bei Fragen zu Art und Ausprägung von Erkrankungen, Schädigungen oder Krankheits- oder Schädigungsfolgen ist es meist sinnvoll, Aussagen zu krankheits- oder störungsbezogenen Ereignissen, Verhaltensweisen, Erfahrungen oder Zuschreibungen auch durch Abgleiche mit situativen Bedingungen oder bereits gesicherten Sachverhalten zu validieren.

Anhaltspunkte für eine suggestive Grundlage des aktuellen Beschwerdevortrags oder für suggestive Einflüsse auf die Entwicklung des Beschwerdevortrags im zeitlichen Verlauf sollten differenziert und vollständig erhoben werden. Die folgenden Quellen können dazu genutzt werden:

- Anhaltspunkte aus der Aktenlage
 - zum Ursprung der beklagten Beschwerden (Zeitachse),
 - zu früher bestehenden bzw. nachgewiesenen psychischen und/oder kognitiven Beschwerden,

- zu zwischenzeitlich erlangter psychischer oder kognitiver Restitution bzw. erhöhter Leistungsfähigkeit,
- zu motivationalen Aspekten im Behandlungsverlauf.
- Anhaltspunkte aus wissenschaftlichen Quellen
 - zum erwartbaren Spontanverlauf unbehandelter Beschwerden und ggf. Therapieverlauf der beklagten Beschwerden,
 - zur Plausibilität des ggf. berichteten Behandlungsverlaufs.
- Anhaltspunkte aus der gutachterlichen Untersuchung
 - zum Ursprung der beklagten Beschwerden (Zeitachse),
 - zu früher bestehenden bzw. nachgewiesenen psychischen und/oder kognitiven Beschwerden,
 - zu zwischenzeitlich erlangter psychischer oder kognitiver Beschwerden-Verbesserung bzw. Leistungsfähigkeit,
 - zu motivationalen Aspekten im Behandlungsverlauf,
 - zur subjektiven Krankheitstheorie,
 - zu lernpsychologischen Faktoren im Rahmen der Behandlung.
- Anhaltspunkte aus der testpsychologischen Diagnostik
 - Motivationale Bereitschaft, sich bei Leistungsprüfungen ausreichend anzustrengen,
 - Bereitschaft und/oder ausreichende „Reflexionsfähigkeit“, in Fragebogenverfahren erlebnisbezogene Angaben bezogen auf den aktuellen Zeitraum zu machen.
 - Diskrepanzen zwischen Angaben im freien Bericht und bei der Bearbeitung von Fragebögen.

H. Die Interpretation validierungsdiagnostischer Informationen erfolgt kontextbezogen auf der Grundlage psychologisch-psychodiagnostischer Erkenntnisse und Prinzipien. Eines der Prinzipien besagt, dass willentliche inhaltliche Verzerrungen und zufallsbedingten Verzerrungen getrennt zu bewerten sind. Validierungsdiagnostische Ergebnisse sollten möglichst eng auf den Sachverhalt oder die Inhalte bezogen werden, zu deren Validierung sie erhoben wurden. Generalisierende personale Zuschreibungen validierungsdiagnostischer Auffälligkeiten im Sinne stabiler, situationsübergreifend wirksamer Eigenschaften sollten vermieden werden.

9. *Einbezug von Validierungsbefunden in die inhaltliche Ergebnisinterpretation*

Ergebnisinterpretation gemäß Qualitätsstandards für psychologische Gutachten

Die Interpretation der auf die inhaltlichen Fragestellungen bezogenen Untersuchungsergebnisse nach vorausgehender validierungsdiagnostischer Bewertung folgt den in den Qualitätsstandards für psychologische Gutachten der Föderation deutscher Psychologinnenvereinigungen festgelegten Empfehlungen.

Besonderes Gewicht liegt bei der psychologischen Eigenschafts- und Funktionsbeurteilung auf der Auswahl geeigneter Normen/Referenzgrößen, die auch zu Validierungszwecken herangezogen werden können. Vielfach sind – nach Kontrolle bzw. Ausschluss motivationaler Verzerrungen – relevante Vergleichsgruppen nicht Patienten oder Personen mit behandlungsbedürftigen Erkrankungen oder psychischen Störungen, sondern (ggf. altersadjustierte) repräsentative Bevölkerungsstichproben.

Einbezug von Akteninformationen in selbst erhobene Validierungsbefunde

Befunde zu eingeschränkter Validität von Selbstauskünften oder Testergebnissen aus früheren Untersuchungen oder anderen Untersuchungsanlässen sollten nur zurückhaltend, ggf. auch gar nicht zur Interpretation der selbst erhobenen Informationen genutzt werden. Gleiches gilt für Hinweise auf eine uneingeschränkte Validität. Validierungsbefunde zur Glaubhaftigkeit von Aussagen oder Testergebnissen sind nur eng situationsbezogen zu interpretieren, sofern sie nicht die Glaubwürdigkeit einer Person betreffen. Von der Validität von Ergebnissen in einer Untersuchung ist nicht auf die Validität der Ergebnisse in einer anderen Untersuchung zu schließen.

Inhaltliche Ergebnisinterpretation bei unzureichenden Mindestvoraussetzungen

Bei konsistenten Hinweisen auf unzureichende kognitive oder emotional-motivationale Mindestvoraussetzungen für die Testbearbeitung sollten die zeitnah in Leistungstests erhobenen Informationen inhaltlich nicht oder nur sehr zurückhaltend interpretiert werden. Der Nachweis mangelnder kognitiver oder motivationaler Mindestvoraussetzungen für die Erhebung diagnostischer Informationen kann die Folgerung begründen, dass die Voraussetzungen für eine hinreichend differenzierte Beurteilung des (kognitiven oder psychomotorischen) Leistungsniveaus oder der erfragten Sachverhalte nicht gegeben sind.

Belege für zufälliges oder logisch inkonsistentes Antworten schließen aber nicht grundsätzlich aus, dass die Person in der Untersuchung zu einem willentlich kontrollierten Antwortverhalten in der Lage war oder gewesen wäre. Auffällige formale Antworttendenzen schließen jedoch aus, dass vorliegende Ergebnisse standardisierter Mess- oder Testverfahren im intendierten Sinne interpretiert werden können. Empirisch ermittelte Standard- oder Normwerte bilden in diesen Fällen die tatsächliche Merkmalsausprägung nicht zuverlässig und valide ab.

Inhaltliche Ergebnisinterpretation bei gesichert willentlich verzerrten Selbstberichten

Willentlich kontrollierte Antwort- oder Reaktionsverzerrungen, die im Einklang mit den Ergebnisinteressen der untersuchten Person stehen, können nach eingehender fachlicher Prüfung zu der Folgerung führen, dass der zu beurteilende Sachverhalt (z. B. Erlebnisnähe eines beschriebenen Ereignisses, Ausprägung einer bestimmten Eigenschaft, Testergebnis in einem Fragebogen, Ergebnisse freier oder strukturierter Exploration) wahrscheinlich nicht hinreichend zuverlässig beurteilt werden kann. Entsprechend kann der psychologische Sachverhalt, auf den sich die Argumentation der gutachterlichen Stellungnahme stützt, als nicht hinreichend gesichert gelten.

Die verminderte Gültigkeit individueller Selbstberichte kann umso höher veranschlagt werden, je genauer und differenzierter der zu beurteilende (inhaltliche) Sachverhalt durch die Ergebnisse erfasst wird (z. B. allgemeine Zuschreibung einer Eigenschaft vs. Zuschreibung einer Eigenschaftsfacette vs. Zuschreibung einer quantitativen Ausprägung der Eigenschaftsfacette).

Auffällige willentliche Antwort- oder Reaktionsverzerrungen, die den Ergebnisinteressen der untersuchten Person zuwiderlaufen, können zu der Folgerung führen, dass die Verzerrungen nicht situationsübergreifend auf sämtliche in der Untersuchung erhobenen Informationen angewendet werden sollten. In diesen Fällen ist der Situationsbezug der Antwortverzerrungen sorgfältig zu prüfen und in die Ergebnisinterpretation einzubeziehen. Einschränkungen der Glaubhaftigkeit sollten dann nur thematisch eng und situationsspezifisch interpretiert werden.

Interpretation von Funktions- und Leistungstestergebnissen bei mangelnder Anstrengungsbereitschaft

Die Ergebnisse kognitiver Leistungstests markieren - unabhängig von der Anstrengungsbereitschaft bei der Testdurchführung - jeweils die untere Grenze der tatsächlichen individuellen Leistungsfähigkeit. Ein durchschnittliches Leistungstestergebnis (z. B. Konzentrationstest, Gedächtnistest, Intelligenztest) kann ohne weitere Prüfung der Anstrengungsbereitschaft als Maß für die tatsächliche kognitive Leistungsfähigkeit ausreichend sein, wenn geprüft werden soll, ob die zu beurteilende Person über eine mindestens durchschnittliche kognitive Leistungsfähigkeit verfügt.

Eine willentliche Verzerrung von Leistungstestergebnissen in einen Leistungsbereich, der über dem Niveau der tatsächlichen Leistungsfähigkeit liegt, ist durch Konstruktionsprinzipien kognitiver Leistungstests nicht möglich.

Die Ergebnisse in einem Performanz-Validierungstest können als Indikatoren für die Anstrengungsbereitschaft in diesem Test interpretiert werden.

Die folgenden Bedingungen sprechen dafür, von einer mutmaßlich mangelnden Anstrengungsbereitschaft im Validierungstest auf eine mangelnde Anstrengungsbereitschaft in einem Leistungstest zu schließen:

- Der Befund mangelnder Anstrengungsbereitschaft ist durch mehrere voneinander unabhängige Validierungsindikatoren gesichert.
- Validierungstests und Leistungstest wurden in zeitlicher Nähe bearbeitet.
- Die von der Testperson gegenüber der Validierungsbedingung gezeigte motivationale Haltung (z. B. verbal oder im Verhalten geäußerte Kooperationsbereitschaft oder Testabwehr) unterschied sich nicht wesentlich von der gegenüber der Leistungstestung kommunizierten Haltung.
- Art oder Gestaltung der Aufgaben des Validierungstests weisen Ähnlichkeiten auf zur Art oder Gestaltung der Aufgaben des Leistungstests auf.

I. Der Einbezug von Validierungsbefunden in die Beantwortung der gutachterlichen Fragestellungen sollte hypothesenbezogen erfolgen und dabei die Konsistenz der Befunde berücksichtigen. Bei inkonsistenten Validierungsbefunden ist zu prüfen, inwiefern diese durch unterschiedliche Untersuchungskontexte erklärt werden können. Wenn inkonsistente Validierungsbefunde weder durch die Messeigenschaften der verwendeten Mess- oder Testverfahren noch durch situative Kontextbedingungen erklärt werden können, kann das den diagnostischen Wert der Befunde in Frage stellen.

10. Beantwortung der rechtlichen Fragestellungen im Gutachten

Beachtung der fachlichen Grenzen

Psychologische Sachverständige beantworten psychologische Fragestellungen auf der Grundlage psychologischer Begrifflichkeiten, Konzepte und Modelle. Vorschläge zur Beantwortung der rechtlichen Fragestellungen dürfen nur formuliert werden, wenn diese explizit von den Auftraggebern gewünscht werden. Die Beantwortung der rechtlichen Fragestellungen ist nicht Aufgabe der Sachverständigen.

Der oder die Sachverständige nutzt im Rahmen gutachterlicher Tätigkeit die Begrifflichkeiten und Verfahrensweisen der rechtlichen Beweiswürdigung nicht. Rechtsbegriffe mit Bezug zur Validierung von

Befunden wie z. B. Beweissicherung, Glaubhaftmachung, Tatsache, Beweismaß der Wahrscheinlichkeit, Vollbeweis usw. sind von Sachverständigen nicht zu verwenden. Gleiches gilt für Verfahrensweisen, die in der Rechtsprechung verwendet werden, um bestimmte Sachverhalte rechtlich zu sichern (z. B. das Prinzip „in dubio pro reo“ oder Entscheidungsbegründungen durch rechtliche Verfahrensabläufe wie z. B. Beweislast oder Beweislastumkehr).

Auch aus validen Befunden zur individuellen Funktions- oder Leistungsfähigkeit resultiert keine direkte Beurteilung des rechtlichen Sachverhalts (z. B. Arbeitsfähigkeit, Erwerbsminderung, Kraftfahreignung, Behinderung, Schuldfähigkeit, Beziehungsfähigkeit usw.)

Darstellung der Validierungsdiagnostik im Gutachten

Das Gutachten sollte zum Schutz der Verfahren keine Details zu den in der Untersuchung verwendeten standardisierten Validierungsmethoden enthalten.

Intermethodale oder qualitative (nicht psychometrische) Vergleiche zur Konsistenzprüfung sollten dargestellt werden, sofern sich die Validitätsbefunde auf diese Daten stützen. Die validierungsdiagnostischen Ergebnisse sollten im Gutachten so dargestellt werden, dass Auftraggeber den Zusammenhang zwischen den Ergebnissen und den daraus abgeleiteten Schlussfolgerungen nachvollziehen können.

Hinweis auf Einbezug validierungsdiagnostischer Befunde in die Stellungnahme

Wenn die rechtliche Fragestellung nur auf Aspekte der psychologischen Beschreibung oder Erklärung von Merkmalen des individuellen Erlebens oder Verhaltens zielt, sollte die Beantwortung der gutachterlichen Fragestellungen den Hinweis enthalten, dass die sachverständige Beantwortung der psychologischen Fragestellungen eine Validierung der erhobenen Information voraussetzt.

Spezifizierung der validierungsdiagnostischen Befunde im Gutachten

Wenn die Validierungsdiagnostik zu dem Ergebnis führt, dass die zur Beantwortung der gutachterlichen Fragestellungen erhobenen Informationen keine inhaltliche Interpretation zulassen, ist dies in der Stellungnahme aufzuführen. Bei Hinweisen auf eine zumindest teilweise eingeschränkte Gültigkeit der erhobenen Informationen sollte in der Stellungnahme differenziert beschrieben werden,

- welche Ergebnisse von Validitätseinschränkungen insbesondere betroffen sind,
- welche Ergebnisse von Validitätseinschränkungen teilweise betroffen sind
- welche Ergebnisse von Validitätseinschränkungen wahrscheinlich nicht betroffen sind.

Stellungnahme zur explizit geforderten rechtlichen Bewertung bei auffälligen Validierungsbefunden

Die rechtliche Bewertung auffälliger Validierungsbefunde hängt zunächst vom rechtlichen Kontext ab, in dem die Befunde genutzt werden. Führen auffällige Validierungsbefunde dazu, dass die untersuchte Partei, die z. B. einen Gesundheitsschaden zur Durchsetzung zivilrechtlicher Ansprüche nachweisen muss, diesen Nachweis nicht erbringen kann, dann geht dieser fehlende Nachweis zu Lasten der Partei.

Im Sozialrecht gilt für Gerichte der Grundsatz der Amtsermittlung. Das heißt: Wenn es aus Sicht des Gerichts überzeugende Gründe gibt, trotz auffälliger Validierungsbefunde vom Vorliegen oder von der kausalen Zuschreibung des Gesundheitsschadens in einer rechtlich relevanten Qualität oder Intensität auszugehen, dann ist diese (rechtliche) Gesamtbewertung maßgeblich.

Wenn das Gericht den Sachverständigen explizit auffordert, Vorschläge für die rechtliche Bewertung zu formulieren, ist es für den Sachverständigen i.d.R. notwendig, die unterschiedlichen Formen der

rechtlichen Beweissicherung in die Kalkulation einzubeziehen. Im Zweifel sollte sich der/die Sachverständige dennoch an den Prinzipien der psychologischen Befundssicherung orientieren. Diese sehen vor, dass konsistent erhöhte oder extrem auffällige Validierungsbefunde eine hinreichend gesicherte Aussage über die zu beurteilenden psychologischen Sachverhalte nicht zulassen.

Stellungnahme mit inhaltlicher Beurteilung trotz Validitätseinschränkungen

Aufgrund der zeitlichen und inhaltlichen Spezifität von Validierungsbefunden ist es möglich, dass auch bei auffälligen Validierungsbefunden gültige/zutreffende Aussagen zu bestimmten Sachverhalten gemacht werden können. Entscheidend ist, dass sich die Beantwortung der gutachterlichen Fragestellung oder ausgewählter Aspekte auf Ergebnisse oder Erkenntnisse stützt, die nach fachlichem Ermessen ausreichend befundlich gesichert sind. Die Stellungnahme sollte deutlich machen, auf welche Befunde sie gestützt ist und inwiefern sich aus der reduzierten Befundlage Einschränkungen (a) der Beschreibung, (b) der Erklärung und ggf. (c) der Vorhersage gesundheitlicher Beeinträchtigungen, ihrer Verarbeitung und ihrer Auswirkungen ergeben.

Funktions- und Fähigkeitsbeurteilung auf der Grundlage von Ausgangsbedingungen, Verhalten und Verhaltenswirkung

Die psychologische Beurteilung von Funktionen und Funktionseinschränkungen setzt validierte Angaben zu (a) Ausgangs- bzw. Anforderungsbedingungen, (b) anforderungsbezogenem Verhalten und (c) Verhaltensergebnissen voraus, die psychologisch sinnvoll (plausibel) aufeinander bezogen werden können. Entsprechend kann sich die Beantwortung der rechtlichen Fragestellungen auf Informationen zu folgenden Variablengruppen beziehen: Anforderungsbedingungen (z. B. Lebensbedingungen, Arbeitsbedingungen), personenbezogene Merkmals- oder Fähigkeitsausprägungen, förderliche und hinderliche Kontextbedingungen (z. B. physikalische, organisatorische, soziale Bedingungen des Verhaltens) sowie auf typische Person-Umwelt-Wechselwirkungen. Auf dieser Grundlage können Funktionen und Funktionseinschränkungen als dynamische Anpassungsprozesse beschrieben werden. Für die genannten Variablengruppen können jeweils unterschiedliche Formen der Validierung notwendig und geeignet sein.

Die Funktionsfähigkeit einer Person kann bei der Beantwortung der rechtlichen Fragestellung sowohl als dynamischer Prozess, als auch im Sinne von Eigenschaften oder Verhaltensdispositionen beschrieben werden. Die Validierung von Aussagen über dynamische Prozesse kann in der Stellungnahme vorwiegend durch die Kontextualisierung bzw. den situativen Bezug gemachter Angaben oder Testergebnisse erfolgen. Die Validierung von Aussagen über stabile Eigenschaften kann v. a. über die Bezugnahme auf validierte Ergebnisse standardisierter Mess- oder Testverfahren erfolgen.

J. Rechtliche und psychologische Begriffe und Konzepte zur Absicherung von Aussagen und Urteilen sind klar zu trennen. Wahrscheinlichkeitsbasierte psychologische Aussagen über die Gültigkeit von Aussagen, Beobachtungen oder Testergebnissen sind nicht gleichzusetzen mit rechtlichen Beweismaßen. Die rechtliche Bewertung von Validierungsbefunden auf Wunsch des Auftraggebers kann nur unter Vorbehalt erfolgen und erfordert den Einbezug rechtlicher Bewertungsmaßstäbe.

Literatur

- Baer, R. A., Miller, J. (2002). Underreporting of psychopathology on the MMPI-2: A meta-analytic review. *Psychological Assessment* 14 (1), 16–26. DOI: 10.1037//1040-3590.14.1.16.
- Bahlo, S. (2014). *Selbstbeschreibung von Patienten mit organischen Hirnleistungsstörungen im MMPI-2 unter besonderer Berücksichtigung der Validitätsskalen: Eine Erprobung des MMPI-2 und seiner Validitätsskalen an Patienten mit Hirnschädigung*. Tübingen, Univ., Diss.
- Boone, K. B. (Hg.) (2021). *Assessment of feigned cognitive impairment. A neuropsychological perspective*. Second edition. New York, London: Guilford.
- Boskovic, I., Merckelbach, H., Vrij, A., Bogaard, G. & Hope, L. (2017). The Verifiability Approach to detection of malingered physical symptoms. In: *Psychology, crime & law*. DOI: 10.1080/1068316X.2017.1302585.
- Chafetz, M. D., Sweet, J. J., Boone, K. B., Cox, D., Hall, V., Kirkwood, M. W., Lafosse J. M., Merten, T. & Oldenburg C. (2025). Neuropsychological review of records in forensic cases: An AACN best practices paper with international perspectives. *Clinical Neuropsychol.* 2025 May; 39 (4) 839-869.
- Dettenborn, H., Merten, T. (Hg.) (2013): *Diagnostik der Beschwerdenuvalidität*. Berlin: Deutscher Psychologen Verlag (Praxis der Rechtspsychologie).
- Dohrenbusch, R. & Brockhaus, R. (2025). Validierung von Selbstauskünften bei erhöhtem Risiko positiv verzerrter Angaben. In R. Dohrenbusch (ed.): *Psychologische Begutachtung*, 869-884. Berlin: Springer.
- Engel, R. R. (Hg.) (2000): *Minnesota Multiphasic Personality Inventory-2. MMPI-2 von S. R. Hathaway und J. C. McKinley*. Manual. Bern: Hogrefe.
- Engel, R. R., Ben-Porath, Y. S., Tellegen, A. (2019). *MMPI-2-RF, Minnesota Multiphasic Personality Inventory-2, Restructured Form*. Deutschsprachige Adaptation des Minnesota Multiphasic Personality Inventory-2, Restructured Form von Yossef S. Ben-Porath und Auke Tellegen. Bern: Hogrefe.
- Engel, R. R., Groves, J. A., Morey, L. C. (2012). *VEI. Verhaltens- und Erlebensinventar: Deutschsprachige Adaptation des Personality Assessment Inventory (PAI)* von Leslie C. Morey: Manual. Bern: Hogrefe.
- Ferrara, S. D., Boscolo-Berto, R., Viel, G. (Hg.) (2016). *Personal Injury and Damage Ascertainment under Civil Law. State-of-the-Art International Guidelines*. Cham: Springer.
- Fokas, K. F., Brovko, J. M. (2020). Assessing symptom validity in psychological injury evaluations using the MMPI-2-RF and the PAI: An updated Review. *Psychological Injury and Law* 13 (4), 370–382. DOI: 10.1007/s12207-020-09393-8.
- Grossman, L. S. & Wasylw, O. E. (1988). A psychometric study of stereotypes: Assessment of malingering in a criminal forensic group. *Journal of Personality Assessment*, 52(3), 549-563.
- Hall, H. V., Poirier, J. G. (2020). *Detecting Malingering and Deception. Forensic Distortion Analysis (FDA-5). Third edition*. Boca Raton: CRC Press Taylor & Francis Group.
- Heilbrun, K., Bennett, W. S., White, A. J. & Kelly, J. (1990). An MMPI-based empirical model of malingering and deception. *Behavioral Sciences & the Law*, 8(1), 45-53.
- Horton, A. M. N., Reynolds, C. R. (Eds.) (2021). *Detection of malingering during head injury litigation*. (3rd ed.,) Springer Nature Switzerland AG.

- Larrabee, G. J. (2012): Performance Validity and Symptom Validity in Neuropsychological Assessment. *Journal of the International Neuropsychological Society*, 18 (4), 625–630. DOI: 10.1017/s1355617712000240.
- Larrabee, G. J. (2015): The multiple validities of neuropsychological assessment. *The American psychologist* 70 (8), S. 779–788. DOI: 10.1037/a0039835.
- Martin, P. K., Schroeder, R. W. (Hg.) (2022). *Validity Assessment in Clinical Neuropsychological Practice. Evaluating and Managing Noncredible Performance*. New York, London: Guilford Press.
- Merten, T. (2014). *Beschwerdenvalidierung*. Göttingen, Bern, Wien: Hogrefe.
- Merten, T. (2023). *Beschwerdenvalidierung in der Begutachtung, Klinik und Rehabilitation*. Göttingen, Bern, Wien: Hogrefe.
- Merten, T. (2025). *Leitfaden für die neuropsychologische Begutachtung*. Göttingen, Bern, Wien: Hogrefe.
- Merten, T., Giger, P., Merckelbach, H., Stevens, A. (2019). *SRSI. Self-Report Symptom Inventory - Deutsche Version: Manual*. Göttingen: Hogrefe.
- Neumann-Zielke, L., Bahlo, S., Diebel, A., Riepe, J., Roschmann, R., Schötzau-Fürwentsches, P., Wetzig, L. (2014). Leitlinie Neuropsychologische Begutachtung. *Zeitschrift für Neuropsychologie*, 26 (4), 2015, 289–306.
- Picard, E., Aparcero, M., Nijdam-Jones, A. & Rosenfeld, B. (2023). Identifying positive impression management using the MMPI-2 and the MMPI-2-RF: A meta-analysis. *The Clinical Neuropsychologist* 37 (3), 545–561. DOI: 10.1080/13854046.2022.2077237.
- Rogers, R., Bender, S. D. (Hg.) (2018). *Clinical assessment of malingering and deception*. Fourth edition. New York, London: The Guilford Press.
- Rohling, M. L., Larrabee, G. J., Greiffenstein, M. F., Ben-Porath, Y. S., Lees-Haley, P., Green, P., Greve, K. W. (2011). A misleading review of response bias: Comment on McGrath, Mitchell, Kim, and Hough (2010). *Psychological Bulletin* 137 (4), 708-12; Authors reply, 713-5. DOI: 10.1037/a0023327.
- Schmidt, T., Watzke, S., Lanquillon, S. & Stieglitz, R.-D. (2019). *SIRS-2. Structured Interview of Reported Symptom*. Deutschsprachige Adaption des Structured Interview of Reported Symptoms, 2nd edition, von Richard Rogers, Kenneth W. Sewell und Nathan D. Gillard. Bern: Hogrefe.
- Schroeder, R. W. (2022). *Validity Assessment in Clinical Neuropsychological Practice. Evaluating and Managing Noncredible Performance*. New York: Guilford Publications.
- Sherman, E., Slick, D., & Iverson, G. (2020). Multidimensional malingering criteria for neuropsychological assessment: A 20-year update of the malingered neuropsychological dysfunction criteria. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, 35(6), 735–764.
- Slick, D. J., Sherman, E. M. S., Iverson, G. L. (1999). Diagnostic criteria for malingering cognitive dysfunction: Proposed standards for Clinical Practice and Research. *Clinical Neuropsychology*, 13, 545-61.
- Stevens, A., Friedel, E., Mehren, G., Merten, T. (2008). Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants. In: *Psychiatry research* 157 (1-3), 191–200. DOI: 10.1016/j.psychres.2007.01.00

- Sweet, J. J., Heilbronner, R. L., Morgan, J. E., Larrabee, G. J., Rohling, M. L., Boone, K. B., Kirkwood, M. W., Schroeder, R. W., & Suhr, J. A. and Conference Participants. American Academy of Clinical Neuropsychology (AACN) 2021 consensus statement on validity assessment: Update of the 2009 AACN consensus conference statement on neuropsychological assessment of effort, response bias, and malingering. *The Clin. Neuropsychol.* 2021 Aug; 35(6), 1053–1106.
- Thies, E. (2012): Der deutsche MMPI-2. Effektivität der Validitätsskalen in der Aufdeckung von Antwortverzerrung. Marburg: Tectum.
- Tippelt, S. (2015). Diagnostik positiver Antwortverzerrungen bei forensischen Begutachtungen. Untersuchung der Validität und Brauchbarkeit der Supernormalität-Skala (SN-S). Online Publikation. https://edoc.ub.uni-muenchen.de/18210/1/Tippelt_Susanne.pdf
- Young, Gerald (2014). Malingering, Feigning, and Response Bias in Psychiatric/Psychological Injury. Implications for Practice and Court. Dordrecht, s. l.: Springer Netherlands (International Library of Ethics, Law, and the New Medicine, 56). Online verfügbar unter <https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=699334>
- Walter, F., Lid, N., Petermann, F. & Kobelt, A. (2017). Einsatz des Beschwerdenuvalidierungstests BEVA in der sozialmedizinischen Begutachtung. *Die Rehabilitation* 56 (3), S. 173–180. DOI: 10.1055/s-0042-121370.
- Wygant, D. B., Sellbom, M., Gervais, R. O., Ben-Porath, Y. S., Stafford, K. P.; Freeman, D. B. & Heilbronner, R. L. (2010). Further validation of the MMPI-2 and MMPI-2-RF Response Bias Scale: findings from disability and criminal forensic settings. *Psychological Assessment* 22 (4), 745–756. DOI: 10.1037/a0020042.

Autorinnen und Autoren

Simone Bahlo
Felix Aßhauer
Robbi Brockhaus
Andrea Diebel
Carsten Dörnberg
Sandro Feuerpeil
Ralf Dohrenbusch