

Diagnostik bei Migrantinnen und Migranten

Débora B. Maehler
Alexandra Shajek
Heinz Ulrich Brinkmann
(Hrsg.)

Ein Handbuch

Diagnostik bei Migrantinnen und Migranten

Débora B. Maehler
Alexandra Shajek
Heinz Ulrich Brinkmann
(Hrsg.)

Diagnostik bei Migrantinnen und Migranten

Ein Handbuch

 **hogrefe**

Dr. Débora B. Maehler, geb. 1977. Studium der Psychologie an der Universität Potsdam. 2011 Promotion. Seit 2012 als Senior Researcherin bei GESIS – Leibniz Institut für Sozialwissenschaften tätig. Post-Doc Fellow am College for Interdisciplinary Educational Research/ CIDER (2013 – 2016). Seit 2016 Leitung des Forschungsdatenzentrum PIAAC (FDZ PIAAC). Forschungsschwerpunkte: Migrations- und Integrationsforschung sowie Bildungsforschung.

Dr. Alexandra Shajek, geb. 1979. Studium der Psychologie in Münster, Uppsala (Schweden) und Berlin, anschließend Mitarbeiterin am Max-Planck-Institut für Bildungsforschung und am Deutschen Institut für Wirtschaftsforschung. 2014 Promotion. Seit 2014 Tätigkeit als bildungspolitische Beraterin bei der VDI/VDE-IT. Arbeitsschwerpunkte: Lebenslanges Lernen, Bildungs- und Innovationsindikatorik sowie quantitative und qualitative Forschungsmethoden.

Dr. Heinz Ulrich Brinkmann, geb. 1946. Studium der Wirtschafts- und Sozialwissenschaften an der Universität zu Köln sowie an ausländischen Universitäten. Dipl.-Volkswirt (soz. wiss. R.) 1978, Dr. rer. pol. 1982. Tätigkeit als wissenschaftlicher Mitarbeiter und Lehrbeauftragter an mehreren deutschen Universitäten. 1988–2008 bei der Bundeszentrale für politische Bildung in Bonn. Seitdem Forschung und Publikationen im sozialwissenschaftlichen Bereich mit dem Schwerpunkt Migration/Integration.

Copyright-Hinweis:

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

Der Nutzer verpflichtet sich, die Urheberrechte anzuerkennen und einzuhalten.

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

Satz: ARThür Grafik-Design & Kunst, Weimar
Format: PDF

1. Auflage 2018

© 2018 Hogrefe Verlag GmbH & Co. KG, Göttingen

(E-Book-ISBN [PDF] 978-3-8409-2786-7; E-Book-ISBN [EPUB] 978-3-8444-2786-8)

ISBN 978-3-8017-2786-4

<http://doi.org/10.1026/02786-000>

Nutzungsbedingungen:

Der Erwerber erhält ein einfaches und nicht übertragbares Nutzungsrecht, das ihn zum privaten Gebrauch des E-Books und all der dazugehörigen Dateien berechtigt. Der Inhalt dieses E-Books darf von dem Kunden vorbehaltlich abweichender zwingender gesetzlicher Regeln weder inhaltlich noch redaktionell verändert werden. Insbesondere darf er Urheberrechtsvermerke, Markenzeichen, digitale Wasserzeichen und andere Rechtsvorbehalte im abgerufenen Inhalt nicht entfernen.

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

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

Das Anfertigen von Vervielfältigungen, das Ausdrucken oder Speichern auf anderen Wiedergabegeräten ist nur für den persönlichen Gebrauch gestattet. Dritten darf dadurch kein Zugang ermöglicht werden.

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

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

Anmerkung:

Sofern der Printausgabe eine CD-ROM beigelegt ist, sind die Materialien/Arbeitsblätter, die sich darauf befinden, bereits Bestandteil dieses E-Books.

Inhaltsverzeichnis

Teil I

Einführung und Grundlagen

- 1 Diagnostische Verfahren für Migrantinnen und Migranten in Deutschland:
Eine Bestandsaufnahme
Débora B. Maehler, Heinz Ulrich Brinkmann & Alexandra Shajek 9
- 2 Testtheoretische Grundlagen der psychologischen Diagnostik:
Ein Überblick
Alexandra Shajek, Débora B. Maehler & Heinz Ulrich Brinkmann 22
- 3 Anwendungsorientierte Einführung in die Übersetzung und Adaptation
von Messinstrumenten
Anouk Zabal & Dorothee Behr 32
- 4 Das Bildungsniveau von Migrantinnen und Migranten: Herausforderungen
in Erfassung und Vergleich
Silke L. Schneider 47

Teil II

Pädagogisch-psychologische Diagnostik bei Kindern und Jugendlichen

- 5 Entwicklungsdiagnostik
Thorsten Macha & Franz Petermann 57
- 6 Sprachstandsdiagnostik
Valentina Reitenbach, Martin Schastak & Dominique Rauch 95
- 7 Leistungsstanddiagnostik
Nicole Haag, Birgit Heppt & Stefan Schipolowski 150

Teil III

Klinische und Persönlichkeits-Diagnostik bei Migrantinnen und Migranten

- 8 Klinische Diagnostik bei Kindern und Jugendlichen
Julian Busch, Birgit Leyendecker & Rainer G. Siefen 197

9 Kultursensible klinische Psychodiagnostik bei Erwachsenen
Yuriy Nesterko & Heide Glaesmer 245

Teil IV

Berufsbezogene Diagnostik bei Migrantinnen und Migranten

10 Diagnostik zur Platzierung von Migrantinnen und Migranten
auf dem Arbeitsmarkt
Jessica Erbe & Rebecca Atanassov 287

11 Gleichwertigkeitsfeststellung für im Ausland erworbene
Berufsqualifikationen
Rebecca Atanassov & Jessica Erbe 296

12 Verfahren zur Erfassung von im Ausland formal und informell
erworbener beruflicher Kompetenzen – Potentiale und Grenzen
technologiebasierter Arbeitsplatzsimulationen
Viola Deutscher & Esther Winther 317

13 Berufliche Eignungstests
*Stefan Krumm, Sibylle Detel, Michela Schröder-Abé, Matthias Ziegler &
Johannes Zimmermann* 339

Die Autorinnen und Autoren des Bandes 391

Sachwortverzeichnis 394

Teil I

Einführung und Grundlagen

1 Diagnostische Verfahren für Migrantinnen und Migranten in Deutschland: Eine Bestandsaufnahme

Débora B. Maehler, Heinz Ulrich Brinkmann & Alexandra Shajek

1.1 Zielsetzungen des Buches

Die weltweiten Fluchtbewegungen der letzten Jahre – ausgelöst u. a. durch Kriege, wirtschaftliche Krisen oder Naturkatastrophen – gehen mit einem starken Anstieg der Zuwanderungen nach Deutschland einher.¹ Aktuelle Zahlen zu den positiven Entscheidungen über Asylanträge (also zur Anzahl von Personen, die sich rechtmäßig in Deutschland aufhalten dürfen) geben ein erstes Bild von den zu erwartenden Veränderungen in der deutschen Gesellschaft: Im Jahr 2016 waren mit 445.210 positiven Entscheidungen in Deutschland dreimal mehr als im Jahr zuvor zu verzeichnen (Eurostat, 2017).² Hierbei hat Deutschland derzeit im internationalen Vergleich eine der höchsten Flüchtlingsaufnahmequoten³ (vgl. Abbildung 1.1). Im Gesamtvergleich der Europäischen Union entfallen auf Deutschland mehr als 60 % der gesamten positiven Entscheidungen. In den Jahren 2015 und 2016 wurde die Zuwanderung in Europa bzw. vor allem in Deutschland quantitativ durch Flüchtlinge aus den arabischen Ländern des Nahen Ostens bestimmt (OECD, 2016).

Mit der Zunahme von Geflüchteten in Deutschland in den letzten Jahren geht eine steigende Nachfrage nach diagnostischen Verfahren für diese Personengruppen – beispielsweise für die klinische Praxis, die Schule oder die soziale Arbeit – einher. Bislang liegen jedoch kaum validierte Verfahren für Personen mit Migrations-

- 1 Die bisherigen Peaks an Zuwanderung nach Deutschland gab es um 1970 sowie um 1990. Für einen Überblick über das bisherige Migrationsgeschehen vgl. Sauer & Brinkmann, 2016, S. 6f.; Brinkmann, 2016, S. 146–150; Bundesamt, 2016, S. 28–36.
- 2 Die Zahlenangaben dieses Satzes beziehen sich auf Entscheidungen in erster Instanz. Sie fassen die Anzahl der Asylenerkennungen, der Flüchtlingsanerkennungen, der Gewährung von subsidiärem Schutz und der Feststellung eines Abschiebungsverbotes zusammen. Diese Zahlen weichen geringfügig von den Zahlen des Bundesamtes (2016) ab.
- 3 Anteil positiver Bescheide an allen gestellten Anträgen.

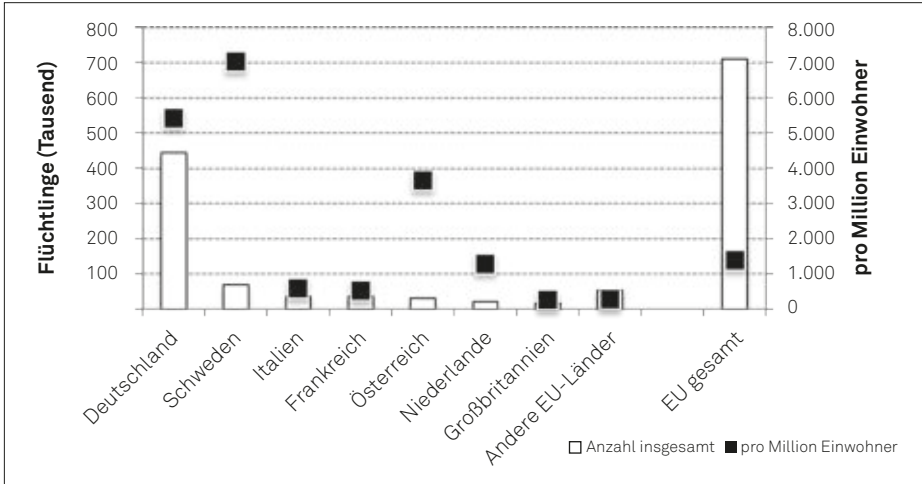


Abbildung 1.1: Positive Entscheidungen über Asylanträge in Ländern der Europäischen Union (2016); Quelle: Eurostat, 2017

hintergrund⁴ vor. Eine Recherche in einschlägigen, bibliographischen Suchmaschinen (FIS, SOWIPORT und PSYNDEX) nach diagnostischen Verfahren für verschiedene Anwendungsbereiche (z.B. Sprachstandsdiagnostik in der Schule oder klinische Diagnostik für Erwachsene) weist nur eine geringe Anzahl an Ergebnissen für die entsprechende Zielgruppe auf.⁵ Tabelle 1.1 zeigt exemplarisch die Ergebnisse für PSYNDEX. Wie ersichtlich wird, finden sich primär Beiträge für die Bereiche Entwicklungsdiagnostik, Sprachstandsdiagnostik und Leistungsstanddiagnostik – sowie mit Einschränkung für die klinische Diagnostik. Obwohl die Anzahl der gefundenen Beiträge in diesen Anwendungsgebieten im zweifeligen Bereich liegt, sind darunter nur wenige diagnostische Verfahren (vgl. die Anzahl der tatsächlich relevanten Artikel in Tabelle 1.1). Es handelt sich überwiegend um Beiträge mit Fokus auf die Zielgruppe, jedoch nicht auf die Darstellung von Verfahren an sich. Insbesondere für den Bereich Arbeitsmarkt zeigt die Suche kaum Treffer an. Bereichsübergreifend fällt außerdem auf, dass es sich um relativ neue Beiträge handelt – was als Hinweis darauf gedeutet werden kann, dass es sich um ein bislang wenig erforschtes Gebiet handelt, das jedoch zunehmend auf Interesse stößt (Goth & Severing, 2017).

Die oben skizzierten Rechercheergebnisse geben einen ersten Eindruck von der Verfügbarkeit diagnostischer Verfahren für Personen mit Migrationshintergrund in Deutschland. Sie spiegeln darüber hinaus den derzeitigen Forschungsbedarf,

⁴ Siehe hierzu die Ausführungen in Abschnitt 1.2.

⁵ Gesucht wurde in Büchern, Zeitschriftenaufsätzen sowie Sammelwerksbeiträgen ohne zeitliche Beschränkung (weitere Details zur Recherche sind in Tabelle 1.1 aufgeführt).

aber auch die laufenden Forschungsaktivitäten in Deutschland (IMIS, 2017⁶; Johansson, Schiefer & Andres, 2016) wider.

Eine weitere Suche im Bestand des Leibniz-Zentrums für psychologische Information und Dokumentation (ZPID)⁷ bestätigt die Ergebnisse der Suchmaschine PSYINDEX: Es liegen tatsächlich nur sehr wenige Verfahren vor, die für die in Deutschland lebende Bevölkerung mit Migrationshintergrund eingesetzt werden können bzw. die entsprechend normiert wurden. Darüber hinaus existiert bislang noch keine Übersicht über die zur Verfügung stehenden Instrumente für diese Zielgruppe. Praktikerinnen und Praktiker verfügen bei ihrer Arbeit in der Regel nicht über die notwendigen Ressourcen, um zeitaufwändig nach passenden Verfahren zu recherchieren. Insbesondere bei der Platzierung von Flüchtlingen im deutschen Bildungssystem und auf dem Arbeitsmarkt ergibt sich zudem das Problem, dass viele Geflüchtete über keine Dokumente verfügen (Brücker, Rother & Schupp, 2016), um ihren Bildungsweg bzw. Bildungsabschluss sowie ihre Arbeitserfahrung nachzuweisen. Hierbei stellt sich nicht nur die Problematik der Überprüfung der von den Zugewanderten gemachten Angaben, der angemessenen Übersetzung der Bildungszeugnisse und der adäquaten Einordnung in das deutsche Bildungs- und Berufssystem, sondern auch die Anpassung ihrer formalen Qualifikation an die Anforderungen in Deutschland.⁸ Hierzu sind zwar inzwischen Lösungsansätze entwickelt worden, die jedoch noch nicht ausreichend bekannt gemacht worden sind.

Ziel dieses Buchprojekts ist es deshalb, eine Bestandsaufnahme der vorhandenen diagnostischen Messinstrumente zur Erfassung des psychischen Wohlbefindens (klinische Diagnostik) sowie sprachbezogener (Sprachstanddiagnostik), schulischer (Leistungsstanddiagnostik) und arbeitsmarkbezogener Kompetenzen (z. B. berufliche Eignungstests, Gleichwertigkeitsfeststellung für im Ausland erworbene Qualifikationen) für Personen mit Migrationshintergrund vorzulegen. Das Handbuch richtet sich nicht nur an Praktikerinnen und Praktiker, sondern gleichermaßen auch an Wissenschaftlerinnen und Wissenschaftler: Um angemessen mit den Herausforderungen umgehen zu können, die sich heute und zukünftig stellen, muss noch viel Forschungsarbeit in diesem Themenfeld geleistet werden.

6 Unter der Leitung von A. Pott, J. Oltmer und C. Schetter wird das Forschungsprojekt „Flucht: Forschung und Transfer. Flüchtlingsforschung in der Bundesrepublik Deutschland“ (Laufzeit: 2016–2018) am Institut für Migrationsforschung und Interkulturelle Studien der Universität Osnabrück durchgeführt. Weitere Informationen unter: https://www.imis.uni-osnabrueck.de/forschung/flucht_und_fluechtlinge/flucht_forschung_und_transfer.html.

7 Die Datenbank beinhaltet etwa 700 Tests und Fragebögen aus dem Bestand der Psychologischen Testothek sowie ca. 4.000 Verfahren aus allen Bereichen der Psychologie sowie aus verwandten Disziplinen. Quelle: <https://www.zpid.de/index.php?wahl=products&uwahl=frei&uuwahl=trier>

8 Vgl. das Kapitel 11 von Atanassov & Erbe.

Tabelle 1.1: Recherche von diagnostischen Verfahren in Suchportal PSYNDEX (Ovid) mit Berücksichtigung von Testverfahren

Erfassungsdatum	Gesamtzahl der gefundenen Artikel	Anzahl der relevanten Artikel	Zeitraum	Bereich	Schlagwörter
25.10.2016	149	5	2005-2015	Entwicklungsdiagnostik	(Migrant or Migration or Flüchtling or Geflüchtete) and (Test or Diagnostik or diagnostische or Fragebogen or Verfahren or Erhebungsinstrument or Instrument or Messung) and (Kinder or Kind or Jugendlichen) and (Entwicklungsdiagnostik or Entwicklung or motorische or kognitive or Intelligenz)
25.10.2016	90	7	2000-2015	Sprachstandsdiagnostik	(Migrant or Migration or Flüchtling or Geflüchtete) and (Test or Diagnostik or diagnostische or Fragebogen or Verfahren or Erhebungsinstrument or Instrument or Messung) and (Kinder or Kind or Jugendlichen) and (Sprache or Sprachstandsdiagnostik or Sprachdiagnostik or Kompetenz or Kenntnisse or Fähigkeiten)
25.10.2016	92	5	2000-2015	Leistungsstandsdiagnostik	(Migrant or Migration or Flüchtling or Geflüchtete) and (Test or Diagnostik or diagnostische or Fragebogen or Verfahren or Erhebungsinstrument or Instrument or Messung) and (Kinder or Kind or Jugendlichen) and (Leistungsstandsdiagnostik or Leistungsstand or Leistung or Kompetenz or Kenntnisse or Fähigkeiten or Lesen or mathematische)
25.10.2016	65	2	2007-2009	Klinische Diagnostik bei Kindern und Jugendlichen	(Migrant or Migration or Flüchtling or Geflüchtete) and (Test or Diagnostik or diagnostische or Fragebogen or Verfahren or Erhebungsinstrument or Instrument or Messung) and (Kinder or Kind or Jugendlichen) and (klinisch or klinische or Wohlbefinden or Trauma or Depression or Depressivität or Selbstwert or Angst)

Tabelle 1.1: Fortsetzung

Erfassungsdatum	Gesamtzahl der gefundenen Artikel	Anzahl der relevanten Artikel	Zeitraum	Bereich	Schlagwörter
25.10.2016	2	0	1994-2007	Klinische Diagnostik bei Erwachsenen	(Migrant or Migration or Flüchtling or Geflüchtete) and (Test or Diagnostik or diagnostische or Fragebogen or Verfahren or Erhebungsinstrument or Instrument or Messung) and (Erwachsen or Erwachsener) and (klinisch or klinische or Wohlbefinden or Trauma or Depression or Depressivität or Selbstwert or Angst)
25.10.2016	5	0	2003-2010	Arbeitsmarkt	(Migrant or Migration or Flüchtling or Geflüchtete) and (Test or Diagnostik or diagnostische or Fragebogen or Verfahren or Erhebungsinstrument or Instrument or Messung) and (Erwachsen or Erwachsener) and (Arbeitsmarkt or Arbeitsplatz or Arbeit or Beruf or Eignungstest or Eigenschaftsdiagnostik or Kompetenz or Kompetenzen or beruflich or berufliche or informell erworbener Kompetenzen or non-formal erworbener Kompetenzen or formale or non-formale or informelle)

Anmerkung: Eine Freitextsuche ergab keine weiteren Treffer.

1.2 Zielbevölkerung mit Migrationshintergrund: Definition und Herausforderung für die Diagnostik

Das vorliegende Handbuch befasst sich mit diagnostischen Verfahren zur Testung von Personen mit Migrationshintergrund in Deutschland. Als Personen mit Migrationshintergrund (oder auch Zuwanderungshintergrund) werden nachfolgend Individuen bezeichnet, die selbst oder deren Vorfahren aus einem anderen Land zugewandert sind. Mit Migrantinnen und Migranten (oder auch: Zuwanderinnen und Zuwanderer bzw. Zugewanderte) sind – in Abweichung von der amtlichen Statistik – hingegen nur Personen gemeint, die über eigene Migrationserfahrung verfügen und somit Angehörige der ersten Migrantengeneration sind (Sauer & Brinkmann, 2016, S. 8).⁹ Insbesondere die letzte Gruppe steht im Mittelpunkt des vorliegenden Handbuches, denn für Migrantinnen und Migranten können rein deutschsprachige Verfahren in der überwiegenden Zahl der Fälle nicht eingesetzt werden, während dies für in Deutschland geborene und aufgewachsene Personen mit Migrationshintergrund – wenn auch unter Umständen mit Einschränkungen – in der Regel möglich ist.¹⁰ Dementsprechend behandeln die folgenden Buchbeiträge primär Verfahren, die für Personen gedacht sind, die über keine oder nicht ausreichende deutsche Sprachkenntnisse verfügen oder bei denen noch keine Akkulturation in die deutsche Gesellschaft stattgefunden hat.

Die größte Gruppe der Personen mit Migrationshintergrund in Deutschland bilden die (Alt-)Aussiedler und Spätaussiedler (zusammen ca. 4,5 Mio.); diese stammen vorwiegend aus der ehemaligen Sowjetunion (insbesondere Kasachstan und Russische Föderation) sowie aus Polen und Rumänien (Sauer & Brinkmann, 2016, S. 9–12). Einen ebenfalls großen Anteil an der Bevölkerung mit Migrationshintergrund in Deutschland stellen Türkeistämmige (überwiegend ethnische Türken und Kurden); 2015 waren dies knapp 2,9 Mio. bzw. 16,7% aller Personen mit Migrationshintergrund (Bundesamt, 2016, S. 161–163). Betrachtet man hingegen nur die Gruppe der Geflüchteten, dann zeigt sich beispielsweise für das Jahr 2016, dass die Herkunftsländer Syrien, Irak, Afghanistan, Eritrea sowie Iran die fünf Länder mit den meisten erstinstanzlich positiven Entscheidungen über Asylanträge sind (Eurostat, 2017).¹¹ In diesen Fällen liegt also ein anerkannter Aufenthaltsstatus vor und es wird davon ausgegangen, dass diese Personen die primäre Zielgruppe der zukünftigen diagnostischen Arbeit in unterschiedlichen Bereichen (z. B. Platzierung

⁹ Hierzu gehören auch Personen, die in die rechtliche Kategorie der Flüchtlinge oder Geflüchteten fallen. Für eine ausführlichere Darstellung der Definition in der amtlichen Statistik bzw. im Mikrozensus wird auf entsprechende Literaturquellen verwiesen (Brinkmann & Maehler, 2015, S. 7–10).

¹⁰ In den jeweiligen Beiträgen werden Abweichungen von diesen begrifflichen Verwendungen oder spezifische Zielgruppen gekennzeichnet.

¹¹ Tabellen verfügbar unter: <http://appsso.eurostat.ec.europa.eu>

auf dem Arbeitsmarkt oder psychodiagnostische Beratung) bilden werden. Dabei ist insbesondere zu berücksichtigen, dass im Vergleich zu Einheimischen ein großer Anteil der in den letzten Jahren nach Deutschland Geflüchteten über einen geringeren Bildungsstand verfügt. Beispielsweise berichten Brücker, Rother und Schupp (2016, S. 6, 55, 63), dass nur 58 % der erwachsenen Geflüchteten über den europäischen Bildungsstandard von zehn Jahren oder mehr in Schule, Ausbildung und Hochschule verfügen, während dies für 88 % der deutschen Wohnbevölkerung der Fall ist. Die Autoren weisen jedoch darauf hin, dass ein Zusammenhang mit der Unterbrechung der Bildungsbiografien durch Krieg, Verfolgung oder Flucht besteht, und dass darüber hinaus die Bildungssysteme nur eingeschränkt vergleichbar sind.

Die Anwendung diagnostischer Verfahren bei Migrantinnen und Migranten dient zum einen der adäquaten Platzierung – z. B. von Kindern in der Schule oder Erwachsenen auf dem Arbeitsmarkt – und zum anderen der Einleitung einer therapeutischen Behandlung von möglicherweise vielen traumatisierten neuen Bürgerinnen und Bürgern (Hettich, 2017). Dabei besteht die Herausforderung nicht nur im beschriebenen Mangel an diagnostischen Verfahren, sondern auch darin, dass die derzeitigen Einwanderergruppen aus Herkunftsländern mit hoher Sprachdiversität stammen. Tabelle 1.2 gibt eine Übersicht über die Sprachdiversität der Geflüchteten aus den fünf häufigsten Herkunftsländern in Deutschland. Bei dieser Auflistung wird deutlich, dass die derzeitigen Migrantinnen und Migranten eines Herkunftlandes nicht zwangsläufig dieselbe Sprache sprechen. So ist beispielsweise Hocharabisch in Syrien zwar die standardisierte schriftliche Form, sie spielt mutmaßlich jedoch kaum eine Rolle bei Personen, die nicht die Schule besucht haben bzw. die in bestimmten Regionen eines Landes aufgewachsen sind. Durch die arabische Diglossie¹² (Hayatli & Lerner, 2017) – z. B. in Syrien und dem Irak – müssten im Prinzip mehrere arabische Versionen eines diagnostischen Verfahrens erstellt werden, um tatsächlich eine valide Messung eines Konstruktes zu erzielen. Eine einzige Version in Hocharabisch wird möglicherweise von Personen mit geringen Grundkompetenzen nicht verstanden. Ebenso liegen unterschiedliche kurdische Sprachen und Alphabete vor: Nordkurdisch (Kurmandschi) in lateinischer Schrift sowie Zentralkurdisch (Sorani) in persischer Schrift.

Die Auflistungen der unterschiedlichen Sprachversionen in Tabelle 1.2 zeigen nicht nur die praktische Herausforderung in der Bereitstellung und Verfügbarkeit diagnostischer Verfahren in unterschiedlichen Sprachen, sondern ebenso die verschiedenen kulturellen Hintergründe und Ethnien, die mit der Sprachenvielfalt einhergehen. Hieraus ergibt sich besonders dringlich die Notwendigkeit einer kultursensitiven Diagnostik (die sprachliche und kulturelle Prägungen des zu mes-

12 Hiermit ist gemeint, dass in einem arabischsprachigen Land zwei unterschiedliche Versionen des Arabischen gesprochen werden: also die Koexistenz von Dialekt und Standardsprache, oder von der gesprochenen Volkssprache und der geschriebenen Hochsprache.

senden Konstrukts berücksichtigt), die über eine Übersetzung der Verfahren hinausgeht und die Messäquivalenz¹³ der eingesetzten Verfahren voraussetzt.

Tabelle 1.2: Sprachdiversität der fünf größten Herkunftsländer der Geflüchteten in Deutschland

Land	Sprache	Unterformen/Dialekte
Syrien	Arabisch	Nordisch-Levantine Najdi
	Kurdisch	Nordkurdisch (Kurmandschi)
Irak	Arabisch	Mesopotamisch Nordisch-Mesopotamisch Najdi
	Azerbaijani	Südlich
	Kurdisch	Nordkurdisch (Kurmandschi) Zentralkurdisch (Sorani)
Afghanistan	Dari (Afghanisches Persisch)	
	Hazaragi	
	Paschtu	Nördlich Südlich
	Turkmenen Usbekisch	Südlich
Eritrea	Tigré	
	Tigrigna	
Iran	Arabisch	Mesopotamisch
	Azerbaijani	Südlich Nördlich
	Bakhtiâri	
	Gilaki	
	Iranisches Persisch (Westliches Farsi)	
	Kurdisch	Zentralkurdisch Südkurdisch
	Laki	
	Luri	Nordisch
	Mazandarani	
Turkmenen		

Anmerkung: In Anlehnung an Hayatli & Lerner, 2017.

13 Zur Messäquivalenz vgl. Maehler & Schmidt-Denter (2013) und van de Vijver & Leung (2011).

1.3 Gliederung des Buches

Das hier vorgelegte Handbuch trägt die derzeit verfügbaren diagnostischen Verfahren – die auf die Bevölkerung mit Migrationshintergrund ausgerichtet sind – für unterschiedliche Anwendungsbereiche zusammen. Alle Beiträge dieses Buches durchliefen ein Peer-Review-Verfahren. Der erste Abschnitt des Buches beinhaltet zunächst einige theoretische Grundlagen: Kapitel 2 fasst die Hauptannahmen psychologischer Tests zusammen und stellt verschiedene Testarten vor. In diesem Zusammenhang wird insbesondere auf die Gütekriterien zur Beurteilung psychologischer Testverfahren eingegangen. Da für die Konstruktion valider Testverfahren eine adäquate Übersetzung und Adaptation eine zentrale Herausforderung darstellt, steht dieses Thema im Mittelpunkt des dritten Kapitels. Im Rahmen einer anwendungsnahen Einführung in geeignete Übersetzungsprozeduren werden im entsprechenden Kapitel Modelle transferfähiger Praxis abgeleitet. Der letzte Beitrag des Einführungsteils (Kapitel 4) behandelt Herausforderungen bei der Erfassung und beim Vergleich des Bildungsniveaus von Migrantinnen und Migranten mit der Wohnbevölkerung. Die adäquate Einschätzung des Bildungsniveaus ist beispielsweise essentiell für die Platzierung von Migrantinnen und Migranten im Bildungssystem und auf dem Arbeitsmarkt. In diesem Kapitel werden zwei zentrale Herausforderungen diskutiert: zum einen das Fehlen formaler Nachweise (Zeugnisse) über die Bildungslaufbahn, und zum anderen die mangelnde Vergleichbarkeit von vorhandenen Abschlüssen aus dem Ausland sowie aus Deutschland. Beide Problematiken werden derzeit in Bezug auf die Integration von Migrantinnen und Migranten auf dem Arbeitsmarkt stark diskutiert, und sie stehen im Mittelpunkt verschiedener Forschungsprojekte. Dieses Thema wird daher nochmals im letzten Abschnitt des Buches aufgegriffen.

Der zweite Abschnitt des Herausgeberwerkes befasst sich zunächst mit pädagogisch-psychologischer Diagnostik bei Kindern und Jugendlichen. Kapitel 5 ist der Entwicklungsdiagnostik für Kinder im Alter von null bis sechs Jahren gewidmet. Es stellt etablierte Entwicklungstests aus dem deutschen Sprachraum dar und bewertet diese vor dem Hintergrund der besonderen Anforderungen bei der Testung von Kindern mit Migrationshintergrund.

Während im Rahmen der Entwicklungsdiagnostik auch Verfahren zum Einsatz kommen, die sprachfrei sind und somit keine Übersetzung für Migrantinnen und Migranten benötigen (z. B.: Nonverbaler Intelligenztest/SON-R 2½-7), steht im Kapitel zur Sprachstanddiagnostik (6) genau dieser Aspekt im Fokus: Existierende Testverfahren im Bereich mündlich-sprachlicher Fähigkeiten werden im Hinblick auf Besonderheiten bei Mehrsprachigkeit diskutiert. Hierbei werden für mehrsprachige Kinder konzipierte Verfahren vorgestellt.

Weiterführend – und ebenfalls mit einem starken Fokus auf den Bereich Sprache – behandelt das letzte Kapitel (7) dieses Abschnittes den Altersbereich der Schul-

kinder; er stellt Verfahren zur Diskussion, die im Rahmen der Leistungsstanddiagnostik in Deutschland zum Einsatz kommen. Hauptgegenstand dieses Kapitels sind die Herausforderungen, die mit einer validen und diskriminierungsfreien Diagnostik kognitiver Leistungen von Kindern mit Migrationshintergrund einhergehen. Die damit verbundenen Anforderungen an die Testverfahren bzw. an die Testsituation werden zunächst dargestellt und anschließend diskutiert. Ein Schwerpunkt liegt dabei einerseits auf der Rolle der Sprache bei der Konzeptualisierung bzw. Testung schulischer Leistungen (z. B. mathematischer Kompetenzen), und andererseits auf der Frage, inwiefern „sprachfreie“ Testungen kognitiver Leistungen überhaupt möglich bzw. wünschenswert sind. Die Darstellung beinhaltet neben Befunden zur Bedeutung sprachlicher Anforderungen an Testaufgaben Studien zur Wirksamkeit sprachlicher Vereinfachungen als Testakkommodationen für Personen mit Migrationshintergrund. Hierzu werden relevante Testverfahren, die separate Normen für Personen mit nicht deutscher Muttersprache bereitstellen, vorgestellt.

Abschnitt III fokussiert auf Testverfahren im Rahmen der klinischen Diagnostik bei Personen mit Migrationshintergrund. Dieser Abschnitt berücksichtigt ebenfalls unterschiedliche Altersgruppen und gliedert sich in die Klinische Diagnostik bei Kindern und Jugendlichen (Kapitel 8) sowie bei Erwachsenen (Kapitel 9). Kultursensitive klinische Diagnostik steht im Mittelpunkt beider Buchbeiträge, die – neben der personalen Diagnostik sowie test- und fragebogengestützter Diagnostik – die Frage nach migrations- und kulturspezifischen Unterschieden in der Diagnostik thematisieren.

Anschließend wird im letzten Buchabschnitt IV die Diagnostik zur Platzierung von Erwachsenen auf dem Arbeitsmarkt dargestellt. Kapitel 10 gibt zunächst einen kurzen Überblick über die jetzigen Strukturen und Zuständigkeiten in Bezug auf diagnostische Instrumente, mit deren Hilfe Kompetenzen von Migrantinnen und Migranten sichtbar gemacht werden sollen. Es folgt die Darstellung des Prozesses der Gleichwertigkeitsfeststellung für Personen mit im Ausland erworbenen Berufsqualifikationen (Kapitel 11). Im Gegensatz zur pädagogischen oder klinischen Diagnostik kommen hierbei keine klassischen Tests, sondern für die psychologische Diagnostik eher ungewöhnliche Verfahren wie die Dokumentenanalyse zum Einsatz. Das Kapitel 12 fokussiert sodann auf Verfahren zur Erfassung von im Ausland formal und informell erworbenen beruflichen Kompetenzen. Anhand simulationsbasierter authentischer Testumgebungen wird ein Modell präsentiert, mit dem entsprechende Kompetenzen von Migrantinnen und Migranten erfasst werden können. Abschließend werden wiederum berufliche Eignungstests für die Migrationsbevölkerung in Deutschland vorgestellt (Kapitel 13).

1.4 Zielgruppe

Neben Therapeutinnen und Therapeuten aus der Praxis soll dieses Buch für Schul- und Berufsberatungsstellen, Arbeitgeberinnen und Arbeitgeber sowie Wissenschaftlerinnen und Wissenschaftler ein wichtiger Leitfad sein.

1.5 Hinweise

Zur besseren Übersicht der vorhandenen Instrumente über die unterschiedlichen Kapitel hinweg, fassen alle Beiträge die ausgewählten Verfahren anhand einer standardisierten Tabelle zusammen. Tabelle 1.3 enthält zunächst grundlegende Informationen über die Autorinnen und Autoren der jeweils herangezogenen Publikationen, das Jahr der Veröffentlichung sowie Anmerkungen zur Konzeption des jeweiligen Verfahrens. Unter Letzterem wird zum Beispiel dargestellt, auf welcher Theorie der Test basiert oder welches Konstrukt erfasst werden soll. Des Weiteren werden Anzahl und Bezeichnung der Untertests dargestellt. Schließlich findet in

Tabelle 1.3: Tabellenvorlage für dargestellte Testverfahren

Testverfahren:						
Autoren:						
Konzeption:						
Untertests:						
Beurteilung der psychometrischen Qualität:						
Alter	Test-sprache	Test-aufbau	zeitlicher Aufwand	Normierungs-stich-probe	Ergän-zende Informa-tionen	Bezug

den Übersichtsinformationen auch eine Beurteilung der psychometrischen Qualität statt. Dabei wird eingeschätzt, ob der Test ausreichend empirisch validiert ist, ob die Kriteriumsvalidität dokumentiert wurde, und inwieweit die Normwerte aktuell oder bereits veraltet sind. Die zweite Tabellenhälfte enthält spezifische Angaben zum berücksichtigten Altersbereich, zur Testsprache(n), zum Testaufbau, zeitlichen Aufwand, zur zugrunde gelegten Normierungsstichprobe sowie ergänzende Informationen (z. B. Informationen darüber, ob es sich um kulturunabhängiges Testmaterial handelt). Zum Schluss wird dargestellt, bei welchen Stellen die vorgestellten Verfahren bezogen werden können (z. B. Testzentrale Göttingen).

Literatur

- Brinkmann, H. U. (2016). Soziodemographische Zusammensetzung der Migrationsbevölkerung. In H. U. Brinkmann & M. Sauer (Hrsg.), *Einwanderungsgesellschaft Deutschland. Entwicklung und Stand der Integration* (S. 145–175). Wiesbaden: Springer VS.
- Brinkmann, H. U. & Maehler, D. B. (2015). Einführung in das Methodenbuch. In D. B. Maehler & H. U. Brinkmann (Hrsg.), *Methoden der Migrationsforschung. Ein interdisziplinärer Forschungsleitfaden* (S. 1–16). Wiesbaden: Springer VS.
- Brücker, H., Rother, N. & Schupp, J. (2016). *IAB-BAMF-SOEP-Befragung von Geflüchteten: Überblick und erste Ergebnisse* (IAB Forschungsbericht 14/2016). Zugriff am 05.05.2017. Verfügbar unter <http://doku.iab.de/forschungsbericht/2016/fb1416.pdf>
- Bundesamt für Migration und Flüchtlinge (2016). *Migrationsbericht des Bundesamtes für Migration und Flüchtlinge im Auftrag der Bundesregierung. Migrationsbericht 2015*, Nürnberg.
- Eurostat (2017). *Asylentscheidungen in der EU. EU-Mitgliedstaaten erkannten im Jahr 2016 über 700 000 Asylbewerber als schutzberechtigt an* (Pressemitteilung 70/2017 – 26. April 2017). Zugriff am 05.05.2017. Verfügbar unter <http://ec.europa.eu/eurostat/documents/2995521/8001720/3-26042017-AP-DE.pdf/08ccec8e-7b7e-4d9f-a5b6-3bc807fd0d4f>
- Goth, G. G. & Severing, E. (Hrsg.). (2017). *Asylsuchende und Flüchtlinge in Deutschland: Erfassung und Entwicklung von Qualifikationen für die Arbeitsmarktintegration*. Bielefeld: Bertelsmann.
- Hayatli, M. & Lerner, M. (2017, März). *Linguistic, cultural, and social considerations when interviewing new MENA populations*. Paper presented at GESIS Symposium on „Surveying the migrant population: Consideration of linguistic and cultural aspects“, Mannheim.
- Hettich, N. (2017). Access to Health Services for Refugees in Germany. In A. Korntheuer, P. Pritchard & D. B. Maehler (Eds.), *Structural Context of Refugee Integration in Canada and Germany* (GESIS Series 15). Köln: GESIS – Leibniz Institute for the Social Sciences.
- IMIS (Institut für Migrationsforschung und Interkulturelle Studien der Universität Osnabrück). (2017). *Wissenschaftler lassen die Notwendigkeit der Stärkung der Flucht- und Flüchtlingsforschung deutlich werden* (Pressemitteilung). Zugriff am 05.05.2017 unter https://www.imis.uni-osnabrueck.de/fileadmin/1_IMIS/Pressemitteilungen/2017-01-19_PM_FFT_St%C3%A4rkung_FFF.pdf
- Johannson, S., Schiefer, D. & Andres, N. (2016). *Was wir über Flüchtlinge (nicht) wissen. Der wissenschaftliche Erkenntnisstand zur Lebenssituation von Flüchtlingen in Deutschland. Eine Expertise im Auftrag der Robert Bosch Stiftung und des SVR-Forschungsbereichs*. Berlin. Zugriff am 26.07.2017. Verfügbar unter http://www.bosch-stiftung.de/content/language1/downloads/RBS_SVR_Expertise_Lebenssituation_Fluechtlinge.pdf

- Maehler, D.B. & Schmidt-Denter, U. (2013). *Migrationsforschung in Deutschland: Leitfaden und psychologische Messinstrumente*. Wiesbaden: Springer VS. <http://doi.org/10.1007/978-3-531-19245-1>
- OECD. (2016). *International migration outlook. Continues trends in international migration*. Paris: OECD Publishing. http://doi.org/10.1787/migr_outlook-2016-en
- Sauer, M. & Brinkmann, H.U. (2016). Einführung: Integration in Deutschland. In H.U. Brinkmann & M. Sauer (Hrsg.), *Einwanderungsgesellschaft Deutschland. Entwicklung und Stand der Integration* (S. 1–21). Wiesbaden: Springer VS.
- van de Vijver, F.J.R. & Leung, K. (2011). Equivalence and bias: A review of concepts, models, and data analytic procedures. In D. Matsumoto & F.J.R. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 17–45). New York: Cambridge University Press.

2 Testtheoretische Grundlagen der psychologischen Diagnostik: Ein Überblick

Alexandra Shajek, Débora B. Maehler & Heinz Ulrich Brinkmann

2.1 Einleitung

Das vorliegende Handbuch behandelt diagnostische Verfahren, die für in Deutschland lebende Migrantinnen und Migranten empfohlen werden. Diesen diagnostischen Verfahren liegen gemeinsame Grundannahmen der psychologischen Testdiagnostik zugrunde, die im Folgenden kurz skizziert werden.¹ Dabei wird zunächst erklärt, was im Allgemeinen unter einem psychologischen Test verstanden wird (2.1.1), welche Anwendungsgebiete in Frage kommen (2.1.2), mit welchem Ziel psychologische Tests eingesetzt werden (2.1.3), aus welchen Bestandteilen ein psychologischer Test besteht (2.1.4) und welche Testarten unterschieden werden (2.1.5). Hieran schließt sich eine kurze Beschreibung der Gütekriterien an, die zur Beurteilung von Testverfahren herangezogen werden (2.2).

2.1.1 Was ist ein psychologischer Test?

Allgemein wird unter einem psychologischen Test ein standardisiertes Instrument bzw. Verfahren verstanden, mit dem ein oder mehrere psychologische Merkmal(e) (z. B. Persönlichkeitseigenschaften, Motivation, Einstellungen) erfasst werden.² Diese Merkmale werden in der Regel als *Konstrukte* bezeichnet.

-
- 1 Für ausführlichere Einführungen in die Testtheorie, auf denen auch die folgenden Ausführungen (sofern nicht anders gekennzeichnet) beruhen, vgl. Schmidt-Atzert & Amelang (2012) sowie Moosbrugger & Kelava (2012).
 - 2 Für eine Übersicht über verschiedene Definitionsansätze und -merkmale vgl. Schmidt-Atzert & Amelang (2012).

2.1.2 Was sind die Anwendungsgebiete psychologischer Tests?

Psychologische Tests werden in den unterschiedlichsten Kontexten angewendet. Im vorliegenden Handbuch werden primär Testverfahren behandelt, die sich folgenden Bereichen zuordnen lassen:

1. *Klinische Diagnostik*: In der klinischen Diagnostik geht es in der Regel darum, mit den entsprechenden Tests psychische Krankheiten festzustellen bzw. auszuschließen. Die entsprechenden Verfahren werden beispielsweise eingesetzt, um eine Posttraumatische Belastungsstörung zu diagnostizieren.³
2. *Pädagogische Diagnostik*: Im Mittelpunkt der Pädagogischen Diagnostik steht die Analyse von Lernprozessen und ihren Voraussetzungen. Hierbei kommen diagnostische Verfahren beispielsweise dann zum Einsatz, wenn geklärt werden soll, ob eine Person an einer Lese-Rechtschreib-Schwäche leidet.⁴
3. *Sprachstanddiagnostik*: Eine besondere Rolle spielt in diesem Band die Sprachstanddiagnostik. Mit den entsprechenden Tests lassen sich beispielsweise die Deutschkenntnisse einer Person ermitteln.⁵
4. *Eignungsdiagnostik*: Die Eignungsdiagnostik spielt beispielsweise eine Rolle im Zusammenhang mit der Integration von Einwanderinnen und Einwanderern in den hiesigen Arbeitsmarkt. Verfahren aus diesem Bereich werden etwa eingesetzt, um festzustellen, ob eine Person, die über keine formalen Qualifizierungsdokumente verfügt, für eine zu besetzende Stelle geeignet ist.⁶

2.1.3 Mit welchem Ziel werden psychologische Tests durchgeführt?

Unabhängig vom konkreten Anwendungsbereich lassen sich psychologische Tests auch danach unterscheiden, mit welchem Ziel sie durchgeführt werden (Schmidt-Atzert & Amelang, 2012). Am häufigsten klassifiziert werden sie anhand folgender Merkmale:

1. *Status- vs. Veränderungsdiagnostik*: Eine *Statusdiagnostik* wird mit dem Ziel durchgeführt, einen momentanen Zustand zu beschreiben – in der Regel um anzuzeigen, ob bestimmte Maßnahmen (z. B. eine Sprachförderung) indiziert sind (Schmidt-Atzert & Amelang, 2012; Petermann & Macha, 2005). Mit einer *Veränderungsdiagnostik* lässt sich überprüfen, ob eine Maßnahme erfolgreich

3 Vgl. Kapitel 8: Busch, Leyendecker & Siefen und 9: Nesterko & Glaesmer.

4 Vgl. Kapitel 5: Macha & Petermann und 7: Haag, Heppt & Schipolowski.

5 Vgl. Kapitel 6: Reitenbach, Schastak & Rauch.

6 Vgl. Kapitel 12: Deutscher & Winther und 13: Krumm et al.

war. Sie kann entweder als Erfolgskontrolle nach Beendigung einer Maßnahme oder aber kontinuierlich im Rahmen einer Verlaufs- oder Prozessdiagnostik erfolgen (Schmidt-Atzert & Amelang, 2012).

2. *Selektion oder Modifikation*: Häufig wird ein diagnostisches Verfahren angewendet, um eine Passung zwischen bestimmten Bedingungen und Personen festzustellen (ebd.). Im Rahmen einer berufsbezogenen Eignungsdiagnostik⁷ soll durch die *Selektion* beispielsweise für eine ausgeschriebene Stelle eine geeignete Bewerberin bzw. ein geeigneter Bewerber ausgewählt oder für eine Person in einer Organisation die bestmögliche Position gefunden werden (Platzierung). Daneben steht im Rahmen der Eignungsdiagnostik jedoch auch manchmal eine *Modifikation* der Person oder der Bedingungen im Vordergrund: In diesem Fall zielt die Diagnostik darauf ab, zu beschreiben, wie sich eine Person (z. B. durch eine Schulung) verändern müsste, um die Anforderungen einer Stelle erfüllen zu können bzw. wie sich ein Arbeitsplatz verändern müsste, damit er von den zur Verfügung stehenden Personen erfolgreich ausgeführt werden kann (ebd.).

2.1.4 Wie ist ein psychologischer Test aufgebaut?

Ein psychologischer Text beinhaltet nach Petermann & Macha (2005) in der Regel drei Bestandteile:

Erstens ein *Testmanual* (auch Handbuch genannt), das die theoretischen Grundlagen des Verfahrens sowie dessen Entwicklung erläutert und der Anwenderin bzw. dem Anwender des Tests die Durchführung, Auswertung und Interpretation des Verfahrens so beschreibt, dass er/sie genau weiß, wie er/sie vorgehen muss (ebd.). Ferner beinhaltet das Manual eine Beschreibung der Qualität des Tests (anhand von Gütekriterien, vgl. Abschnitt 2.2) und die für den Test relevanten Normwerte (vgl. Abschnitt 2.2.2.6).

Zweitens das *Testmaterial*, das von der zu testenden Person bearbeitet wird (ebd.). Dieses kann je nach Test sehr unterschiedlich aussehen – so enthält ein Intelligenztest bspw. Bilder, die sortiert werden müssen oder Figuren, die in eine bestimmte Reihenfolge gebracht werden sollen. Die einzelnen Aufgaben in einem Test bzw. die einzelnen Aussagen in einem Fragebogen werden dabei als *Items* bezeichnet. Bei einem Set von Items, die ein bestimmtes Merkmal messen, wird von einer *Skala* gesprochen.

Drittens *Protokollblätter* zum Protokollieren der Testdurchführung sowie ggf. *Auswertungshilfen* wie Schablonen oder Messgeräte (ebd.).

7 Vgl. Kapitel 13: Krumm et al.

2.1.5 Welche Arten von psychologischen Tests gibt es?

In Abhängigkeit von dem zu messenden Merkmal werden im Allgemeinen unterschiedliche Arten von Tests voneinander abgegrenzt (Moosbrugger & Kelava, 2012). Zu den am häufigsten angewandten gehören:

Leistungstests: Diese messen Aspekte der kognitiven Leistungsfähigkeit einer Person und sind üblicherweise so konstruiert, dass die Antworten der Getesteten als logisch richtig oder falsch klassifiziert werden können (Moosbrugger & Kelava, 2012). Zu den bekanntesten Leistungstests gehören Aufmerksamkeits- und Konzentrationstests sowie Intelligenztests.

Persönlichkeitstests bzw. Persönlichkeitsfragebogen: Bei diesen wird nicht das kognitive Leistungsvermögen erfasst, sondern die zu testenden Personen werden aufgefordert, ihr typisches Verhalten einzuschätzen (Moosbrugger & Kelava, 2012). Neben klassischen Persönlichkeitstests werden solche Verfahren zu dieser Kategorie gezählt, die beispielsweise Interessen einer Person oder ihre Einstellungen erfassen (Moosbrugger & Kelava, 2012).

2.2 Gütekriterien für psychologische Tests

Um die Qualität in der Entwicklung, Bewertung und Anwendung psychologischer Tests sicherzustellen, sind von unterschiedlichen Organisationen Standards und Richtlinien erarbeitet worden, die zum Teil auch auf Deutsch vorliegen. Zu den bekanntesten gehören die Richtlinien der International Test Commission (ITC) (2016; s. a. Schmidt-Atzert und Amelang, 2012). Üblicherweise werden zur Beurteilung der Qualität eines psychologischen Tests bestimmte Qualitätskriterien herangezogen. In der Regel werden drei Hauptgütekriterien – Objektivität, Reliabilität und Validität – sowie verschiedene Nebengütekriterien unterschieden. Diese werden in den Abschnitten 2.2.1 und 2.2.2 kurz beschrieben.

2.2.1 Hauptgütekriterien

2.2.1.1 Objektivität

Ein psychologischer Test ist dann objektiv, wenn das Ergebnis des Tests unabhängig davon ist, wer den Test durchgeführt, ausgewertet und interpretiert hat (Moosbrugger & Kelava, 2012). Dementsprechend lassen sich drei Aspekte der Objektivität unterscheiden:

Die *Durchführungsobjektivität* ist dann gegeben, wenn die Testdurchführung unabhängig ist von „Besonderheiten, die der Diagnostiker in die Testsituation einbringt und die die Testleistung beeinflussen“ (Petermann & Macha, 2005, S. 38), bzw. „wenn das Testergebnis nicht davon abhängt, welcher Testleiter den Test mit der Testperson durchführt“ (Moosbrugger & Kelava, 2012, S. 9). Dies ist umso eher der Fall, je standardisierter der Test ist. Aus diesem Grund enthält das Testmanual eines psychologischen Tests genaue Anweisungen darüber, wie der Testleiter den Test durchzuführen hat.

Unter *Auswertungsobjektivität* wird verstanden, dass das Testergebnis auch nicht davon abhängt, welche Person das Ergebnis der Testdurchführung bewertet (Moosbrugger & Kelava, 2012). Um dies sicherzustellen, enthält das Testmanual detaillierte Auswertungsregeln. Besonders leicht möglich ist dies bei Multiple-Choice-Aufgaben, da hier in den Auswertungsregeln beispielsweise einfach die jeweils korrekte Antwort beschrieben wird.

Schließlich sollte ein psychologischer Test nicht davon abhängig sein, wer ein nach der Testdurchführung und -auswertung vorliegendes Ergebnis interpretiert – das heißt, verschiedene Testleiter müssen auf der Grundlage desselben Testergebnisses dieselben Schlussfolgerungen ziehen (Moosbrugger & Kelava, 2012, S. 10). In diesem Zusammenhang wird von *Interpretationsobjektivität* gesprochen.

2.2.1.2 Reliabilität

Mit der Reliabilität bzw. Zuverlässigkeit eines Tests wird dessen Messgenauigkeit beschrieben. Je höher die Reliabilität, desto weniger ist das Testergebnis mit Messfehlern behaftet. Die Reliabilität eines Tests wird üblicherweise in Form eines Koeffizienten dargestellt, der zwischen 0 und 1 variiert, je höher, desto reliabler ist der Test. Zur Bestimmung der Reliabilität existieren verschiedene Methoden, die unterschiedliche Realitätsaspekte abbilden:

Um zu überprüfen, wie zuverlässig ein Test bei ein und derselben Person zu unterschiedlichen Zeitpunkten zum selben Ergebnis kommt (vorausgesetzt, das zu messende Merkmal bleibt konstant), wird die *Retest-Reliabilität* ermittelt. Hierzu wird der Test mit denselben Testpersonen zu zwei Zeitpunkten durchgeführt und die Übereinstimmung der Ergebnisse als Korrelationskoeffizient angegeben (Moosbrugger & Kelava, 2012). Problematisch hierbei ist allerdings, dass die Ergebnisse in Abhängigkeit vom Zeitintervall zwischen den beiden Messungen durch unterschiedliche Einflüsse wie Übungs- und Erinnerungseffekte beeinflusst werden.

Die *Paralleltest-Reliabilität* wird ermittelt, indem dieselben Versuchspersonen zwei gleichwertige Tests oder auch vergleichbare Testversionen (Paralleltests) bearbeiten. Die Übereinstimmung zwischen den Ergebnissen der beiden Tests wird als Korrelationskoeffizient berechnet (Petermann & Macha, 2005, S. 40). Eine Ver-

allgemeinerung der Paralleltest-Methode ist die Konsistenzanalyse (Moosbrugger & Kelava, 2012): Hier werden nicht zwei Testhälften korreliert, sondern jedes einzelne Item (d. h. jede Aufgabe) wird als Testteil betrachtet und eine Übereinstimmung mit den anderen Items analysiert. Je höher die Korrelation der einzelnen Items untereinander, umso höher ist die interne Konsistenz (Cronbachs Alpha) des Verfahrens.

Da häufig keine zwei vergleichbaren Testversionen zur Verfügung stehen, kann die Reliabilitätsbestimmung auch mit der *Split-Half-Reliabilität* mittels Testhalbierungsmethode erfolgen (Moosbrugger & Kelava, 2012, S. 12). Der Test wird in zwei Hälften unterteilt und die Übereinstimmung der Ergebnisse wird wiederum – unter Berücksichtigung eines Korrekturfaktors – als Korrelationskoeffizient der beiden Testhälften bestimmt (ebd.).

2.2.1.3 Validität

Die Validität eines Tests gibt Auskunft darüber, wie sehr ein Test das Merkmal misst, das er messen soll (Moosbrugger & Kelava, 2012, S. 13). Bei der Validität handelt es sich um das wichtigste Gütekriterium, denn nur wenn der Test ausreichend valide ist, lassen sich aus dem Ergebnis eines Tests Schlussfolgerungen für das Verhalten von Personen außerhalb der Testsituation ziehen. Mit der Validität wird somit die Übereinstimmung des in der Testsituation Gezeigten mit dem Verhalten außerhalb der Testsituation angegeben. Um die Validität eines Tests möglichst umfassend beurteilen zu können, werden verschiedene Aspekte der Validität untersucht (ebd., S. 15):

Unter der *Inhaltsvalidität* wird verstanden, dass ein Test das Merkmal, das er erfassen soll, in den relevanten Aspekten vollständig erfasst (Moosbrugger & Kelava, 2012). Eine hohe Inhaltsvalidität liegt beispielsweise vor, wenn die Aufgaben eines Tests „einen unmittelbaren Ausschnitt aus dem Verhaltensbereich darstellen, über den eine Aussage getroffen werden soll (wenn z. B. Rechtschreibkenntnisse anhand eines Diktates überprüft werden oder die Eignung eines Autofahrers anhand einer Fahrprobe ermittelt wird)“ (Moosbrugger & Kelava, 2012, S. 15).

Die *Konstruktvalidität* gibt Auskunft darüber, „ob und inwieweit mit einem Test abgeleitete Hypothesen bestätigt werden können“ (Petermann & Macha, 2005, S. 41). Ist dies der Fall, ist davon auszugehen, dass die Messung eines psychologischen Merkmals (des Konstrukts) präzise gemessen wird. Um zu überprüfen, ob ein Test als konstruktvalide gelten kann, werden in der Regel zwei Aspekte bestimmt: 1. Mit einer Überprüfung der *konvergenten Validität* wird sichergestellt, dass die Ergebnisse eines Tests mit den Ergebnissen von Instrumenten, die dieselben oder verwandte Konstrukte messen, übereinstimmen. Darüber hinaus muss jedoch ebenfalls gewährleistet werden, dass ein Test mit Verfahren, die ein anderes Merkmal messen, nicht oder – im Fall eines verwandten Merkmals – nur leicht

korreliert ist. Lässt sich dies nachweisen, so verfügt ein Test über eine hohe diskriminante bzw. divergente Validität. 2. Die *Kriteriumsvalidität* gibt an, inwieweit sich ein Test eignet, um auf der Grundlage des Tests ein Kriterium – d. h. ein bestimmtes Verhalten oder Erleben von Personen – außerhalb der Testsituation vorhersagen zu können (Moosbrugger & Kelava, 2012). Liegt dieses Kriterium zeitgleich mit dem Testzeitpunkt, wird dabei von konkurrenzer bzw. Übereinstimmungsvalidität gesprochen, liegt es stattdessen in der Zukunft, wird von Vorhersagevalidität bzw. prognostischer Validität gesprochen (s. http://www.methodenpsychologie.de/kriterienbezogene_validitaet.html).

Mit der *Augenscheinvalidität* wird schließlich das Ausmaß bezeichnet, in dem ein Test auch Laien als valide erscheint (Moosbrugger & Kelava, 2012). Die Augenscheinvalidität ist zunächst einmal unabhängig von den oben beschriebenen, wissenschaftlich relevanten Aspekten der Validität; sie ist aber dennoch von Bedeutung, denn ein Instrument mit einer hohen Augenscheinvalidität genießt in der Regel eine hohe Akzeptanz bei der Versuchsperson.

2.2.2 Nebengütekriterien

2.2.2.1 Testökonomie

Das Gütekriterium der Testökonomie ist dann gegeben, wenn die mit einem Test verbundenen Ressourcen (in der Regel sind dies finanzielle und zeitliche) in einem angemessenen Verhältnis zum diagnostischen Erkenntnisgewinn stehen (Moosbrugger & Kelava, 2012, S. 21). Günstig auf die Testökonomie können sich beispielsweise die Durchführung mittels Computer oder die Möglichkeit des adaptiven Testens (hierbei werden jeder Testperson nur die Aufgaben vorgelegt, die für die Person den größten Informationsgewinn beinhalten) auswirken (ebd., S. 21).

2.2.2.2 Nützlichkeit

Das Gütekriterium der Nützlichkeit (Utilität) gilt dann als erfüllt, wenn „für das von ihm gemessene Merkmal praktische Relevanz besteht und die auf seiner Grundlage getroffenen Entscheidungen (Maßnahmen) mehr Nutzen als Schaden erwarten lassen“ (Moosbrugger & Kelava, 2012, S. 22).

2.2.2.3 Zumutbarkeit

Das Kriterium der Zumutbarkeit stellt sicher, dass die Testperson erstens weder absolut noch relativ zum Nutzen des Tests und zweitens nicht in psychischer, in körperlicher oder in finanzieller Hinsicht über Gebühr belastet wird (Moosbrugger & Kelava, 2012, S. 22).