

Karl-Heinz Renner

Differentielle Psychologie, Persönlichkeitspsychologie und Diagnostik

Kurseinheit 2:
Grundlagen psychologischer Diagnostik

Fakultät für
Psychologie

Das Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung und des Nachdrucks, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Kein Teil des Werkes darf in irgendeiner Form (Druck, Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung der FernUniversität reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Der Inhalt dieses Studienbriefs wird gedruckt auf Recyclingpapier (80 g/m², weiß), hergestellt aus 100 % Altpapier.

Einführung zum Kursbestandteil

Bitte lesen Sie diese Einführung sorgfältig durch!

Der Kursbestandteil 03418 umfasst den vorliegenden Studienbrief und einige Pflichtliteratur-Kapitel aus bewährten Lehrbüchern zur Psychologischen Diagnostik. Der Studienbrief führt in die Themen der Pflichtliteratur ein und verweist auf weitere diagnostische Inhalte, die im Modul 7 vermittelt werden. Dieser Studienbrief soll als roter Faden für die Psychologische Diagnostik dienen, wie sie im Modul 7 und in anderen Modulen gelehrt wird. Die Liste der Pflichtliteratur-Kapitel finden Sie in der *Moodle*-Umgebung zu diesem Kurs.

Soweit es die urheberrechtlichen Bestimmungen erlauben, werden wir die Pflichtliteratur in *Moodle* einstellen; wir möchten allerdings ausdrücklich empfehlen, dass Sie sich einige diagnostische Lehrbücher anschaffen. Gerade in den letzten Jahren sind sehr gute neue oder in neuen Auflagen erschienene Lehrbücher veröffentlicht worden, die Ihnen im weiteren Verlauf des Studiums, aber auch in Ihrer späteren praktischen Arbeit sehr nutzen werden

Insgesamt weisen die Textteile des vorliegenden Studienbriefs ein unterschiedliches Abstraktionsniveau auf: einige Passagen geben lediglich einführende Überblicke zur Pflichtliteratur, andere Abschnitte sind dagegen wesentlich detaillierter ausgeführt, insbesondere dann, wenn grundlegende Inhalte wiederholt oder die Pflichtliteratur ergänzt wird.

Wir empfehlen Ihnen sehr, die Inhalte dieses Kurses gemeinsam mit anderen Studierenden zu erarbeiten. Eine sehr gute Möglichkeit zu diesem kooperativen Lernen bietet die virtuelle Lehr-Lernumgebung *Moodle*, die an der FernUniversität in Hagen standardmäßig eingesetzt wird. Sie finden den Zugang zu Moodle über den entsprechenden Link: <https://moodle-psy.fernuni-hagen.de>

Sie werden innerhalb von *Moodle* zu einzelnen Abschnitten dieses Kurses Aufgaben finden, die Sie bearbeiten sollten. Sie werden dabei und auch bei anderen inhaltlichen Fragen zum Kurs von einem Tutor/einer Tutorin und von uns unterstützt. Ziel ist es aber, dass Sie zunächst in Kooperation mit anderen Studierenden die Inhalte und Aufgaben des Kurses bearbeiten. Jenseits von *Moodle* bieten einzelne Studienzentren eine weitere Möglichkeit zur Zusammenarbeit mit anderen Studierenden und z.T. auch mit MentorInnen.

Die Lektüre und Verarbeitung dieses und der anderen Kursbestandteile im Modul 7 sind mit Arbeit und Zeitaufwand verbunden! Bitte erarbeiten Sie die Inhalte und die Aufgaben möglichst kontinuierlich sowie plan- und regelmäßig. Ein „Einpauken“ des gesamten Stoffes in wenigen Tagen ist nicht möglich. Planen Sie deshalb bitte möglichst feste Zeiten ein, zu denen Sie sich mit dem Studienmaterial auseinandersetzen.

In diesem Kurs wird aus Gründen der besseren Lesbarkeit häufig das generische Maskulinum verwendet. In solchen Fällen sind aber immer alle Geschlechter gemeint.

Lernziele

Der vorliegende Kurs vermittelt Wissen zu grundlegenden Fragestellungen und Themen der Psychologischen Diagnostik. Im Einzelnen sollen Sie...

- die Wichtigkeit und Relevanz der Psychologischen Diagnostik begründen können (Kapitel 1).
- Psychologische Diagnostik definieren und explizieren können sowie Aufgabenfelder, Arten/Strategien, Dimensionen und Zielsetzungen der Psychologischen Diagnostik kennen (Kapitel 2).
- wichtige Meilensteine in der Geschichte der Psychologischen Diagnostik kennen (Kapitel 3).
- testtheoretische Grundlagen der Psychologischen Diagnostik und Methoden zur Interpretation von Testresultaten kennen und anwenden können (Kapitel 4 und 5).
- Wissen um diagnostische Verfahren und deren Klassifikation sowie Erhebungsstrategien erworben haben (Kapitel 6).
- ein handlungstheoretisches Modell des diagnostischen Prozesses sowie Modelle der diagnostischen Urteilsbildung gelernt haben (Kapitel 7).
- Grundzüge der klinisch-psychologischen Diagnostik und Intervention kennen (Kapitel 8).
- einige neuere Entwicklungen der Psychologischen Diagnostik kennen, kritisch beurteilen können und dabei Ihre Kompetenzen in der Lektüre englischer Fachliteratur vertiefen (Kapitel 9).

Wir wünschen Ihnen viel Erfolg bei der Erarbeitung des Kurses und viel Spaß beim Studieren!

Karl-Heinz Renner

Inhaltsverzeichnis

Einführung zum Kurs	1
Inhaltsverzeichnis	3
1 Einführung: Warum ist Psychologische Diagnostik wichtig?	4
2 Was ist Psychologische Diagnostik?	6
2.1 Definitionen und Aufgabenfelder	6
2.2 Arten, Dimensionen und Zielsetzungen	10
3 Stichpunkte zur Geschichte der Psychologischen Diagnostik	13
4 Testtheoretische Grundlagen der Diagnostik	15
5 Interpretation von Resultaten aus psychologischen Tests (Einzelfalldiagnostik)	16
6 Datenquellen und diagnostische Verfahren	19
6.1 Datenquellen der Psychologie	19
6.2 Weitere Klassifizierungen für Daten	25
6.3 Erläuterungen zum Begriff des Tests	26
6.4 Probleme der Veränderungsmessung	30
7 Grundlagen des diagnostischen Prozesses und der diagnostischen Urteilsbildung	33
8 Klinisch-psychologische Diagnostik und Intervention	36
9 Neuere Entwicklungen	38
9.1 Internetgestützte Diagnostik	38
9.2 Ambulatorisches Assessment	39
9.3 Renaissance der objektiven Persönlichkeitsdiagnostik	41
Literaturempfehlungen	42
Zeitschriften	42
Datenbanken und Internetquellen	42
Zitierte Literatur	43

1 Einführung: Warum ist Psychologische Diagnostik wichtig?

Psychologische Diagnostik als Methodenlehre und Querschnittsdisziplin

Warum ist die Psychologische Diagnostik so wichtig? Bei *jeder* psychologischen Fragestellung müssen Informationen bzw. Daten gesammelt werden. Mit welchen Methoden bei welchen Fragestellungen relevante Daten sachgerecht erhoben, weiterverarbeitet und interpretiert werden können, ist Gegenstand der Psychologischen Diagnostik. Die Psychologische Diagnostik wird als zentrale Methodenlehre im Dienste der Angewandten Psychologie definiert (s.u.). Daten werden aber selbstverständlich auch in der psychologischen Forschung benötigt. Da sich die Psychologische Diagnostik durch alle Grundlagen- und Anwendungsfächer zieht, kann sie auch als *Querschnittsdisziplin* verstanden werden.

Garbage In, Garbage Out

Die sachgemäße und valide Feststellung von psychischen Zuständen und Eigenschaften ist eine zentrale Voraussetzung für psychologische Forschung und für die Vorbereitung von praktisch-psychologischen Entscheidungen. Um die hohe Bedeutung der Diagnostik für die gesamte Psychologie deutlich zu machen, möchte ich auf eine „Weisheit“ zurückgreifen, die ursprünglich aus der Informatik und den Computerwissenschaften stammt: *Garbage In, Garbage Out* (abgekürzt GIGO). Auf unser Thema übertragen, bedeutet dieser Ausspruch: Wer ungültige und/oder ungenaue Daten bzw. wörtlicher genommen „Datenmüll“ erhebt, der wird auch ungültige Ergebnisse erzielen, sei es in der Forschung oder in der psychologischen Praxis. Psychologische Diagnostik produziert gewissermaßen den „Input“ für die Bearbeitung einer Fragestellung; die Qualität und die Gültigkeit des Outputs, also der Antwort auf eine psychologische Fragestellung, hängt nach dem GIGO-Prinzip ganz entscheidend von der Qualität und Gültigkeit des Inputs ab. Ein „Garbage Out“ kann dann gerade in der Psychologie mehr Leiden und finanzielle Verluste bedeuten. Damit z.B. eine psychische Störung erfolgreich behandelt werden kann, ist es in den meisten Fällen zunächst notwendig, genau zu diagnostizieren, welche Störung überhaupt vorliegt. Wird eine vorliegende psychische Störung nicht erkannt bzw. falsch diagnostiziert, dann besteht die Gefahr, dass auch die an die Diagnose anschließende Therapie ineffizient ist. Für Unternehmen kann die Fehlbesetzung einer Führungsposition, aber auch anderer Stellen hohe finanzielle Einbußen zur Folge haben. Deshalb ist es wichtig, den richtigen Mann bzw. die richtige Frau für eine gegebene Stelle auszuwählen; dazu ist eine fundierte Eigenschaftsdiagnostik notwendig.

Verminderung von Leiden, Verhinderung finanzieller Verluste und Ressourcenorientierung

Die Psychologische Diagnostik hilft aber nicht nur dabei, Leiden zu vermindern und finanzielle Verluste zu verhindern; sie ist nicht ausschließlich defizitorientiert. Vielmehr zielt Psychologische Diagnostik auch darauf ab, Ressourcen und Potentiale von Individuen und Gruppen aufzudecken. Diese ressourcenorientierte Sichtweise ist in den letzten Jahren im Zuge des Positive Psychology Movement zunehmend mehr auch für die Diagnostik fruchtbar gemacht worden. So haben Shane Lopez und C.R. Snyder bereits 2003 ein Buch zum Thema „Positive Psychological Assessment“ herausgegeben, in dem Ansätze zur Diagnostik menschlicher Stärken

(z.B. Optimismus, Hoffnung, Selbstwirksamkeit, Mut, Selbstwert, Empathie, Kreativität etc.) vorgestellt werden.

Es gibt noch einen weiteren Grund, warum Psychologische Diagnostik wichtig ist. Psychologen und Psychologinnen wird von Laien eine hohe diagnostische Kompetenz zugeschrieben. Ich möchte an dieser Stelle noch einmal die Anekdote wiederholen, die ich bereits im Kurs 3400 an den Anfang des Diagnostik-Kapitels gestellt habe: Vor einigen Jahren habe ich bei einer Geburtstagsfeier meines Bruders einige seiner Kommilitoninnen und Kommilitonen aus dem Fachbereich Jura kennen gelernt. Etwas verwundert musste ich feststellen, dass sie sich mir gegenüber zwar höflich, aber sehr zurückhaltend benahmen. Im Nachhinein erzählte mir mein Bruder, dass er seinen Mitstudierenden erzählt habe, ich sei Psychologe und würde sie sofort durchschauen – und offensichtlich haben sie das auch geglaubt. Der Satz: „Dann haben Sie mich schon durchschaut!?!“, fiel in der Vergangenheit noch einige Male, nachdem ich mich selbst als Psychologe „geoutet“ hatte. In der Tat scheinen Laien Psychologinnen und Psychologen einen „Röntgenblick“ zuzuschreiben. Auch wenn die Laienperspektive die tatsächlichen Möglichkeiten der Psychologischen Diagnostik (insbesondere das „Durchschauen auf den ersten Blick“) überschätzt, macht sie doch zumindest sehr deutlich, dass diagnostische Kompetenzen von einer ausgebildeten Psychologin bzw. einem ausgebildeten Psychologen erwartet werden.

Diagnostische Kompetenzen werden erwartet

In den mittlerweile auslaufenden Diplomstudiengängen Psychologie ist die Diagnostik mit Ausnahme der Testtheorie erst im Hauptstudium gelehrt worden. Dass die Diagnostik bei der Umstellung auf das konsekutive Studiensystem gemäß den Empfehlungen der Deutschen Gesellschaft für Psychologie nun als Pflichtmodul bereits im B.Sc. vermittelt wird, ist ein weiterer Beleg für die Bedeutung und Relevanz dieses Faches.

2 Was ist Psychologische Diagnostik?

Im Folgenden werden einige Abschnitte des Diagnostik-Kapitels aus Kurs 3400 wiederholt und teilweise erweitert. Einerseits handelt es sich um so grundlegende Aspekte der Psychologischen Diagnostik, dass sie in dieser Einführung nicht einfach deshalb weggelassen werden können, weil sie bereits im Modul 1 dargestellt wurden. Andererseits sind im Modul 7 etliche Studierende eingeschrieben, denen das erste Modul aufgrund vorhandener Studienleistungen aus dem Diplomstudengang Psychologie anerkannt wurde und die z.T. mit dem Vordiplom überhaupt erst mit dem siebten Modul Ihr B.Sc.-Studium an der FernUniversität beginnen. Auch um alle Belegerinnen und Beleger des siebten Moduls auf denselben Wissensstand zu bringen, ist es daher notwendig, erneut auf die nachfolgenden Inhalte einzugehen, die auch in dem zu diesem Kapitel gehörigen Pflichtliteratur-Kapitel behandelt werden.

2.1 Definitionen und Aufgabenfelder

Einige Definitionen

Was ist Psychologische Diagnostik? In einschlägigen Einführungs- und Lehrbüchern werden folgende Definitionen gegeben:

Psychodiagnostik ist eine Methodenlehre im Dienste der Angewandten Psychologie. Soweit Menschen die Merkmalsträger sind, besteht ihre Aufgabe darin, interindividuelle Unterschiede im Verhalten und Erleben sowie intraindividuelle Merkmale und Veränderungen einschließlich ihrer jeweils relevanten Bedingungen so zu erfassen, [dass] hinlänglich präzise Vorhersagen künftigen Verhaltens und Erlebens sowie deren evtl. Veränderungen in definierten Situationen möglich werden (Amelang & Schmidt-Atzert, 2006, S. 3).

Die Inhalte und Methoden der Psychologischen Diagnostik beziehen sich auf die *regelgeleitete Sammlung* und *Verarbeitung* von gezielt erhobenen Informationen, die für das Verständnis menschlichen Erlebens und Verhaltens bedeutsam sind. Aus den gewonnenen Informationen sollen Fragestellungen (eines Auftraggebers) bearbeitet und Entscheidungen getroffen werden. Die *Prinzipien der Entscheidungsfindung* müssen wissenschaftlichen Kriterien entsprechen. Die Schritte der Entscheidungsfindung müssen nachvollziehbar sein und die Schlussfolgerungen *ethischen Standards* genügen. Die Fragestellungen der Psychologischen Diagnostik können sich dabei auf die

- *Beschreibung* und
- *Klassifikation*,
- *Erklärung*,
- *Vorhersage* (Prognose) und
- *Evaluation* von *Zuständen* und/oder *Verläufen*

beziehen (Eid & Petermann, 2006, S. 16, Hervorhebungen im Original).

Psychologische Diagnostik ist das systematische Sammeln und Aufbereiten von Informationen mit dem Ziel Entscheidungen und daraus resultierende Handlungen zu begründen, zu kontrollieren und zu optimieren. Die Entscheidungen basieren auf einem komplexen Informationsverarbeitungsprozess. In diesem Prozess wird auf Regeln, Anleitungen, Algorithmen usw. zurückgegriffen. Man gewinnt damit psychologisch relevante Charakteristika von Merkmalsträgern und integriert die gegebenen Daten zu einem Urteil (Diagnose, Prognose). Als Merkmalsträger gelten Einzelpersonen, Gruppen, Institutionen, Situationen, Gegenstände etc. (Jäger & Petermann, 1999, S. 11).

Psychologische Diagnostik ist eine Methodenlehre innerhalb der Psychologie, bildet damit also einen Bereich z.B. mit der Statistik oder der Versuchsplanung. Anders als diese beiden Disziplinen stellt sie jedoch primär ein System von Verfahrensweisen im Dienste der Angewandten Psychologie dar. Neben dieser primären Funktion erfüllt Diagnostik aber auch Aufgaben für die Grundlagendisziplinen der Psychologie, insbesondere für die Differentielle Psychologie.

Beim psychologischen Diagnostizieren geht es (...) nicht, wie der psychologische Laie vielleicht meinen könnte, um das Erkennen des „Wesens“ eines Menschen, sondern um das Erfüllen eines praktischen (und damit eingegrenzten) Auftrags. Tatsächlich ist Diagnostizieren nicht primär ein Erkenntnisvorgang (im Alltagsverständnis dieses Begriffs), sondern (...) ein Handlungs- und Entscheidungsprozess (...). In einen Entscheidungsprozess mündende Aufträge können etwa darin bestehen, unter mehreren Bewerbern den für eine bestimmte Position geeignetsten herauszufinden, Eltern hinsichtlich des für ihr Kind passenden Schulzweigs zu beraten, gesundheitsrelevante Einstellungen einer Person zu erheben, um evtl. ein Programm zur Modifikation ungünstiger Einstellungen einzuleiten, oder zu bestimmen, ob bei einem Klienten eine behandlungsbedürftige Ausprägung von Depression vorliegt (...). (Krohne & Hock, 2015, S. 1)

Psychologisches Diagnostizieren ist ein Prozess, der unter Zuhilfenahme verschiedener Verfahren zielgerichtete Informationen über psychische Eigenschaften einer Person gewinnen will. Der Prozess bezieht sich auf

- Klärung der Fragestellung
- Auswahl diagnostischer Verfahren
- Anwendung und Auswertung der Verfahren
- Interpretation und Gutachtenerstellung
- Festsetzen der Intervention

Psychologische Diagnostik ist die wissenschaftliche Disziplin („Lehrfach“), die psychologisches Diagnostizieren für die Praxis vorbereitet (Kubinger, 2006, S. 3).

Bestimmungsstücke der Psychologischen Diagnostik

Aus diesen Definitionsvorschlägen lassen sich folgende Bestimmungsstücke der Psychologischen Diagnostik ableiten:

1. Die Psychologische Diagnostik ist eine Methodenlehre im Dienste der Angewandten Psychologie.
2. Ihr Gegenstand ist die gezielte und regelgeleitete Sammlung und Verarbeitung von Daten, die für die Bearbeitung von Fragestellungen relevant sind. Spezielle diagnostische Verfahren (z.B. Tests, Fragebögen, Interview, Verhaltensbeobachtung), die sowohl empirisch und praktisch bewährt, als auch theoretisch fundiert sein müssen, werden zur Datensammlung eingesetzt. Da nicht nur in der Angewandten Psychologie, sondern auch in der Grundlagenforschung Daten erhoben und verarbeitet werden, ist die Psychologische Diagnostik auch für die Grundlagendisziplinen, insbesondere die Differentielle Psychologie, bedeutsam.
3. Zu den grundlegenden Fragestellungen der Diagnostik zählen die Beschreibung, Klassifikation, Vorhersage und Evaluation von Unterschieden zwischen und innerhalb von Personen im Hinblick auf psychische Zustände (z.B. Angst), Eigenschaften (z.B. Intelligenz) und deren Veränderungen sowie damit einhergehender relevanter Bedingungen.
4. Psychologische Diagnostik untersucht nicht nur Einzelpersonen, sondern auch Gruppen, Organisationen, Situationen und andere Merkmalsträger.
5. Psychologische Diagnostik bereitet Entscheidungen (z.B. zur Berufseignung, zur Wirksamkeit von Interventionsprogrammen) nach wissenschaftlichen Kriterien und ethischen Standards vor.
6. Psychologisches Diagnostizieren ist ein Prozess, der mehrere Phasen umfasst: Klärung der Fragestellung, Auswahl von psychologisch-diagnostischen Verfahren, Anwendung, Auswertung, Interpretation, Gutachtenerstellung, Interventions- bzw. Maßnahmenvorschlag.

Enge Verbindung zwischen Diagnostik und Intervention

Der im sechsten Merkmal genannte Interventions- und Maßnahmenvorschlag als Teil des diagnostischen Prozesses verweist schon auf die enge Verbindung von *Diagnostik und Intervention*. Dass Diagnostik und Intervention zusammenhängen, lässt sich einerseits mit einem *erweiterten Interventionsbegriff* begründen. In einem engeren Sinn wird unter Intervention eine systematische, auf Veränderung abzielende therapeutische oder pädagogische Maßnahme verstanden. Nach Jäger (2003, S. 230) schließt Intervention in einem erweiterten Sinn dagegen jede Maßnahme ein, die für den Probanden eine Wirkung nach sich zieht, z.B. auch die Entscheidung für eine bestimmte Berufsausbildung oder einen Arbeitsplatz, die jemand aufgrund psychologisch-diagnostischer Beratung trifft. Zudem sind die *Übergänge zwischen Diagnostik und Intervention* im engeren Sinn mehr oder weniger fließend. Schon wenn ein Proband einen Fragebogen ausfüllt, können durch die Iteminhalte Reflexionsprozesse ausgelöst und in einer anschließenden Therapie- oder Coaching-Maßnahme genutzt werden. Noch deutlicher lässt sich der fließende Übergang zwischen Diagnostik und Intervention am Beispiel der

Selbstbeobachtung veranschaulichen, die z.B. in verhaltenstherapeutisch orientierten Interventionen zum Einsatz kommt (z.B. Reinecker, 1986). Ein Klient, der sich das Rauchen abgewöhnen möchte, kann z.B. den Auftrag erhalten, in einem Verhaltenstagebuch zu registrieren, unter welchen externen und internen Bedingungen er die Tendenz hat, zur nächsten Zigarette zu greifen. Diagnostisch gesehen geht es dabei um die Erfassung der Häufigkeit und der Anlässe des Rauchens. Selbstbeobachtung ist aber zumeist auch reaktiv, d.h. die Beobachtung und Registrierung des eigenen Verhaltens führt häufig auch zu einer Veränderung des Verhaltens. Anders ausgedrückt wirkt sich das Beobachten auf das Beobachtete aus. Wer bei jedem Impuls zu rauchen, zunächst die externen und internen Umstände in einem Verhaltenstagebuch registriert, der wird nach diesem Selbstbeobachtungsvorgang u.U. gar nicht mehr zu dieser protokollierten Zigarette greifen.

Psychologische Diagnostik stellt nicht nur Verfahren zur Erhebung und Verarbeitung von Daten für die Grundlagen- und Anwendungsfächer zur Verfügung, sondern profitiert umgekehrt auch von den Entwicklungen in den unterschiedlichen psychologischen Teildisziplinen. Nach Hossiep und Wottawa (1993, S. 132) konstituieren Grundlagen- und Anwendungsfächer sowie Methoden ein *diagnostisches Dreieck*, in dem die psychologische Diagnostik entsteht und fruchtbar wird. Von „außen“ wirken kulturelle, soziale, wirtschaftliche und technische Rahmenbedingungen auf das diagnostische Dreieck ein.

Diagnostisches Dreieck

Als Methodenlehre im Dienste der Angewandten Psychologie ist die Psychologische Diagnostik wie oben bereits erwähnt für alle Fragestellungen relevant. Eine exemplarische Übersicht der *diagnostischen Aufgabenfelder und Fragestellungen in einigen wichtigen Anwendungsfächern* soll dies verdeutlichen (vgl. Schmidt-Atzert & Amelang, 2012, S. 5ff.):

Diagnostische Aufgabenfelder und Fragestellungen

- ABO-Psychologie, z.B. Personalselektion und –entwicklung
- Pädagogische Psychologie, z.B. Eignung für weiterführende Schulen und Studiengänge, Erziehungsprobleme
- Klinische Psychologie, z.B. Diagnose psychischer Störungen, Auswahl und Evaluation von Interventionsmaßnahmen
- Forensische Psychologie, z.B. Glaubwürdigkeitsdiagnostik, Bewährungsprognose
- Markt- und Werbepsychologie, z.B. Diagnose der „Anmutungsqualität“ eines neuen Artikels
- Ökologische Psychologie, z.B. subjektive Wahrnehmung von Wohn-, Arbeits-, und Schulumwelten
- Verkehrspsychologie, z.B. Erfassung der Fähigkeit zum Führen von Fahrzeugen

2.2 Arten, Dimensionen und Zielsetzungen

Arten/Strategien der Diagnostik	In den eben skizzierten diagnostischen Aufgabenfeldern und Fragestellungen der unterschiedlichen angewandten Teildisziplinen kehren zwei „Themen“ immer wieder, die Schmidt-Atzert und Amelang (2012, S. 16ff.) als <i>grundlegende Arten bzw. Strategien der Diagnostik</i> bezeichnen: Die <i>Selektionsdiagnostik</i> und die <i>Modifikationsdiagnostik</i> .
Selektionsdiagnostik	Die <i>Selektionsdiagnostik</i> zielt im Rahmen personal-, organisations-, oder pädagogisch-psychologischer Aufgaben darauf ab, geeignete Personen oder Bedingungen auszuwählen (zu selektieren). Bei der <i>Personenselektion</i> sollen geeignete Kandidatinnen und Kandidaten ermittelt werden, die bestimmten Anforderungen genügen, die z.B. mit einem Arbeitsplatz, einer Schulart oder einem Studienfach verbunden sind. Bei der Personenselektion sind die Anforderungen fixiert und die Personen sozusagen variabel. Bei der <i>Bedingungsselektion</i> ist es dagegen umgekehrt: Für Personen mit einem bestimmten Fähigkeits- und Merkmalsprofil, das als stabil angenommen wird, sollen geeignete Bedingungen ausgewählt werden, die zu dem jeweils persönlichen Profil passen. So können beispielsweise Empfehlungen für bestimmte Berufe oder geeignete Arbeitsplätze ausgesprochen werden.
Modifikationsdiagnostik	<i>Modifikationsdiagnostik</i> steht dagegen im Mittelpunkt klinisch-psychologischer Fragestellungen. Bei dieser Form der Diagnostik soll ermittelt werden, welche Erlebens- und Verhaltensweisen einer Person verändert werden sollen oder welche externen Bedingungen verändert werden müssen, damit ein Problemverhalten reduziert werden kann. Im ersten Fall dient die Diagnostik der Verhaltensmodifikation, im zweiten Fall der Bedingungsmodifikation.
	Selektionsdiagnostik basiert hauptsächlich auf dem Eigenschaftsmodell, während sich Modifikationsdiagnostik in erster Linie auf verhaltensdiagnostische Prinzipien bezieht.
Eigenschaftsmodell	Das <i>Eigenschaftsmodell</i> geht davon aus, dass Personen auf bestimmten (eignungsrelevanten) Dimensionen (z.B. Intelligenz, soziale Kompetenz, Gewissenhaftigkeit) verglichen werden können. Individuelle Ausprägungen einer Eigenschaft werden zu den Ausprägungen einer Normstichprobe in Beziehung gesetzt und es kann abgeschätzt werden, ob die Person unter-, über-, oder durchschnittliche Werte auf der interessierenden Dimension aufweist. Wie man bei diesen Schätzungen vorgeht, haben Sie im Kurs 03421 zur Testkonstruktion erfahren bzw. erfahren Sie im Kurs 3419: Persönlichkeitskonstrukte und Persönlichkeitsmessung. Wegen der relativen zeitlichen Stabilität und transsituativen Konsistenz von Eigenschaften sind Prognosen möglich, z.B. lässt sich der Studienerfolg mit einer gewissen Wahrscheinlichkeit aus der Intelligenz und bestimmten Persönlichkeitsmerkmalen vorhersagen.
Verhaltensdiagnostik	Die wesentlichen Modellannahmen der <i>Verhaltensdiagnostik</i> besagen, dass Verhalten erlernt ist, von Situation zu Situation variiert und sich prinzipiell verändern

lässt. Eine zentrale diagnostische Frage ist es, zu identifizieren, welche situativen Bedingungen ein Verhalten auslösen und aufrechterhalten. Unter welchen situativen Bedingungen z.B. zeigt ein Kind aggressives Verhalten. Nicht, was eine Person „hat“ (Eigenschaften), sondern, was eine Person tut, ist in der Verhaltensdiagnostik von Interesse.

Eigenschafts- und Verhaltensdiagnostik bzw. die damit einhergehende Selektion oder Modifikation von Personen/Verhalten und Bedingungen lassen sich anhand von vier Dimensionen und damit verbundenen diagnostischen Zielsetzungen unterscheiden (Pawlik, 1976, S. 236).

Dimensionen und Zielsetzungen der Diagnostik

1. *Status- vs. Prozessdiagnostik*: Statusdiagnostik zielt auf die Erfassung eines Ist-Zustandes ab. In der Regel werden mehrere als stabil angenommene Eigenschaftsausprägungen gemessen, die eine probabilistische Vorhersage zukünftiger Erlebens- und Verhaltensweisen ermöglichen. Dagegen ist es das Ziel der Prozessdiagnostik, Veränderungen interessierender Verhaltensweisen (z.B. Häufigkeit aggressiver Handlungen) im Zeitverlauf zu erheben. Das Verhalten muss daher zu mehreren Zeitpunkten registriert werden.
2. *Normorientierung vs. Kriteriumsorientierung*: Wie oben bereits skizziert, wird in der Eigenschaftsdiagnostik ein individuelles Untersuchungsergebnis mit dem Durchschnittswert einer Normstichprobe verglichen. Eine Norm- oder Eichstichprobe ist eine möglichst repräsentative Vergleichsstichprobe von Personen, bei denen die interessierende Eigenschaft ebenfalls erhoben wurde. In der Verhaltensdiagnostik wird eine Person im Hinblick auf ein vorgegebenes Erlebens- oder Verhaltensziel (z.B. ein pädagogisches Ziel oder Therapieziel), das sogenannte Kriterium, untersucht. Kann beispielsweise ein Klient nach einem Selbstsicherheitstraining bestimmte Kriteriumssituationen (z.B. sich beschweren; die eigene Meinung vertreten) erfolgreich bewältigen? Ist die Fehlerzahl in einem Diktat nach einem Rechtschreibtraining bei einem Schüler unter eine bestimmte Grenze abgesunken?
3. *Testen vs. Inventarisieren*: Eigenschaftsausprägungen werden ermittelt („getestet“), indem aus der Grundgesamtheit des Erlebens- und Verhaltensrepertoires eine Stichprobe gezogen wird. Um eine latente Eigenschaft zu erfassen, wird einer Person in einem Fragebogen z.B. eine Reihe von Items mit Aussagen vorgelegt, die die zu diagnostizierende Eigenschaft indizieren. Der Proband wird gebeten, sich selbst auf Feststellungen wie „In sozialen Situationen werde ich leicht unsicher“ oder „Ich habe Schwierigkeiten, anderen in die Augen zu schauen“ einzuschätzen (z.B. auf einer Skala mit den Abstufungen trifft gar nicht zu – trifft etwas zu – trifft weitgehend zu – trifft vollständig zu). Die unterschiedlichen Feststellungen repräsentieren eine Stichprobe relevanter Erlebens- und Verhaltensweisen, die in unserem Beispiel die Eigenschaft soziale Ängstlichkeit indizieren. In der Verhaltensdiagnostik geht es im Gegensatz dazu um eine möglichst vollständige

Inventarisierung des gesamten, für eine Fragestellung relevanten Verhaltensrepertoires und der damit verbundenen auslösenden und aufrechterhaltenden Bedingungen, z.B. welche aggressiven Verhaltensweisen führt ein Kind wie häufig gegen wen unter welchen Bedingungen aus?

4. *Diagnostik als Messung vs. Diagnostik als Information für und über Behandlung*: Ziel der Eigenschaftsmessung ist es, die Ausprägung eines Persönlichkeitsmerkmals vor dem Hintergrund bestimmter Testgütekriterien (Objektivität, Reliabilität, Validität) möglichst genau und valide abzuschätzen. In der Verhaltensdiagnostik ist das Ziel dagegen, Informationen zu erfassen, die helfen sollen, eine bestimmte Interventionsmaßnahme auszuwählen und/oder deren Effektivität zu bestimmen.

Selektionsdiagnostik (Eigenschaftsmodell) kann vor dem Hintergrund dieser vier Dimensionen als normorientierte Statusdiagnostik zur Messung bzw. Schätzung des Ausprägungsgrades von Eigenschaften expliziert werden. Demgegenüber ist Modifikationsdiagnostik (Verhaltenstheorie) kriteriumsorientierte Prozessdiagnostik zur Entscheidungs- und Behandlungsoptimierung.

Übergänge zwischen
Selektions- und Modifi-
kationsdiagnostik

In den ersten Zeilen dieses Abschnitts wurde die Selektionsdiagnostik der Arbeits- und Organisationspsychologie, die Modifikationsdiagnostik dagegen der Klinischen Psychologie zugeordnet. Diese Zuordnung ist eher exemplarisch zu verstehen, da sie den tatsächlichen diagnostischen Aktivitäten in den genannten Disziplinen nicht gerecht wird. In den meisten angewandten Teildisziplinen geht die Selektionsdiagnostik oft in Modifikationsdiagnostik über, weil es nicht (immer) nur darum geht, einen Ist-Zustand zu erfassen, sondern auch Veränderungen einzuleiten. In der Arbeits- und Organisationspsychologie steht bei der Personalauswahl natürlich die Selektionsdiagnostik im Mittelpunkt. Bei Personalentwicklungsmaßnahmen, die z.B. auf eine Veränderung bestimmter Kompetenzen der Mitarbeiter abzielen, kommt aber auch Modifikationsdiagnostik zum Einsatz. Bevor in der Klinischen Psychologie eine Therapie überhaupt beginnt, steht die Frage nach der Selektion der erfolgversprechendsten therapeutischen Maßnahme angesichts einer vorliegenden Störung im Vordergrund (sogenannte differentielle Indikation).

3 Stichpunkte zur Geschichte der Psychologischen Diagnostik

Im Pflichtliteratur-Kapitel zur Geschichte der Psychologischen Diagnostik werden Sie zunächst einige sehr frühe Beispiele für „psychologisches Diagnostizieren“ kennen lernen. So wurden bereits vor 3000 Jahren in China Leistungsprüfungen zur Eignungsfeststellung für gehobene Positionen im öffentlichen Dienst durchgeführt, die noch bis 1905 zum Einsatz gekommen sind (siehe auch den Abschnitt 1.7 Meilensteine der psychologischen Diagnostik im Pflichtliteratur-Kapitel des vorherigen Abschnitts). Hintergrund dieser Leistungsprüfungen war, dass es in China keine erbliche Führungsschicht gab. Im schriftlichen Teil wurde Wissen zu den Themenbereichen Recht, militärische Angelegenheiten, Landwirtschaft, Finanzen, Geographie, Rechnen und Literatur geprüft. Im Handlungsteil standen Verhaltensstichproben im Reiten, Musizieren und Bogenschießen im Mittelpunkt. Die Lösungen eines Kandidaten im schriftlichen Teil wurden von zwei unabhängigen Beurteilern bewertet und damit ansatzweise schon das Prinzip der intersubjektiven Übereinstimmung (Objektivität) umgesetzt. Im 19. Jahrhundert wurde diese Prozedur von den Engländern zur Auswahl von Mitarbeitern für die East India Company kopiert.

Leistungsprüfungen im alten China

Ein weiterer Meilenstein in der Geschichte der Psychologischen Diagnostik war die Messung allgemeiner psychischer Gesetzmäßigkeiten zur Fundierung der Psychologie als Wissenschaft im 19. Jahrhundert. Hier ist insbesondere der Ihnen schon bekannte Gustav Theodor Fechner zu nennen, der mit seiner Fundamentalformel, nach der die Sinnesempfindung S eine logarithmische Funktion der Reizstärke R ist ($S = k \times \log R$), eine Behauptung Kants widerlegte. Kant zufolge könne die Psychologie niemals eine Wissenschaft werden, weil psychische Phänomene nach seiner Auffassung nicht quantifizierbar sind.

Messung allgemeiner psychischer Gesetzmäßigkeiten

Während bei Fechner noch die Messung *allgemeiner* Gesetzmäßigkeiten durch laborexperimentelle Forschung und Reaktionszeitmessungen im Mittelpunkt stand, begann spätestens mit Francis Galton (1822 – 1911) die systematische Erforschung und Erfassung interindividueller Unterschiede. Galton war mit Charles Darwin verwandt und entwickelte vor dem Hintergrund der Evolutionstheorie die Idee einer biologischen Fundierung kognitiver Fähigkeiten. Zur Erfassung der Intelligenz setzte Galton elementare Maße ein, z.B. Reaktionszeitmessungen, Tests zur Diskriminationsfähigkeit im visuellen, akustischen und kinästhetischen Bereich sowie Gedächtnistests. Berühmt geworden ist die nach ihm benannte Galton-Pfeife zur Prüfung der Hörschwelle. Ein weiterer Pionier der Intelligenzmessung war der Wundt-Schüler James McKeen Cattell, der die Bezeichnung Mental Test für eine von ihm entwickelte Aufgabenbatterie einführte.

Galton und die 1. Periode der Diagnostik

Während die sogenannte 1. Periode der Diagnostik in den Experimentallabors stattfand, war die zweite Phase durch praktische Problemstellungen aus der Pädagogik und Psychiatrie gekennzeichnet. Hier sind insbesondere Hermann Ebbinghaus

Die 2. Periode der Diagnostik

(1850 -1909) und Alfred Binet (1857 – 1911) zu nennen. Letzteren werden Sie auch noch einmal genauer im Kurs 03419 kennen lernen.

Anfänge der Intelligenz und Leistungsmessung

Ebbinghaus war bekanntlich Experimentalpsychologe und ist mit seinen Studien zum Gedächtnis in die Geschichte der Psychologie eingegangen. Er führte aber auch im Auftrag der Stadt Breslau eine Untersuchung durch, in der er feststellen sollte, ob der Vor- oder der Nachmittagsunterricht mit einer größeren Ermüdung der Schüler verbunden ist. Er führte zur Beantwortung dieser Frage Gruppen-Intelligenztests, u.a. mit Rechen- und Gedächtnisaufgaben durch. Er stellte u.a. eine deutliche Steigerung der Leistungen mit dem Alter und eine positive Beziehung zwischen Test- und Schulleistung fest.

Alfred Binet (1857 – 1911) erhielt 1904 vom französischen Unterrichtsministerium den Auftrag, lernschwache Schüler zu identifizieren, für die Sonderschulen eingerichtet werden sollten. Er entwickelte gemeinsam mit Théodore Simon einen Intelligenztest, für den er auch einige Aufgaben von Ebbinghaus adaptierte. Zudem führte er das Konzept des Intelligenzalters ein, das von William Stern zum sogenannten Intelligenzquotienten und von Wechsler zum Abweichungs-IQ weiterentwickelt wurde. Diese Konzepte werden in der Pflichtliteratur und auch im Kurs 03419 ausführlich erläutert.

Meilensteine in der Geschichte der Persönlichkeitsdiagnostik

Einige „Meilensteine“ in der Geschichte der Persönlichkeitsdiagnostik, die Sie kennen müssen, sind

- Der Personal Data Sheet von Woodworth (1918), der als erster Persönlichkeitstest in Fragebogenform gilt (116 Items) und zur Eignungsfeststellung von Soldaten für den 1. Weltkrieg eingesetzt wurde.
- Das MMPI, Minnesota Multiphasic Personality Inventory (Hathaway & McKinley, 1943), das ebenfalls genauer im Kurs 03419 vorgestellt wird. Der Rorschach-Test (Rorschach, 1921), der zu den umstrittenen projektiven Verfahren zählt. Der „Rorschach“ (bitte immer ohne „h“ nach dem „o“!).
- Theorien und Modelle über Persönlichkeitsmerkmale: die Giant Three, die Big Five, und die 16 PF.

4 Testtheoretische Grundlagen der Diagnostik

Die Grundlagen der klassischen Testtheorie und einige Grundzüge von probabilistischen Testtheorien haben Sie bereits im Kurs 03421 (Grundlagen der Testkonstruktion) kennen gelernt bzw. studieren Sie gerade, wenn Sie Modul 6 parallel zu Modul 7 belegt haben. In diesem Kurs wurden auch Gütekriterien zur Beurteilung von diagnostischen Verfahren ausführlich vorgestellt. Da einige von Ihnen Modul 7 möglicherweise vor Modul 6 studieren und wieder andere erst mit dem siebten Modul ihr Studium im B.Sc. an der FernUniversität beginnen, muss – ähnlich wie bei den Inhalten im zweiten Kapitel dieses Studienbriefs – für einen ungefähr vergleichbaren Wissensstand gesorgt werden. Deshalb werden Sie zu diesem Abschnitt ein Pflichtliteratur-Kapitel lesen und erarbeiten, in dem die wichtigsten Testgütekriterien vorgestellt werden. Für viele von Ihnen wird die Lektüre sicherlich eine Wiederholung sein, die aber bei diesem so zentralen und grundlegenden Thema auf keinen Fall schadet.

An dieser Stelle sollen nur stichpunktartig die wichtigsten Begriffe und Konzepte aufgelistet werden, die Sie insbesondere nach dem erfolgreichen Abschluss von Modul 6 kennen, diskutieren und anwenden können müssen:

Übersicht zu Testgütekriterien

- Objektivität: Durchführungs-, Auswertungs-, und Interpretationsobjektivität
- Reliabilität: Retest-, Paralleltest-, Testhalbierungs-Reliabilität, interne Konsistenz
- Validität: Inhalts-, Augenschein-, Konstrukt- und Kriteriumsvalidität
- Skalierung
- Normierung (Eichung)
- Testökonomie
- Nützlichkeit
- Zumutbarkeit
- Unverfälschbarkeit
- Fairness

5 Interpretation von Resultaten aus psychologischen Tests (Einzelfalldiagnostik)

Standardmessfehler,
Standardschätzfehler,
Konfidenzintervalle und
kritische Differenzen

Ebenfalls im Kurs 03421 haben Sie im Abschnitt 5.4 unter der Überschrift „Gütekriterien für die Einzelfalldiagnostik“ gelernt, dass individuelle Testwerte messfehlerbehaftet sind und deshalb nicht einfach im Sinne des wahren Wertes interpretiert werden dürfen. Vielmehr gilt es unter Berücksichtigung des Standardmessfehlers bzw. Standardschätzfehlers Konfidenzintervalle zu bestimmen, in denen der wahre Wert eines Probanden mit einer festgelegten Wahrscheinlichkeit vermutet werden kann. Zudem lassen sich kritische Differenzen bestimmen, die mindestens erreicht werden müssen, damit man zwei Testwerte als unterschiedlich interpretieren kann. Ähnlich wie bei den Testgütekriterien ist es auch im Hinblick auf die Einzelfalldiagnostik notwendig einen vergleichbaren Kenntnisstand bei allen Studierenden im Modul 7 zu sichern. Deshalb ist wieder ein Pflichtliteratur-Kapitel zu lesen und zu erarbeiten, in dem die genannten Konzepte ausführlich erklärt und mit vielen Beispielen erläutert werden.

Interpretation von Test-
profilen

In dem Pflichtliteratur-Kapitel wird zudem ein Formelapparat vorgestellt mit dem sich *Unterschiede zwischen Testprofilen* ermitteln und interpretieren lassen. Ein Testprofil resultiert, wenn individuelle Testwerte aus einem Test oder Inventar, in dem mehrere Konstrukte erfasst werden, gemeinsam betrachtet und z.B. in ein Profildiagramm übertragen werden. Testprofile lassen sich z.B. in Tests, die verschiedene Intelligenzfacetten erfassen oder in mehrdimensionalen Persönlichkeitsinventaren (z.B. zur Erfassung der Big Five) erstellen. Dabei kann dann z.B. interessieren, ob ein individuelles Testprofil über die Zeit stabil geblieben ist oder sich verändert hat (z.B. infolge von Entwicklungseinflüssen oder Interventionsmaßnahmen).

Bei der Beurteilung der Stabilität bzw. Veränderung von wiederholt erfassten Testprofilen müssen zwei Faktoren berücksichtigt werden: die Profilstalt und die Profilhöhe (vgl. Kristof, 1958, Huber & Pfleger, 1994, Lienert & Raatz, 1998). Die Profilstalt betrifft die relativen Positionen der Profilmkmale zueinander, also deren Rangreihe. Mit der Stabilität bzw. Veränderung einer Profilstalt ist folglich die Stabilität bzw. Veränderung der relativen Positionen der Profilmkmale über die Zeit gemeint. Dieser Aspekt wurde von Renner (2002) als ipsativ-differentielle Stabilität bzw. Veränderung bezeichnet. Wäre man nur daran interessiert, würde es reichen, eine sogenannte Q-Korrelation über die individuellen oder gruppenbezogenen Profilmkmale zu berechnen. Dabei wird allerdings die Profilhöhe nicht berücksichtigt. Die Profilhöhe bezieht sich auf die möglichen absoluten Differenzen zwischen den einzelnen wiederholt gemessenen Profilmmerkmalen und wurde von Renner (2002) als ipsativ-absolute Stabilität bzw. Veränderung interpretiert.

Der Formelapparat von
Kristof (1958)

Eine Möglichkeit, die sowohl die Profilstalt als auch die Profilhöhe bei der Stabilitätsbestimmung berücksichtigt, bietet ein von Kristof (1958) entwickelter Formelapparat. Dabei werden zwei individuelle oder gruppenbezogene Profile

zuerst auf „Deckungsgleichheit“ getestet (globaler Profilvergleich). Resultiert ein signifikanter globaler Profilunterschied, kann weiter geprüft werden, ob die beobachtete Diskrepanz auf Unterschieden in der Profilhöhe und/oder auf Unterschieden in der Profilstalt basiert. Um diese stufenweise Prüfprozedur durchführen zu können, müssen zwei Voraussetzungen erfüllt sein: (1) Die Ausprägungen der Profilmerekmale müssen vor der Verrechnung auf der Basis einer Eichstichprobe normiert werden. (2) Außerdem müssen die (differentiellen) Stabilitäten (Retest-Korrelationen) der Profilmerekmale bekannt sein. Da der Vergleich von Testprofilen im Pflichtliteratur-Kapitel nur für den globalen Profilvergleich demonstriert wird, soll im Folgenden der gesamte Formelapparat vorgestellt werden.

Die nachfolgende Darstellung der Formeln orientiert sich an Huber und Pfleger (1994), die mit Hilfe des Formelapparates die Stabilität von clusteranalytisch ermittelten kardiovaskulären Reaktionsmustern überprüften. Eine Anwendung des Formelapparats auf Selbstinterpretationsstile bei redeängstlichen Studentinnen findet sich bei Renner (2002).

Die Prüfgröße für den globalen Profilvergleich lässt sich wie folgt berechnen:

$$Ch^2 = (N / 2\sigma^2) \sum_{j=1}^m [D_j^2 / (1 - r_{jj})], \quad df = m \quad (1)$$

Die Indizes $j = 1, \dots, m$ bezeichnen die verwendeten Profilmerekmale. D_j ist die Differenz zwischen zwei korrespondierenden Profilmittelwerten für die Messgröße j ; r_{jj} ist der zugehörige Stabilitätskoeffizient. σ^2 symbolisiert die Populationsvarianz, die bei einer T-Normierung der Messwerte 100 beträgt. Unterschiede in der Profilhöhe zweier Gruppenprofile lassen sich mit folgender Formel bestimmen:

$$Ch^2 = (Nv^2 / 2\sigma^2) \sum_{j=1}^m [1 / (1 - r_{jj})], \quad df = 1 \quad (2)$$

Die Differenz $v = h_1 - h_2$ zwischen den Profilhöhen h_1 und h_2 lässt sich mit den nachfolgenden Formeln bestimmen: Die Profilhöhe h eines einzelnen Gruppenprofils berechnet sich mit $M_j = \sum y_{ij} / N$ für $j = 1, \dots, m$ aus:

$$h = \frac{\sum_{j=1}^m [M_j / (1 - r_{jj})]}{\sum_{j=1}^m [1 / (1 - r_{jj})]} \quad (3)$$

Unterschiede in der Profilstalt zweier Gruppenprofile werden mit folgender Prüfgröße getestet:

$$\text{Chi}^2 = (N / 2\sigma^2) \sum_{j=1}^m (D_j - v)^2 / (1 - r_{jj}), \quad df = m - 1 \quad (4)$$

Zur Überprüfung der Stabilität bzw. Veränderung von zwei individuellen Merkmalprofilen (des Selbstinterpretationsstils bei einer Person) muss in den Gleichungen (1), (2), (3) und (4) $N = 1$ gesetzt werden. Zusätzlich kann die Stabilität eines Profils durch einen Stabilitätskoeffizienten r_{pp} beschrieben werden. Dieser Koeffizient lässt sich aus dem kritischen Chi^2 -Wert K für eine gegebene Signifikanzzahl α und dem beobachteten Chi^2 -Wert, wie folgt berechnen:

$$r_{pp} = K / (K + \text{Chi}^2_{\text{beob.}}) \quad (5)$$

r_{pp} rangiert zwischen 0 und 1 und lässt sich sowohl mit den empirischen Chi^2 -Werten für den globalen Profilvergleich (1), den Vergleich der Profilhöhen (2) und den Vergleich der Profilgestalten (4) bestimmen. Sind die beiden Profile identisch, dann resultiert ein $\text{Chi}^2_{\text{beob.}} = 0$ und für $r_{pp} = 1$. Für $\text{Chi}^2_{\text{beob.}} = K$ resultiert $r_{pp} = .5$. Je größer $\text{Chi}^2_{\text{beob.}}$ im Vergleich zu K ausfällt umso mehr nähert sich r_{pp} dem Wert Null an. r_{pp} -Werte kleiner .5 indizieren stets signifikante Instabilitäten und damit eine Veränderung des Merkmalsprofils (vgl. Huber & Pfleger, 1994, S. 169).

6 Datenquellen und diagnostische Verfahren

6.1 Datenquellen der Psychologie

Im Abschnitt 6 des Kurses 03401 wurden folgende 10 Datenquellen und Erhebungstechniken nach Pawlik (2006) skizziert:

1. Biographische und Aktaurdaten
2. Verhaltensspuren
3. Verhaltensbeobachtung
4. Verhaltensbeurteilung
5. Ausdrucksverhalten
6. Interview
7. Projektive Tests/Verfahren
8. Objektive Tests
9. Fragebogen
10. Psychophysiologische Diagnostik

Für die neu hinzugekommenen Studierenden im Modul 7 seien die Kurzbeschreibungen dieser Datenquellen und Erhebungstechniken an dieser Stelle wiederholt:

1. *Biographische und Aktaurdaten*: umfassen sozio-demographische Daten, wie Alter, Geschlecht, Familienstand und Beruf einer Person, die reaktionsobjektiv erfasst werden können. Auch andere objektive Daten zur Lebensgeschichte, der Schul- und Berufsbildung oder auch die Krankengeschichte zählen zu dieser Datenquelle. Biographische und Aktaurdaten können mit speziellen biographischen Fragebogen, in einem sogenannten anamnestischen Interview oder durch Rückgriff auf bestehende Aktenquellen erhoben werden. Pawlik (2006) ordnet biographische und Aktaurdaten der Modalität „Verhalten“ zu. Verhalten ist hier in einem weiten Wortsinn zu verstehen, was z.B. für den Familienstand, die Schul- und Berufsbildung noch nachvollziehbar ist, da mit diesen Informationen bestimmte Verhaltensweisen einhergehen. Das Alter und das Geschlecht lassen sich zwar schwer als „Verhalten“ auffassen, jedoch können diese Informationen bestimmte alters- und geschlechtstypische Verhaltensweisen indizieren – allerdings nur sehr allgemein und grob.

Biographische und Aktaurdaten

2. *Verhaltensspuren*: sind direkt beobachtbare Nachwirkungen (Produkte) menschlichen Verhaltens. Solche Nachwirkungen sind zum einen am äußeren Erscheinungsbild einer Person direkt beobachtbar: An einer Person, die mit Turnschuhen bekleidet, verschwitzt und außer Atem aus dem Wald kommt, sind Spuren des Verhaltens „Laufen“ sichtbar. Zum anderen fallen intentionale Verhaltensprodukte, wie z.B. Zeichnungen oder andere im weitesten Sinne künstlerische Werke in diese Datenquelle. Verhaltensspuren entstehen oft aber auch mehr oder weniger zufällig, so z.B., wenn man sich den Zustand eines Kinderzimmers nach einer

Verhaltensspuren

Geburtstagsparty oder den Arbeitsplatz eines Studierenden ansieht. Psychologisch interessant ist, dass Verhaltensspuren mit Persönlichkeitsmerkmalen assoziiert sein können. So ist es plausibel anzunehmen, dass das Verhalten einer hoch gewissenhaften Person Spuren in ihrem Arbeitsbereich hinterlassen wird, z.B. einen aufgeräumten und ordentlichen Schreibtisch. Ein solcher Schreibtisch kann für andere Personen wiederum als Hinweis für die Gewissenhaftigkeit des Besitzers dienen. Vor dem Hintergrund solcher Überlegungen zeigten Gosling und Kollegen (2002) einer Gruppe von Fremdbeurteilern die Büroräume von unbekanntem und nicht anwesenden Zielpersonen. Nach dem Besuch der Büros schätzten die Fremdbeurteiler bestimmte Persönlichkeitsmerkmale der Besitzer (Zielpersonen) ein. Die Besitzer der Büroräume schätzten dieselben Persönlichkeitsmerkmale zudem selbst ein. Es resultierten substantielle Übereinstimmungen, d.h. Korrelationen zwischen der selbst- und fremdeingeschätzten Offenheit, Gewissenhaftigkeit und Extraversion. Die Selbst- und die Fremdeinschätzungen dieser Eigenschaften waren zudem mit bestimmten Merkmalen bzw. Verhaltensspuren in den Büroräumen korreliert, die ein anderes Team von Beobachtern registriert hatte, z.B. mit der Farbigkeit (Offenheit, Extraversion) und der Ordentlichkeit (Gewissenhaftigkeit).

Verhaltensbeobachtung

3. *Verhaltensbeobachtung*: ist ein unmittelbar einleuchtender Begriff, weil wir auch im Alltag aktiv beobachten können, was andere tun. Wissenschaftliche Verhaltensbeobachtung ist im Gegensatz zur Alltagsbeobachtung aber methodisch kontrolliert und systematisiert, muss sich an bestimmten Gütekriterien messen lassen und ist zumeist mit einer weiterführenden quantitativ-statistischen Analyse der protokollierten Verhaltensdaten verbunden. Das zu beobachtende Verhalten umfasst alle visuell und akustisch wahrnehmbaren Aktivitäten und Veränderungen des Zustands einer Person, z.B. Körperbewegungen, Laut- und Sprachäußerungen, aber auch physiologische Reaktionen, z.B. Erröten oder Schwitzen. Verhalten kann von einer Person selbst oder von einer anderen, „fremden“ Person beobachtet werden. Im ersten Fall spricht man von *Selbstbeobachtung*, im zweiten Fall von *Fremdbeobachtung*. Wissenschaftliche Verhaltensbeobachtungen sind in der Regel sogenannte *reduktive Deskriptionen* (vgl. Mees, 1977), die sich auf bestimmte, theoretisch relevante Beobachtungseinheiten (Kategorien, Verhaltensklassen) beschränken. Mehrere solcher Beobachtungseinheiten oder Kategorien bilden ein sogenanntes *Zeichensystem*, wenn nur bestimmte, aber nicht alle Verhaltensweisen interessieren und beobachtet werden. Ein *Kategoriensystem* zielt dagegen darauf ab, jede Verhaltensweise innerhalb einer gegebenen Verhaltensstichprobe durch Beobachtung zu erfassen. Die Handhabung und Anwendung eines Kategoriensystems muss innerhalb eines *Beobachtertrainings* eingeübt werden. Beobachtung im Sinne reduktiver Deskription bedeutet dann, durch das Zeichen- oder Kategoriensystem definierte Verhaltensweisen zu identifizieren und den unterschiedlichen Kategorien des Beobachtungssystems zuzuordnen. Dabei muss genau und eindeutig definiert sein, welche Verhaltensweisen zu welcher Beobachtungskategorie gehören. In einem vom Autor entwickelten Beobachtungssystem zur Registrierung von Verhaltensindikatoren für Redeangst ist z.B. die Kategorie „Verlegenheitsgesten“

folgendermaßen definiert: „Die Probandin fasst sich an die Nase, ins Haar, ans Ohr, die Brille oder kratzt sich.“

4. *Verhaltensbeurteilung*: betrifft mehr oder weniger subjektive und summarische Einschätzungen und Bewertungen der Häufigkeit, Intensität und Ausprägungsform des eigenen Verhaltens oder des Verhaltens anderer Personen. Wenn eine Person ihr eigenes Verhalten einschätzt, dann tut sie dies auf der Basis mentaler Repräsentationen des zumeist bereits vergangenen Verhaltens. Aber auch die Beurteilung des Verhaltens anderer Personen basiert auf mentalen Repräsentationen, da das Verhalten nicht im Hinblick auf einzelne Mikroaspekte registriert und in diesem Sinne gemessen, sondern summarisch eingeschätzt wird, nachdem das Verhalten einer fremden Person u. U. über längere Zeiträume hinweg wahrgenommen wurde. Verhaltensbeurteilungen geben im Gegensatz zu Verhaltensbeobachtungen nicht oder zumindest nicht so genau vor, auf welche Mikroaspekte des Verhaltens die Aufmerksamkeit fokussiert werden soll. Verhaltensbeurteilungen lassen sich von Verhaltensbeobachtungen abgrenzen, je mehr subjektive Wertungen und Interpretationen in die Verhaltensbeurteilung eingehen (vgl. Ellgring, 1996). Letzteres ist z.B. immer dann der Fall, wenn die Abstufungen einer Ratingskala nicht quantitativ definiert sind (wie häufig ist „häufig“ oder „manchmal“, wie lange ist „kurzdauernd“) und wenn wenig verhaltensnahe Beurteilungseinheiten zugrunde liegen: Anstelle der Kategorien „Blickkontakt“ und „Versprecher“ könnte das Verhalten bei einer öffentlichen Rede z.B. auch im Hinblick auf „Unsicherheit“ oder „Souveränität“ eingeschätzt werden, z.B. auf einer numerisch und verbal verankerten Skala: 0 = gar nicht – 1 = etwas – 2 = ziemlich – 3 = sehr unsicher (bzw. mit derselben Skala, aber in einem eigenen Item: 0 = gar nicht – 1 = etwas – 2 = ziemlich – 3 = sehr souverän). Verhaltensbeurteilungen basieren häufig auf solchen abstrakten und auch komplexeren Merkmalen. Die notwendigen subjektiven Interpretationen bei der Einschätzung komplexer Merkmale nutzen – positiv gewendet – die menschliche Fähigkeit, Verhaltensindikatoren zu „verschmelzen“ und integrierend zu interpretieren (vgl. Langer & Schulz v. Thun, 1974).

Verhaltensbeurteilung

5. *Ausdrucksverhalten*: wird von Pawlik (2006) als eigene Datenquelle gelistet, obwohl es auch der Verhaltensbeobachtung oder –beurteilung hätte zugeordnet werden können. Ausdrucksverhalten umfasst solche Variationen der Mimik (Gesichtsausdruck), der Stimme und Sprechweise sowie der Ganzkörpermotorik, in denen sich Gefühle, Stimmungen und Affekte äußern. Auch die Handschrift als Verhaltensspur zählt zum Ausdrucksverhalten. Rückschlüsse auf die Persönlichkeit können aus (bestimmten Merkmalen) der Handschrift jedoch *nicht* abgeleitet werden; wohl aber verändert sich die Handschrift in der Regel unter Stress und spiegelt damit eine bestimmte aktuelle Befindlichkeit wider. Ein international verbreitetes Beobachtungssystem zur Registrierung der mimischen Muskelbewegungen ist das *Facial Action Coding System* (FACS), das 1976 von Ekman und Friesen erstmals vorgelegt und seitdem kontinuierlich weiterentwickelt wurde. Im FACS sind mehrere sogenannte Action Units (AUs) definiert. Es handelt sich dabei um kleinste, gerade noch sichtbare mimische Veränderungen bzw. Bewegungseinheiten, z.B.

Ausdrucksverhalten

Heben der Augenbraue innen (AU1), Zusammenpressen der Lippen (AU23) oder Runzeln der Kinnhaut (AU17). Wichtig ist, dass diese Action Units zunächst ohne Bewertung objektiv erfasst werden. Erst in einem zweiten Schritt werden die Action Units dann mit Hilfe eines Auswertungsprogramms bestimmten Emotionen zugeordnet. Grundlage für diese Zuordnung ist Ekman's Annahme, nach der die Ausdrucksformen bestimmter Primäremotionen wie Angst, Wut und Ekel biologisch determiniert sind und damit universell gelten. Diese Annahme ist nicht unumstritten. Ekman und Friesen selbst konnten aber in einer Reihe von Studien zeigen, dass zumindest bestimmte Primäremotionen, z.B. Angst, Wut und Ekel kulturübergreifend erkannt und richtig interpretiert werden. Die Handhabung des FACS ist sehr aufwändig und erfordert ein langes Training.

Interview

6. *Interview*: Neben der standardisierten Beobachtung gilt das Interview als zweiter traditioneller Königsweg zumindest innerhalb der psychologischen Diagnostik. Das Interview ist nämlich in der psychologischen Berufspraxis das am häufigsten eingesetzte diagnostische Instrument (vgl. Kici & Westhoff, 2000). Ein Interview ist eine zielgerichtete mündliche Kommunikation zwischen einem oder mehreren Befragern und einem oder mehreren Befragten, wobei eine Informationssammlung über das Verhalten und Erleben der zu befragenden Person(en) im Vordergrund steht (Keßler, 1999). Interviews können im Hinblick auf den Wortlaut, die Anzahl und Abfolge der Fragen mehr oder weniger standardisiert sein. Dasselbe gilt für die Antwortmöglichkeiten des Probanden, das Verhalten des Interviewers und die Auswertung. Je nach dem Grad der Standardisierung werden verschiedene Interviewformen unterschieden. Bei einer freien Exploration sind nur einige wenige Fragen bzw. „Aufträge“ festgelegt (z.B. „Erzählen Sie mir ihr Leben!“), mögliche Nachfragen hängen vom Verlauf der Exploration ab. In einem halbstrukturierten Interview sind dagegen einzelne Bereiche und Fragen festgelegt, die auf jeden Fall abgehandelt werden müssen, zusätzliche Nachfragen sind natürlich erlaubt. Standardisierte Interviews werden z.B. in der Diagnostik psychischer Störungen eingesetzt. Hier sind die Fragen und deren Abfolge genau festgelegt und der Interviewer hat u.U. auch die Aufgabe die Antworten zu kodieren, um z.B. den Schweregrad eines Symptoms zu quantifizieren. Dass das Interview eine derart beliebte Methode in der Praxis und auch schon bei Studierenden ist, liegt vermutlich auch an der *vermeintlichen* Einfachheit dieser diagnostischen Methode. Fragen stellen kann schließlich jeder. Ein Interview zu konstruieren, durchzuführen und auszuwerten, so dass valide Daten resultieren, ist aber viel komplexer und schwieriger als einen Fragebogen vorzulegen und auszuwerten. In einer Interviewsituation laufen Prozesse auf mehreren Ebenen ab. Der Interviewte und der Interviewer nehmen sich gegenseitig wahr, stellen Hypothesen auf und nehmen implizit Beurteilungen vor. Dabei spielen die wahrgenommene(n) Macht, Kompetenz und Absichten des Interviewers und auch des Interviewten eine Rolle. Sozialpsychologisch gesehen können Prozesse der Ersteindrucksbildung und Stereotype wirken, etwa wenn eine noch relativ junge Interviewerin eine ältere Person interviewt. Motivationspsychologisch können Bedürfnisse nach Kontrolle und Komplexitätsreduktion, Selbstdarstellungstendenzen, aber auch Interesse relevant sein. Emotionale Gestimmtheiten wie

Angst, Sympathie, Argwohn u.a., aber auch habituelle Tendenzen, also Persönlichkeitsmerkmale wie z.B. soziale Ängstlichkeit, Selbstwertschätzung oder soziale Kompetenzen können Einfluss auf die Interviewsituation nehmen. Konkret wird der Interviewte mehr oder weniger über folgende Fragen nachdenken: Was geht hier vor? Was will man von mir? Wie soll ich mich am besten verhalten? Warum will der Interviewer bestimmte Dinge wissen?

7. *Projektive Tests* oder vorsichtiger formuliert projektive Verfahren geben mehrdeutiges Reizmaterial vor (z.B., abstrakte Klecksbilder, mehrdeutige Bilder von Personen, frei gestaltbares Spielmaterial u.a.), die von Probanden bearbeitet, d.h. interpretiert, ergänzt oder gestaltet werden sollen. Projektive Verfahren sind im Kontext der Psychoanalyse bzw. Tiefenpsychologie entstanden und sollen helfen, sogenannte Abwehrmechanismen (vgl. psychoanalytisches Paradigma, Kurs 03400) zu umgehen und Zugang zu wichtigen psychologischen Informationen (z.B. Konflikte, Impulse) ermöglichen, die den Probanden nicht bewusst sind. Der Wert projektiver Verfahren ist ziemlich umstritten und wird in den meisten Lehrbüchern als eher niedrig eingeschätzt. In der Tat sind die klassischen Gütekriterien, also Objektivität, Reliabilität und Validität bei vielen projektiven Verfahren nicht gegeben. Weil projektive Verfahren in der Praxis aber nach wie vor zum Einsatz kommen, muss sich jeder Psychologie-Studierende mit dieser Datenerhebungsmethode auseinandersetzen, um im konkreten Fall Argumente anführen zu können, warum ein bestimmtes projektives Verfahren nicht verwendet wird oder aber doch. Denn es wäre falsch, projektive Verfahren als vollkommen unbrauchbar zu bewerten, da sie in der Praxis gerade in der Arbeit mit Kindern eine nützliche Hilfe zur Exploration oder Kontakt- und Beziehungsanbahnung sein können. So enthält z.B. der Szeno-Test Spielmaterial (Puppen, Bausteine etc.), das gerade für Kinder einen hohen Aufforderungscharakter bietet und einen indirekten Weg darstellt, über das Spiel Hypothesen über mögliche Probleme zu generieren, die dann freilich mit anderen diagnostischen Methoden genauer untersucht, erhärtet und möglicherweise auch wieder verworfen werden müssen. Zudem spielen projektive Verfahren, insbesondere der Thematische Apperzeptionstest, in der Forschung zur impliziten Machtmotivation eine wichtige Rolle (Winter, 1973; Schultheiss & Pang, 2007). Implizite Motive sind im Gegensatz zu expliziten Motiven nicht bewusst repräsentiert (McClelland et al., 1989) und können deshalb auch nicht mit Hilfe von Verfahren erfasst werden, die einen solchen bewussten Zugang voraussetzen (z.B. Fragebogen).

Projektive Tests

8. *Objektive Tests*: Pawlik (2006) bezeichnet objektive Tests als an Testgütekriterien überprüfte Stichproben z.B. von Leistungsaufgaben, über die eine Verhaltensstichprobe der untersuchten Person im interessierenden Merkmal erhoben und ausgewertet wird. In diesem Sinne wäre ein Intelligenztest ein objektiver Test und manchmal werden auch diagnostische Verfahren, die Persönlichkeitsmerkmale mit Hilfe von Fragebogen erheben, als Persönlichkeitstests bezeichnet (vgl. z.B. Brähler et al., 2002). Es gibt aber noch eine andere Variante von objektiven Tests, die auf Raymond B. Cattell zurückgeht und solche Verfahren meint, deren

Objektive Tests

Messintention für den Probanden nicht per Augenschein erschließbar ist, die also “undurchschaubar” sind. Der Test soll damit “unverfälschbar” z.B. im Sinne sozialer Erwünschtheit sein. Wenn in einem Fragebogen ein Item vorgelegt wird wie z.B. „Ich empfinde häufig Furcht und Angst“, dann ist ziemlich offensichtlich, welches Merkmal damit erfasst werden soll und der Proband kann seine Antwort in diejenige Richtung verzerren, die er für günstig hält. Ein Beispiel für einen objektiven Test sensu Cattell ist der Subtest “Flächengrößen Vergleichen” aus der Testbatterie Arbeitshaltungen von Kubinger und Ebenhöf (1996). Es handelt sich dabei um einen Test, der computergestützt administriert wird und mit dem ein kognitiver Stil, nämlich “Impulsivität vs. Reflexivität” undurchschaubar erfasst werden soll. Impulsivität vs. Reflexivität ist die Tendenz, in Problemsituationen entweder schnell und fehlerreich bzw. langsam und fehlerarm zu arbeiten. 20 Items mit jeweils zwei Flächen müssen in 30 Sekunden verglichen werden. Reflexive bearbeiten in dieser kurzen Zeit zwar nur wenige Items, aber machen dabei wenige Fehler. Impulsive bearbeiten dagegen viele oder sogar alle Items und machen dabei vergleichsweise viele Fehler. In den letzten Jahren hat es in der Forschung eine Renaissance der Entwicklung und des Einsatzes von solchen objektiven Tests gegeben (vgl. Ortner et al., 2007).

Fragebogen

9. *Fragebogen:* Nach Mummendey (2003) ist ein Fragebogen ein Erhebungsinstrument, bei dem festgelegte Antwortmöglichkeiten auf klar vorgegebene Fragen oder Feststellungen angekreuzt werden. Zu Beginn eines Fragebogens wird eine kurze Instruktion gegeben, in der das Ausfüllen erläutert wird, und manchmal auch ein Beispiel-Item (vgl. die nachfolgende Textbox).

Im Folgenden finden Sie eine Reihe von Feststellungen, mit denen man **sich selbst beschreiben** kann. Bitte lesen Sie jede Feststellung durch und wählen Sie aus den Antwortmöglichkeiten diejenige aus, die am besten angibt, was **im Allgemeinen** auf Sie zutrifft. Kreuzen Sie bitte bei jeder Feststellung das entsprechende Kästchen der von Ihnen gewählten Antwort (durch einen Klick) an. Beispiel:

	trifft gar nicht zu	trifft etwas zu	trifft weit- ge- hend zu	trifft voll- stän- dig zu
Ich möchte von Leuten, die mir wichtig sind, anerkannt werden	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

In einem Fragebogen ist also *keine offene Beantwortung* vorgesehen wie in einem Interview. Zudem werden die Antworten auf einzelne Fragen in der Regel nicht für sich ausgewertet und interpretiert, sondern es werden stets die Antworten auf eine größere Zahl von Fragen oder Feststellungen nach bestimmten statistischen

Prinzipien in einem einzigen Messwert zusammengefasst. Fragebogen sind die am häufigsten eingesetzte und bevorzugte Methode zur Erfassung von Persönlichkeitsmerkmalen sowie Motivations- und Interessensvariablen. Der Fragebogen kennzeichnet in der Klassifikation von Pawlik eindeutig eine Erhebungsmethode und keine Datenquelle. Die einem Fragebogen zugrunde liegende Datenquelle ist in den meisten Fällen die Selbsteinschätzung der eigenen Persönlichkeitsmerkmale. Fragebögen können aber auch herangezogen werden, um die Persönlichkeitsmerkmale von anderen Personen einschätzen zu lassen. In diesem Fall sind Fremdeinschätzungen die Datenquelle.

10. *Psychophysiologische Diagnostik*. Veränderungen des Erlebens und Verhaltens kovariieren mit organismischen Variablen, z.B. mit dem Blutdruck, der Herzfrequenz, der Ausschüttung bestimmter Hormone und der Aktivität in bestimmten Hirnarealen. Wer zum ersten Mal vor einer größeren Gruppe von fremden und statushöheren Personen eine Rede halten soll, der wird die dabei in der Regel auftretende Aufregung auch körperlich spüren, wobei nicht alle organismischen Veränderungen bewusst werden. Psychophysiologische Daten betreffen solche im weiten Sinne körperlich-organismischen Veränderungen und werden mit Hilfe elektrophysiologischer Registriermethoden (elektrodermale Aktivität; Herz-Kreislauf-Aktivität; Muskelaktivität; Hirnaktivität) oder über biochemische Messgrößen (Hormonspiegel im Blut, immunbiologische Analyse des Speichels, Messung der regionalen zerebralen Durchblutung oder des regionalen zerebralen Stoffwechsels) gemessen. In den letzten Jahren haben solche psychophysiologischen Maße und Erfassungsmethoden innerhalb des kognitions- und neuropsychologischen Paradigmas zentrale Bedeutung gewonnen.

Psychophysiologische
Diagnostik

6.2 Weitere Klassifizierungen für Daten

Pawlik (2006) klassifiziert die 10 Datenquellen nach den folgenden *drei Kriterien*: (a) der Datenmodalität, (b) der erfassbaren Varianz und (c) der Reaktionsobjektivität. Nach dem Kriterium der Datenmodalität wird unterschieden, ob es sich bei den Daten lediglich um mentale Repräsentationen des Erlebens und Verhaltens einer Person, um direkt beobachtbares Verhalten oder um psychophysiologisch erfassbare Variablen handelt. Unter erfassbarer Varianz wird klassifiziert, ob die Datenquelle im Labor und/oder im Feld erhoben werden kann. Die Reaktionsobjektivität betrifft dagegen das Ausmaß, in dem eine Datenquelle durch Ziele, Werte und Einstellungen eines Probanden beeinflusst oder sogar verfälscht werden kann.

L-, Q- und T-Daten

Eine frühere Klassifikation von Datenquellen, die Sie in vielen Lehrbüchern finden werden, stammt von Raymond B. Cattell (1958), der zwischen L-, Q- und T-Daten unterscheidet:

L-Daten (Life Record Data)

- objektive Lebensdaten (z.B. Geschwisterzahl, Lebensereignisse)
- Verhaltensbeobachtung
- Fremdbeurteilung
- Projektive Verfahren
- Ausdrucksanalyse
- Morphologische Methoden

Q-Daten (Questionnaire Data): Selbstbeurteilung des Individuums

- Fragebogen
- Interview

T-Daten (Test-Data): Daten aus objektiven Tests

- Papier-Bleistift-Tests (z.B. Intelligenz)
- Apparative Anordnungen
- Physiologische Messungen
- Objektive Tests in der speziellen Bedeutung von Cattell!

Obwohl Cattell sehr stark zwischen Fragebögen und anderen Datenquellen unterscheidet und sie eindeutig den Q-Daten zuordnet, lassen sich bestimmte Anwendungen von Fragebögen auch anders einordnen (z.B. Entscheidungszeit pro Item bei einem Fragebogen als T-Daten; vgl. Kapitel 7.2.4 in Stemmler et al., 2016).

6.3 Erläuterungen zum Begriff des Tests

Zum Begriff „Test“

Der Begriff „*Test*“, der bei den Klassifikationen von Datenquellen und Verfahren immer wieder auftaucht, wird in der Psychologie mehrdeutig verwendet. Einigkeit besteht lediglich darin, dass es sich bei einem Test um ein Verfahren zur Gewinnung diagnostisch relevanter Daten handelt. In einem engen Sinn ist ein Test ein Verfahren, mit dem Daten unabhängig von den subjektiven Urteilen und Einschätzungen der Probanden erhoben werden können. Lediglich Fähigkeits- und Leistungstests sowie objektive Tests im Sinne von Cattell (vgl. Cattell & Warburton, 1967), deren Messintention für einen Probanden nicht per Augenschein erschließbar ist, die also „undurchschaubar“ sind, genügen diesem Kriterium. Dagegen lässt sich aus einer Definition von Lienert und Raatz (1998, S. 1) eine breitere Verwendung des Begriffs „psychologischer Test“ ableiten. Demnach handelt es sich bei einem psychologischen Test um

ein wissenschaftliches Routineverfahren zur Untersuchung eines oder mehrerer empirisch abgrenzbarer Persönlichkeitsmerkmale mit dem Ziel einer möglichst quantitativen Aussage über den relativen Grad der individuellen Merkmalsausprägung.

Tests dienen also der Erfassung von Persönlichkeitsmerkmalen, die in den meisten Fällen mit Hilfe von Fragebögen zur subjektiven Selbsteinschätzung erfolgt. Ein wissenschaftliches Routineverfahren ist ein Test deshalb, weil es bestimmten Testgütekriterien genügt und weil die Durchführung, Auswertung und zum Teil auch die Interpretation der erhobenen Daten standardisiert erfolgt und routinemäßig wiederholt werden kann. Ein Test liefert eine quantitative Aussage über ein interessierendes Merkmal, da die Antworten eines Probanden mit Zahlen verknüpft und zu einem Gesamtwert (Score) zusammengefasst werden. Der relative Grad der individuellen Merkmalsausprägung kann aus diesem Gesamtwert ermittelt werden, indem er zu einer Vergleichsgruppe (Norm- bzw. Eichstichprobe, siehe vorheriger Absatz) oder einem Kriterium in Beziehung gesetzt wird.

Die breitere Verwendung des Testbegriffs korrespondiert mit einer gängigen Klassifikation psychologischer und pädagogischer Tests, die Rolf Brickenkamp seinem gleichnamigen Handbuch zugrunde gelegt hat. Der „Brickenkamp“ liegt mittlerweile in der dritten, nunmehr zweibändigen Auflage vor und wird von Brähler et al. (2002) herausgegeben. In Tabelle 1 werden die im Brickenkamp verwendeten *drei Klassen psychologischer Tests* mit jeweils mehreren Subkategorien genannt.

3 Klassen psychologischer Tests nach Brickenkamp

Bei der im Kapitel 2 vorgenommenen Explikation der Psychologischen Diagnostik wurde betont, dass es dabei nicht nur um Einzelpersonen geht, sondern auch um Gruppen, Organisationen, Situationen und andere Merkmalsträger. Gruppen- und interaktionsdiagnostische Verfahren „verbergen“ sich z.T. zwar auch innerhalb der Subkategorien der drei Klassen psychologischer Tests nach Brickenkamp; sie werden in diesem Werk aber nicht als eigene Klasse ausgewiesen. Weil sich Gruppen- bzw. Interaktionsdiagnostik aber offensichtlich schon einmal deswegen von „normaler“ Diagnostik unterscheidet, weil dabei eben nicht nur eine Person im Mittelpunkt steht, ist diesem Thema in den meisten Lehrbüchern ein eigenes Kapitel gewidmet. Interaktionsdiagnostische Verfahren werden Sie im Kurs 03420, KE 2 sowie im Modul 10 kennen lernen.

Gruppen- und Interaktionsdiagnostik

Tabelle 1

3 Klassen psychologischer Tests nach Brickenkamp

Leistungstests	
1.1 Entwicklungstests	
1.2 Intelligenztests	
1.3 Allgemeine Leistungstests	
1.4 Schultests	
1.4.1 Einschulungstests	
1.4.2 Spezielle Schuleignungstests	
1.4.3 Mehrfächertests	
1.4.4 Lesetests	
1.4.5 Rechtschreibtests	
1.4.6 Mathematik- und Rechentests	
1.4.7 Sonstige Schultests	
1.5 Spezielle Funktionsprüfungs- und Eignungstests	
Psychometrische Persönlichkeitstests	
2.1 Persönlichkeits-Struktur-Tests	
2.2 Einstellungs- und Interessentests	
2.3 Klinische Tests	
2.3.1 Fragebogen	
2.3.2 Interviews	
2.3.3 Sonstige klinische Verfahren	
Persönlichkeitsentfaltungs-Verfahren (Projektive Tests)	
3.1 Formdeutungsverfahren 3.2 Verbal-thematische Verfahren 3.3 Zeichnerische und Gestaltungsverfahren	

Die Items von Leistungstests umfassen Aufgaben, die richtig oder falsch gelöst werden können. In einem psychometrischen Persönlichkeitstest werden selbstbeschreibende Aussagen vorgegeben, zu denen ein Proband auf einer mehrstufigen Antwortskala Stellung nehmen soll (im einfachsten, nämlich zweistufigen Fall, z.B. trifft zu – trifft nicht zu). Wichtig ist in diesem Zusammenhang die Unterscheidung zwischen „typischer Performanz (bzw. typisches Verhalten) und „maximaler Performanz bzw. Leistung“. Während es in Intelligenz- und Leistungstests darum geht, die maximale Leistung eines Probanden zu erfassen, zielen Persönlichkeitstests zu meist darauf ab, das typische, durchschnittliche Verhalten zu erheben (eben das, was per Instruktion „im Allgemeinen“ auf eine Person zutrifft).

Typische vs. maximale Performanz

Konzentrationstests werden oft als *speed-tests*, d.h. mit Zeitbegrenzung, vorgegeben. Die Schwierigkeit der Aufgaben ist gleichbleibend und eher niedrig. Bei einem *Power- oder Niveautest* erfolgt dagegen keine oder eine sehr großzügige Zeitvorgabe zur Lösung von Aufgaben mit steigender Schwierigkeit, z.B. in einem Intelligenztest. Ziel ist es, das intellektuelle Niveau, die „Denkkraft“ des Probanden zu erfassen. Die meisten Leistungstests enthalten speed- und power-Komponenten. Aufgaben werden beispielsweise mit steigender Schwierigkeit und einer Zeitbegrenzung vorgegeben.

speed-tests

power- oder Niveautests

Leistungs- und psychometrische Persönlichkeitstests können individuell, d.h. im Sinne einer *Einzeltestung*, oder aber einer *Gruppe* vorgegeben werden. Projektive Tests werden in der Regel nur als Einzeltests vorgegeben, da die Antwortmöglichkeiten der Probanden bei diesen Verfahren deutlich weniger standardisiert sind und häufig auch noch eine Verhaltensbeobachtung empfohlen wird.

Einzel- und Gruppentestung

Die Testsituation ist so zu gestalten, dass die interessierenden Merkmale optimal erfasst werden können. Wenn z.B. die maximale Leistung erfasst werden soll, ist es hinderlich, wenn Test- und Prüfungsangst durch die diagnostische Situation hervorgerufen werden. Deshalb empfiehlt es sich eine „Aufwärmphase“ einzuplanen, in der sich ein Proband mit der Situation vertraut machen kann und in der eine konstruktive Beziehung zwischen Testand und Testleiter aufgebaut wird. Weitere Hinweise zur Gestaltung der Testsituation sind dem Kapitel 4 bei Schmitd-Atzert und Amelang (2012) zu entnehmen.

Gestaltung der Testsituation

In der Praxis kommt eigentlich nur bei der Vorselektion großer Bewerberzahlen vor, dass nur ein einziges diagnostisches Verfahren eingesetzt wird. In der Einzelfalldiagnostik kommen i.d.R. mehrere diagnostische Verfahren zum Einsatz, was verschiedene Vorteile, aber auch einige Probleme mit sich bringt. Bei der optimalen Gestaltung einer Testsituation ist auch zu entscheiden, ob die diagnostische Untersuchung in einer einzigen Sitzung (einstufig) oder über mehrere Sitzungen (mehrstufig) erfolgen soll (vgl. Punkt 6.4 „Probleme der Veränderungsmessung“, S. 30ff. in diesem Studienbrief).

6.4 Probleme der Veränderungsmessung

Ein Spezialfall von mehrstufiger Datenerhebung liegt vor, wenn Veränderungen bzw. Unterschiede über die Zeit gemessen werden sollen, was z.B. immer dann der Fall ist, wenn Trainings- und Interventionseffekte evaluiert werden sollen. Dabei ist u.a. das Problem der Irreliabilität von Differenzwerten zu beachten. Auf dieses Problem gehen auch Schmidt-Atzert und Amelang (2012) im Kapitel 4.5 ein; es sei aber an dieser Stelle etwas genauer erläutert und bewertet. Der „statistische Stein des Anstoßes“ in der Veränderungsmessung betrifft die Bildung von Differenzwerten zwischen den Merkmalsausprägungen von Personen zu zwei Messzeitpunkten. Diese Differenzwerte sind mit Problemen behaftet, die manche Autoren zu der resignierten Schlussfolgerung führten, besser nicht zu versuchen, Veränderung zu messen, sondern ihre Fragestellungen anders zu formulieren (z.B. Cronbach & Furby, 1970: How should we measure change - or should we?).

Mangelnde Reliabilität der Differenzwerte

Das Hauptproblem ist wie bereits erwähnt die mangelnde Reliabilität der Differenzwerte. Bereits die einmalige Erfassung von psychischen Merkmalen ist generell mehr oder weniger messfehlerbehaftet und deshalb ungenau. Bei Differenzwerten verschärft sich die Messungenauigkeit, da in die Differenzen von zwei Messungen die Messfehlervarianzen beider Messzeitpunkte additiv eingehen. Dieser doppelte Messfehleranteil wirkt sich insbesondere dann reliabilitätsmindernd aus, wenn Pre- und Posttestmessung hoch miteinander korrelieren. Die Reliabilität von Differenzwerten lässt sich nach der Formel:

$$r_{dd} = (r_{jj} + r_{kk} - 2r_{jk}) / 2(1 - r_{jk}) \quad (6)$$

bestimmen (Guilford, 1954, S. 394). r_{dd} steht für die Reliabilität der Differenzwerte, r_{jj} und r_{kk} für die Reliabilitäten zu den beiden Messzeitpunkten und r_{jk} für die Interkorrelation der Messwerte. Bei einer Reliabilität von $r_{jj} = r_{kk} = 0.8$ zu beiden Messzeitpunkten und einer Korrelation von $r_{jk} = 0.7$ zwischen beiden Messwertereihen beträgt die Reliabilität der Differenzwerte lediglich $r_{dd} = 0.33$. Die Reliabilität der Differenzwerte kann nur dann angemessene Werte erreichen, wenn die beiden Messwertereihen möglichst niedrig oder am besten gar nicht korrelieren. Bei hohen Korrelationen wird die Reliabilität der Differenzen selbst dann relativ niedrig ausfallen, wenn die Reliabilitäten der Einzelmessungen relativ hoch sind. Unter diesen Umständen scheint es sehr problematisch, Veränderung mit Differenzwerten messen zu wollen.

In Defense of Difference Scores

In einem mit der Überschrift „In Defense of Difference Scores“ (S. 730ff) versehenen Abschnitt können Rogosa et al. (1982) allerdings zeigen, dass wahre Veränderungen in einer Stichprobe auch dann sehr präzise gemessen werden können, wenn die Reliabilität der Differenzwerte nahezu 0 entspricht. Die Reliabilität der Differenzen kann als ein Maß verstanden werden, das angibt, wie gut Personen auf der Basis ihrer Differenzwerte in eine Rangreihe gebracht werden können. Dies gelingt dann umso besser, wenn sich die wahren Differenzwerte deutlich unterscheiden,

also bei hoher Streuung der Differenzwerte. Ist die Streuung der Differenzwerte niedrig, wird die Reliabilität der Differenzen ebenfalls niedrig ausfallen; das bedeutet aber nicht, dass die nahezu identischen wahren Veränderungen in der Stichprobe unpräzise gemessen werden. Andere mit der Messung von Veränderungen verbundene Probleme wie z.B. die oftmals negativen Korrelationen zwischen Anfangswerten und Differenzwerten oder die Regression zur Mitte werden von Rogosa et al. ebenfalls diskutiert und in ihrem Problemgehalt relativiert. Insgesamt kommen Rogosa et al. zu Schlussfolgerungen, die den bisherigen kritischen und z.T. pessimistischen Sichtweisen zur Veränderungsmessung widersprechen: Differenzwerte zwischen verschiedenen Messzeitpunkten liefern sinnvolle, unverzerrte Schätzungen für wahre Veränderungen, insbesondere wenn Daten nicht nur für zwei sondern für drei oder mehr Messzeitpunkte vorliegen (vgl. auch Willet, 1989). Andere Veränderungsmaße, wie z.B. Regressionsresiduen sind unter praktischen Gesichtspunkten zur Veränderungsmessung weniger geeignet. Somit können „klassische“ Verfahren, die auf individuellen Differenzmaßen basieren, wie z.B. t-Tests für abhängige Stichproben und Varianzanalysen mit Messwiederholungen, zur statistischen Überprüfung von Veränderungshypothesen sinnvoll eingesetzt werden.

Das „Problem“ der sozialen Erwünschtheit: Das vielleicht „vermeintliche“ Problem der sozialen Erwünschtheit wurde ebenfalls im Kurs 03421 diskutiert. Immer dann, wenn Probanden in einer Studie oder Bewerber auf eine Stelle explizit aufgefordert werden, sich selbst zu beschreiben – sei es in einem Fragebogen oder in einem Interview – wird eine möglichst zutreffende, d.h. authentische Vermittlung von Persönlichkeitsmerkmalen erwartet bzw. verlangt. Anders als bei Leistungstests, in denen Probanden die Aufgaben lediglich absichtlich nicht oder falsch lösen können (faking bad), besteht bei Persönlichkeitsfragebögen die Möglichkeit, dass sich ein Proband so darstellt, wie er glaubt, dass es für ihn am günstigsten ist (faking good). In einer Bewerbungssituation mag ein Proband z.B. Items, die Gewissenhaftigkeit und Leistungsmotivation widerspiegeln, eher in Schlüsselrichtung beantworten (also eher in Richtung „trifft vollkommen auf mich zu“). Die gesamte Persönlichkeitsdiagnostik basiert auf einer Authentizitäts- oder Echtheitsunterstellung und versucht, mögliche Täuschungen, Über- oder Untertreibungen seitens der Probanden mit einem eigens entwickelten Methodenarsenal abzuschätzen und einzugrenzen - mit mäßigem Erfolg. Immer dann, wenn ein Persönlichkeitspsychologe seinem routinemäßigem Tagewerk nachgeht und versucht, Persönlichkeitsmerkmale zu erfassen, muss er sich nolens volens auch mit Selbstdarstellung befassen. Da die Vermittlung von Bildern der eigenen Person sowohl in authentischer, als auch in täuschender Absicht erfolgen kann, stellt sich bei der Erfassung von Persönlichkeitsmerkmalen im Fragebogen eigentlich immer die Frage, welche Form der Selbstdarstellung dabei zugrunde lag. Wenn vermutet werden kann, dass Probanden im Allgemeinen dazu neigen, ihre Antworten, z.B. in Bewerbungssituationen, in Richtung sozialer Erwünschtheit zu verzerren, dann schließt sich eine weitere Frage an: „Ist das wirklich ein Problem?“. Bernd Marcus (2003) hat gezeigt, dass sozial erwünschte Selbstbeschreibungen in Persönlichkeitsfragebögen, die im Rahmen der Eignungsdiagnostik eingesetzt werden, nicht wirklich so

Soziale Erwünschtheit

problematisch sind, wie gemeinhin angenommen wird. Vielmehr zeugt es von rezeptiver Selbstdarstellungskompetenz, wenn ein Bewerber die Anforderungen der Situation und die Erwartungen der relevanten Entscheidungsträger zutreffend erkennt und sich dementsprechend darstellt. In Berufen, die viel Kundenkontakt und Emotionsbewältigung erfordern (z.B. Verkäufer, Flugbegleiter), ist Selbstdarstellungskompetenz eine zentrale Anforderung. Wenn jemand in einem Fragebogen in der Lage ist, die relevanten Anforderungen zu erkennen, dann indiziert das möglicherweise, dass er oder sie das auch außerhalb des Fragebogens kann. Diese Hypothese könnte den empirischen Befund erklären, nach dem Selbstdarstellung – obwohl das gemeinhin angenommen wird – die Validität verfälschbarer Persönlichkeitsfragebogen in der Personalauswahl nicht beeinträchtigt (Marcus, 2003).

7 Grundlagen des diagnostischen Prozesses und der diagnostischen Urteilsbildung

Auch der diagnostische Prozess und die diagnostische Urteilsbildung wurden im Kurs 3400 in Grundzügen bereits dargestellt und sollen im vorliegenden Kurs durch zwei Pflichtliteratur-Kapitel differenziert und vertieft werden. Im Kurs 3400 wurden mit Jäger (2006, S. 89) folgende Merkmale des diagnostischen Prozesses herausgestellt:

Merkmale des diagnostischen Prozesses

- die zeitliche, organisatorische, strategische und personale Erstreckung zwischen vorgegebenen zunächst allgemeinen und später präzisierten diagnostischen Fragestellungen sowie deren Beantwortung
- Die Beantwortung erfolgt in Form einer Diagnose oder Prognose
- Im Verlauf des diagnostischen Prozesses können aus der Beantwortung einer Teilfragestellung neue Fragen resultieren, so dass eine Verschachtelung von Fragen und Beantwortungen resultiert
- Die Beantwortung der Fragestellungen ist zugleich Ausgangspunkt für eine Entscheidungshilfe bzw. Entscheidung für den Auftraggeber

Zudem wurde der diagnostische Prozess mit der Durchführung einer empirischen Untersuchung verglichen. Wie eine empirische Untersuchung beginnt auch ein diagnostischer Prozess mit einer Fragestellung, die in die psychologische Terminologie übersetzt und damit präzisiert werden muss. Im Anschluss werden vor dem Hintergrund der Fragestellung und evtl. vorliegender Vorinformationen Hypothesen aufgestellt und mit geeigneten diagnostischen Verfahren geprüft. Grundsätzlich sollten Daten nach den Prinzipien der Sparsamkeit und Eignung erhoben werden. Es sollten also nur so viele Daten wie nötig und auch nur solche Daten, die für die Fragestellung relevant sind, erhoben werden (vgl. Schmitt & Gschwendner, 2006). Im Unterschied zu den meisten empirischen Studien zu Forschungszwecken, die an größeren Stichproben durchgeführt werden, betrifft der diagnostische Prozess häufig aber nur eine zu untersuchende Person. Natürlich können auch Gruppen Gegenstand eines diagnostischen Prozesses sein; in vielen Fällen steht aber eben nur eine einzelne Person im Mittelpunkt. Folglich geht es in einem diagnostischen Prozess um die Überprüfung idiografischer (d.h. einzelfallbezogener) Hypothesen. Ein weiterer Unterschied zu einer forschungsbezogenen Studie besteht in der Möglichkeit und Notwendigkeit, weiteren Hypothesen nachzugehen, die sich im Verlauf des diagnostischen Prozesses ergeben. In einer empirischen Studie werden in der Regel nur diejenigen Hypothesen geprüft, die a priori aufgestellt wurden. Neue, im Verlauf der Forschung resultierende Hypothesen müssen dann auch in einer neuen Studie untersucht werden. Ein diagnostischer Prozess kann demgegenüber mehrere „Teiluntersuchungen“ umfassen. Bei der Beantwortung der Ausgangsfragestellung

Diagnostischer Prozess als empirische Untersuchung

muss das diagnostische Urteil wieder in die Sprache des Auftraggebers rückübersetzt und in mündlicher und/oder schriftlicher Form mitgeteilt werden.

Handlungstheoretisches Modell des diagnostischen Prozesses

In Kapitel 7 bei Krohne und Hock (2015) wird nun der diagnostische Prozess im Rahmen eines handlungstheoretischen Modells von Kaminski (1970) als ein sequentieller Arbeitsprozess mit Rückmeldungsschleifen expliziert, der erst mit der Erreichung eines Zielkriteriums abgeschlossen ist. Der Arbeitsprozess umfasst die folgenden Phasen (vgl. Krohne & Hock, 2015, S. 193ff):

- „Eingangsdaten und Datenbeschaffung“
- „Hypothesenbildung und –bearbeitung“
- „Hypothesenbeurteilung“
- „Planung der Datenbeschaffung“
- „Planung der praktischen Phase“
- „Praktische Phase“
- „Prüfungen der praktischen Phase“

In dem genannten Kapitel von Krohne und Hock (2015) werden Implikationen des Modells sowie Details der entsprechenden Phasen ausführlich vorgestellt. Dabei wird deutlich, welche Anforderungen ein diagnostischer Praktiker bei seiner Arbeit zu berücksichtigen und wie einzelne Arbeitsschritte zu organisieren sind, damit ein zuvor festgelegter Zielzustand möglichst gut erreicht werden kann.

Diagnostische Urteilsbildung

Ein besonders zentraler und kritischer Aspekt des diagnostischen Prozesses ist die diagnostische Urteilsbildung. Es handelt sich dabei um einen

...Teil eines mehrstufigen und rückgekoppelten Vorgangs, der die erste Analyse sozusagen ‚von selbst eingehender‘ Daten, die Hypothesenbildung, die Herstellung von Untersuchungssituationen, die Auswahl geeigneter diagnostischer Verfahren, die Datensammlung und –erhebung, die Datenbewertung und –kombination im Hinblick auf die Hypothesen, die diagnostische Entscheidung, sowie die Überprüfung der Folgen von Entscheidungen beinhaltet. (...) (Krohne & Hock, 2015, S. 133)

Validität diagnostischer Urteile: Klinische vs. statistische Urteilsbildung

In Kapitel 5 bei Krohne und Hock (2015) wird zunächst die Frage der Validität diagnostischer Urteile erörtert. In diesem Zusammenhang werden zwei widerstreitende Modelle vorgestellt und anhand empirischer Studien bewertet: die klinische und die statistische Urteilsbildung bzw. Vorhersage (Meehl, 1954). *Klinische Urteilsbildung* stützt sich auf intuitive, auf klinischem Erfahrungswissen basierende Informationsverarbeitung. Erhobene Daten werden auf der Basis klinischer Erfahrungen intuitiv zusammengefasst. Bei der *statistischen Urteilsbildung* erfolgt dagegen eine Kombination und Gewichtung der Vorhersage-Variablen anhand von statistischen Beziehungen, die vorher empirisch ermittelt wurden.

Im Abschnitt 5.2 gehen Krohne und Hock (2015) der Frage nach, wie diagnostische Urteile eigentlich zustande kommen. Zur Klärung dieser Frage wird untersucht, wie Diagnostiker in der Praxis vorgehen, um zu einem Urteil zu gelangen. Diagnostiker sind in diesen Studien also die Probanden! In sogenannten paramorphen Modellen wird dann formal beschrieben, wie Daten von Diagnostikern gewichtet und kombiniert werden. Vorgestellt werden lineare Modelle und Konfigurationsmodelle. Bei den paramorphen Modellen geht es *nicht* darum zu ermitteln, ob eine Diagnose „richtig“ ist: „Vielmehr soll ein Algorithmus (im einfachsten Fall eine Gleichung) erstellt werden, der es ermöglicht, die Diagnose bei gegebenen Daten zu reproduzieren. Ein gutes paramorphes Modell liefert bei gleichen Daten, die gleichen Vorhersagen wie der modellierte Diagnostiker“ (Krohne & Hock, 2015, S. 149). Im Abschnitt 5.3 werden dann die Vorteile von expliziten Modellen (statistische Modelle und paramorphe Modelle) zusammengefasst, die u.a. Transparenz, Optimierbarkeit, Lernen und Erfahrungsaustausch sowie Entlastung und letztlich auch Validität gewährleisten.

Paramorphe Modelle

Schließlich gehen Krohne und Hock (2015) noch auf die Frage ein, wie sich die klinische Urteilsbildung zur Optimierung von diagnostischen Urteilen nutzen lässt. Vor dem Hintergrund empirischer Studien und einer Meta-Analyse lässt sich eigentlich rational rechtfertigen, dass formelle (statistische) Methoden der diagnostischen Urteilsbildung – wann immer möglich – der informellen klinischen Urteilsbildung vorgezogen werden sollten. „Wann immer möglich“ bedeutet, wann immer empirisch und statistisch gesicherte Vorhersagemodelle verfügbar und anwendbar sind. Dagegen sind „klinische“ Erfahrungen für die Identifikation relevanter Prädiktorvariablen notwendig, deren tatsächliche Validität – in weiteren Schritten – statistisch abgesichert werden muss. In vielen Fällen wird eine Kombination klinischer und statistischer Urteilsbildung notwendig sein, da oft nur teilweise statistisch gesicherte Vorhersagemodelle vorliegen. Zudem können klinische bzw. „intuitive“ Daten einen wertvollen Beitrag zur diagnostischen Urteilsbildung leisten, wenn sie in geeigneter Weise aufbereitet und verwertet werden.

Optimierung von diagnostischen Urteilen

Als Beispiel dafür stellen Krohne und Hock die Shedler-Westen-Assessment-Procedure (SWAP-200, Shedler & Westen, 1998) vor. In einer Version des Verfahrens wird zuerst ein klinisches, tiefenpsychologisches Interview zu Persönlichkeitsstörungen durchgeführt. Im Anschluss werden die klinischen und intuitiven Eindrücke im Rahmen eines sogenannten Q-Sort-Verfahrens 200 persönlichkeitsbeschreibenden Feststellungen zugeordnet und eingeschätzt. Eine individuelle Zuordnung und Einschätzung kann dann mit fünf prototypischen Mustern verglichen werden, die distinkte Persönlichkeitsstörungen repräsentieren (antisozial-psychopathisch, emotional dysreguliert, vermeidend, narzisstisch, gehemmt-selbstkritisch). Die prototypischen Muster wurden im Rahmen von professionellen Beurteilungen von mehreren hundert klinisch auffälligen Adoleszenten ermittelt.

Shedler-Westen-Assessment-Procedure

8 Klinisch-psychologische Diagnostik und Intervention

Im B.Sc. Psychologie an der FernUniversität in Hagen wird die Klinische Psychologie als eigenes Anwendungsmodul *nicht* angeboten. Sie sollen in Ihrem Studium aber zumindest Grundkenntnisse der klinischen Diagnostik und Intervention erwerben, die z.T. auch in nicht-klinischen Berufsfeldern relevant sind. Bereits im Kurs 03400 haben Sie einige grundlegende Konzepte und Inhalte der Klinischen Psychologie kennen gelernt. Im vorliegenden Kurs sollen diese Inhalte durch die Lektüre und Erarbeitung eines weiteren Pflichtliteratur-Kapitels erweitert und vertieft werden. In dem unten angegebenen Beitrag von Thomas Fydrich geht es in erster Linie um die klinisch-psychologische Diagnostik, es werden jedoch auch die wichtigsten klinisch-psychologischen Interventionsmaßnahmen kurz skizziert.

Aufgaben der klinisch-psychologischen Diagnostik

Als wichtigste Aufgaben der klinisch-psychologischen Diagnostik werden genannt (Fydrich, 2012, S. 504f):

1. Qualitative und quantitative Beschreibung der vorliegenden (psychischen) Störung
2. Die Klassifikation der psychischen Störung
3. Exploration von besonderen lebensgeschichtlichen Bedingungen bei der Entstehung und dem bisherigen Verlauf der Störung
4. Beobachtung des Verlaufs der Intervention und der Veränderung der Symptomatik (adaptive Diagnostik, Verlaufsdiagnostik)
5. Überprüfung des Therapieerfolgs (Qualitätssicherung)

Sodann werden Rahmenbedingungen der klinisch-psychologischen Diagnostik diskutiert und das diagnostische Gespräch (Interview) als Erhebungsmethode vorgestellt.

ICD und DSM

Die Klassifikation psychischer Störungen wird vor dem Hintergrund der beiden international verbreiteten Systeme, dem ICD und dem DSM erörtert. Das *ICD (International Classification of Diseases)* ist ein von der WHO herausgegebenes Klassifikationssystem für alle Krankheiten. Kapitel V (F) betrifft die klinisch-diagnostischen Leitlinien für psychische und Verhaltensstörungen. Die Diagnostik erfolgt kategorial-polythetisch, d.h. jede Störung wird anhand mehrerer Symptome bzw. Kriterien beschrieben; die Identifikation einer bestimmten Anzahl von Symptomen führt zur Diagnose einer bestimmten Störung, die nach einer alphanumerischen Klassifikation kodiert wird. Das *DSM-5 (Diagnostic and Statistical Manual of Mental Disorders)* ist das aktuelle Klassifikationssystem der American Psychiatric Association. In den grundlegenden Zielen und Diagnosenamen stimmen die

WHO- und die APA-Klassifikation zwar überein; bei der Definition einzelner Diagnosen gibt es jedoch teilweise Unterschiede.

Sowohl im ICD als auch im DSM-IV TR (aber nicht mehr im DSM-5!) wird eine *multiaxiale Diagnostik* vorgenommen, die darauf abzielt, nicht nur die klinische Symptomatik einer betroffenen Person zu erfassen, sondern auch psychosoziale und medizinische Merkmale, die zum besseren Verständnis einer psychischen Störung beitragen können. Seit dem ICD-10 ist auch für die Diagnose der psychischen und Verhaltensstörungen (Kapitel V, F) eine multiaxiale Diagnostik vorgesehen, allerdings lediglich auf 3 Achsen, nämlich: Achse I: psychiatrische und somatische Diagnosen; Achse II: Beeinträchtigungen der psychosozialen Funktionsfähigkeit (WHO-Disability Diagnostic Scale) und Achse III: Belastungsfaktoren (Z Codes im ICD-10).

Multiaxiale Diagnostik

Einen breiten Raum in dem Beitrag von Fydrich (2012) nimmt die überblicksartige Darstellung verschiedener klinisch-diagnostischer Verfahren ein. Im Kurs 3400 haben Sie gelernt, dass zur Untersuchung psychologischer Fragestellungen verschiedene Paradigmen herangezogen werden können. Auch in der Klinischen Psychologie wird Diagnostik unter der „Schirmherrschaft“ verschiedener Paradigmen bzw. - bescheidener formuliert – verschiedener Ansätze durchgeführt. So lassen sich diagnostische Verfahren und Ansätze unterscheiden, die vor dem Hintergrund verhaltenstheoretischer und kognitiver, klientenzentrierter, psychodynamischer, systemischer und interpersoneller Modelle entwickelt wurden. Beispiele dafür werden bei Fydrich vorgestellt und diskutiert. Zudem wird auf Beobachtungsmethoden, die Rolle von Persönlichkeitstests sowie klinisch-diagnostische Methoden der biologischen Psychologie, Neuropsychologie und Psychophysiologie eingegangen. Am Ende werden Verfahren zur Erfolgskontrolle und Qualitätssicherung vorgestellt und das Kriterium der klinisch bedeutsamen Verbesserung erörtert.

Klinisch-diagnostische Verfahren

9 Neuere Entwicklungen

Nach Auffassung des Autors dieses Studienbriefs gab es in den letzten 10-20 Jahren drei wichtige Neuentwicklungen in der Psychologischen Diagnostik:

1. Internetgestützte Diagnostik
2. Ambulatorisches Assessment
3. Die Renaissance der objektiven Persönlichkeitsdiagnostik

9.1 Internetgestützte Diagnostik

Internetgestützte Diagnostik

Die meisten von Ihnen haben in unserem virtuellen Labor bereits an Fragebogenstudien teilgenommen, um Versuchspersonen-Stunden zu erwerben. Ohne das Internet hätten diese Fragebögen im Papier-Bleistift-Format zunächst gedruckt, dann von Ihnen ausgefüllt, und anschließend von unseren Hilfskräften in SPSS eingegeben werden müssen. Ohne das Internet wären unsere Umfragen deutlich teurer, unflexibler und aufwendiger. Wir hätten Ihnen z.B. die Fragebögen zuschicken müssen, Sie hätten uns die Fragebögen wieder zurückschicken oder wir hätten die Administration im Rahmen von Präsenzveranstaltungen durchführen müssen. Die Vorteile der internetgestützten Datenerhebung wurden bereits im Kurs 3401 dargestellt und seien an dieser Stelle noch einmal stichpunktartig aufgezählt: Asynchronizität und Alokaltät, Flexibilität, Automatisierbarkeit, Objektivität, Dokumentierbarkeit als medieninhärentes Merkmal, Ökonomie, höhere Repräsentativität und Diversität. Im Kurs 3401 wurden lediglich zwei Nachteile der internetgestützten Datenerhebung genannt: höhere Abbruchquoten und die Möglichkeit der wiederholten Teilnahme durch ein- und dieselbe Person. Darüber hinaus sollte u.a. auch die Frage nach der Äquivalenz von computer- bzw. internetgestützten Versionen vorliegender Persönlichkeitsinventare, die als *paper-pencil*-Versionen entwickelt wurden, gestellt werden (vgl. Fahrenberg, 2006).

Trotz der genannten potentiellen Probleme ist das Internet für die psychologische Diagnostik eine wichtige Ressource, deren Bedeutung in Zukunft sowohl für die Forschung als auch die Praxis vermutlich noch zunehmen wird. So können großangelegte interkulturelle Studien mit mehreren tausend Teilnehmerinnen und Teilnehmern im Internet mit relativ geringem finanziellem Aufwand durchgeführt werden. Ein anderes Beispiel ist das psychologiegestützte Recruiting im Internet (Wottawa et al., 2005). Bei dieser Form der web-basierten Personalanwerbung und -selektion werden auch berufsbezogene Persönlichkeitsmerkmale erhoben, deren Relevanz für die Eignungsdiagnostik unbestritten ist. Dabei geht es nicht nur darum, Persönlichkeitsmerkmale zu erheben und eignungsdiagnostisch zu verwerten, sondern

auch um die möglichst effiziente Rückmeldung der Ergebnisse an die Bewerberinnen und Bewerber und daraus abgeleitete Empfehlungen.

9.2 Ambulatorisches Assessment

Das Erleben und Verhalten eines Menschen wird in der psychologischen Diagnostik häufig nicht unmittelbar dort erfasst, wo es stattfindet – im Feld. Dagegen wird Erleben und Verhalten, das eigentlich im Feld stattfindet – sicherlich aus verständlichen ökonomischen Gründen – mit Hilfe eines Fragebogens retrospektiv und im Sinne einer mentalen Repräsentation erfasst. Solche Auskünfte im Fragebogen können als komplexe Rekonstruktionen des Erlebens und Verhaltens im Feld interpretiert werden, die von zahlreichen Verzerrungen überlagert werden (z.B. Erinnerungseffekte, Antworttendenzen und –heuristiken, Stereotype etc.). Kurzum: Mittels Fragebogen erfasstes Erleben und Verhalten ist nicht gleichbedeutend mit Erleben und Verhalten im Feld, d.h. im alltäglichen Kontext einer Person. Zudem kann sich Erleben und Verhalten das unter Laborbedingungen gezeigt und erfasst wird, ganz erheblich vom Erleben und Verhalten im Alltag unterscheiden. Ein gutes Beispiel dafür ist das aus der Medizin bekannte Phänomen der „Weißkittel-Hypertonie“. Demnach wird beim Arzt-im-weißen-Kittel bei ca. 10% der Personen fälschlicherweise ein höherer Blutdruck diagnostiziert, als er im Alltag tatsächlich der Fall ist. Deshalb wird mittlerweile bei Verdacht auf Hypertonie ein engmaschiges 24-Stunden-Blutdruck-Monitoring im Alltag durchgeführt, bevor eine in der Regel medikamentöse Therapie eingeleitet wird.

In der Psychologie sind die genannten Probleme eigentlich auch schon länger bekannt; aber erst in den letzten Jahren hat sich im Zuge des technologischen Fortschritts im Bereich von portablen Mikrocomputern und Handhelds das sogenannte ambulatorische/ambulante Assessment zunehmend mehr verbreitet. „Ambulantes Assessment erfasst Selbstberichte, Verhaltensweisen oder physiologische Messwerte mit computer-unterstützten Methoden, während die Untersuchten ihrem normalen Tagesablauf nachgehen“ (Fahrenberg et al., 2007, S. 12).

Zwei Beispiele (vgl. Renner, 2005, S. 150f) sollen verdeutlichen, was mit Ambulantes Assessment gemeint ist:

Zwei Beispiele

Beispiel 1: In einer Feldstudie von Buse und Pawlik (1994) protokollierten Schüler u.a. ihr Verhalten im Alltag mithilfe eines tragbaren Taschencomputers. Zu vorprogrammierten, unregelmäßigen Zeitpunkten (durchschnittlich einmal pro Stunde) wurden die Schüler durch ein Piepsen des Computers dazu aufgefordert, ihr in den vorausgegangenen fünf Minuten aufgetretenes Verhalten auf einer vorgegebenen Verhaltensliste zu registrieren. Dabei wurden verhaltensbezogene Items auf dem Computer-Display dargeboten, die mit „ja“ oder „nein“ zu beantworten waren (z.B. „für die Schule geübt“; „jemandem ein Kompliment gemacht“). Die in den Items beschriebenen Verhaltensakte gelten als prototypisch für bestimmte Eigenschaften

(z.B. Gewissenhaftigkeit, Freundlichkeit). Die Analyse der so gewonnenen Verhaltensdaten zielte auf eine zentrale persönlichkeitspsychologische Fragestellung ab: die Bestimmung der Stabilität bzw. Konsistenz der Verhaltensakte über verschiedene Zeitpunkte bzw. Situationen.

Beispiel 2: In einer Feldstudie von Mehl und Pennebaker (2003) wurden die „Sounds of Social Life“ erfasst. Aufgezeichnet wurden akustische Informationen (z.B. Gespräche) aus dem Alltag von 52 Studierenden, die im Abstand von vier Wochen je zwei Tage lang einen kleinen Audiorekorder bei sich trugen, der sich alle 12 Minuten für die Dauer von 30 Sekunden einschaltete. Die aufwändige Analyse der 30-Sekunden-Intervalle umfasste im ersten Schritt die Registrierung des Orts, der Aktivität und der Art der Interaktion sowie eine computergestützte Kategorisierung des Sprachgebrauchs der Probanden. Die anschließende Bestimmung der Stabilität und Konsistenz der so aufbereiteten Daten, basiert, anders als bei der Selbstbeobachtung des eigenen Verhaltens im Beispiel 1, auf der Perspektive eines außenstehenden „Zuhörers“.

Die technischen Möglichkeiten ambulatorischer Messungen entwickeln sich derzeit rasend schnell. Die Möglichkeiten verschiedene Arten von Daten zu gewinnen weiten sich ständig aus und umfassen inzwischen die Bereiche, die in der folgenden – sicherlich nicht umfassenden - Liste genannt sind:

- physiologische Maße wie Blutdruck, Herzrate, Atmung, Hautleitfähigkeit oder EEG
- Fragebogendaten durch Applizierung auf Handhelds oder per SMS auf Handys (vgl. Beispiel 1)
- Akustische Aufnahmen (vgl. Beispiel 2) oder optische Aufnahmen, wobei hier besondere Rücksicht genommen werden muss auf die Privatsphäre und die Persönlichkeitsrechte des Versuchsteilnehmers, als auch von Personen, die sich in der Umgebung aufhalten und somit auch aufgezeichnet werden
- Bewegungsmaße durch Beschleunigungsmesser
- Lokalisationen durch GPS-Systeme

Mittlerweile gibt es auch eine „Society for Ambulatory Assessment“, auf deren Website Sie weitere Informationen zum Thema finden:

<http://www.ambulatory-assessment.org/>

9.3 Renaissance der objektiven Persönlichkeitsdiagnostik

Wie im Abschnitt 4 bereits erwähnt, hat es in den letzten Jahren eine Renaissance in der Entwicklung und des Einsatzes von objektiven Tests gegeben. In einem aktuellen Sammelband von Ortner et al. (2007) werden objektive Persönlichkeitstests zu folgenden Konstrukten vorgestellt:

- Feldabhängigkeit
- Emotionale Belastbarkeit
- Reflexivität/Impulsivität (Arbeitshaltungen)
- Berufliche Interessen
- Spontanflexibilität
- Risikoverhalten
- Leistungsmotivation
- Realitätsangemessenheit der Selbsteinschätzung
- Hyperkinetisches Syndrom

Die Renaissance der objektiven Persönlichkeitsdiagnostik

Zudem werden in diesem Band auch sogenannte Implizite Assoziationstests (IAT) zur Erfassung von Ängstlichkeit und Selbstwertschätzung vorgestellt. Den IAT haben Sie vermutlich bereits im Sozialpsychologie-Modul kennen gelernt. Zum IAT sind in den letzten zehn Jahren eine wahre Flut von Forschungsartikeln publiziert worden. Ob IATs jemals auch in der diagnostischen Praxis eingesetzt werden können, ist aber nach wie vor eine offene Frage.

Impliziter Assoziationstest (IAT)

Implizite Assoziationstests zielen darauf ab interindividuelle Unterschiede in automatischen Bewertungen bestimmter Konzepte oder Kategorien (z.B. alt vs. jung) zu messen, die auf der Grundlage ihrer Assoziation mit bestimmten Attributen (z.B. gut vs. schlecht), zustande kommen. Beispiele für solche automatischen bzw. impliziten Bewertungen sind wertbehaftete Einstellungen gegenüber farbigen und weißen Menschen, gegenüber Inländern und Ausländern (vgl. Greenwald et al., 1998). Ein IAT umfasst in der Regel fünf Schritte, die bei Krohne und Hock (2015) im Abschnitt 11.6.2 sehr gut beschrieben werden.

Literaturempfehlungen

- Schmidt-Atzert, L., & Amelang, M. (2012). *Psychologische Diagnostik* (5. Aufl.). Springer.
- Eid, M., & Diener, E. (2006). *Handbook of multimethod measurement in psychology*. APA.
- Krohne, H. W., & Hock, M. (2015). *Psychologische Diagnostik. Grundlagen und Anwendungsfelder* (2. Aufl.). Kohlhammer.
- Kubinger, K. D. (2019). *Psychologische Diagnostik* (3. Aufl.). Hogrefe.
- Petermann, F., & Eid, M. (Hrsg.). (2006). *Handbuch der Psychologischen Diagnostik*. Hogrefe.
- Westhoff, K., Hagemester, C., Kersting, M., Lang, F., Moosbrugger, H., Reimann, G., & Stemmler, G. (Hrsg.). (2010). *Grundwissen für die berufsbezogene Eignungsbeurteilung nach DIN 33430* (3. Aufl.). Pabst.
- Moosbrugger, H., & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion* (2. Aufl.). Springer.

Zeitschriften

- Diagnostica
- Journal of Individual Differences
- European Journal of Psychological Assessment
- Psychological Assessment

Datenbanken und Internetquellen

- PSYINDEX Tests (via Bibliotheksserver der FernUniversität: Home > Datenbanken und Lieferdienste > Datenbanken)
- <https://www.testzentrale.de> (kommerzielles Angebot)
- <https://www.schuhfried.at> (kommerzielles Angebot)
- Diagnostische Verfahren in der Konstruktion (z.B. via Website der Fachgruppe Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik, <https://www.dgps.de/index.php?id=158>)

Zitierte Literatur

- Amelang, M., & Schmidt-Atzert, L. (2006). *Psychologische Diagnostik und Intervention*. Springer.
- Brähler E., Holling H., Leutner D., & Petermann F. (Hrsg.). (2002). *Brickenkamp Handbuch pädagogischer und psychologischer Tests* (2 Bde., 3., vollst. überarb. und erw. Aufl.). Hogrefe.
- Buse, L., & Pawlik, K. (1994). Differenzierung zwischen Tages-, Setting- und Situationskonsistenz ausgewählter Verhaltensmerkmale, Maßen der Aktivierung, des Befindens und der Stimmung in Alltagssituationen. *Diagnostica*, 40, 2–26.
- Cattell, R. B. (1958). What is “objective” in “objective personality tests?” *Journal of Counseling Psychology*, 5(4), 285–289. <https://doi.org/10.1037/h0046268>
- Cattell, R. B., & Warburton, F. W. (1967). *Objective personality and motivation tests. A theoretical introduction and practical compendium*. University of Illinois Press.
- Cronbach, L. J., & Furby, L. (1970). How we should measure “change“ - or should we? *Psychological Bulletin*, 74(1), 68–80. <https://doi.org/10.1037/h0029382>
- Eid, M., & Diener, E. (2006). *Handbook of multimethod measurement in psychology*. APA.
- Ellgring, H. (1996). Verhaltensbeurteilung als Methode der Differentiellen Psychologie. In K. Pawlik (Hrsg.), *Enzyklopädie der Psychologie. Differentielle Psychologie und Persönlichkeitsforschung* (Bd. 1, pp. 395–425). Hogrefe.
- Fahrenberg, J. (2006). Assessment in daily life. A review of computer-assisted methodologies and applications in psychology and psychophysiology, years 2000-2005. *Undefined*. <https://www.semanticscholar.org/paper/Assessment-in-daily-life.-A-Review-of-Methodologies-Fahrenberg/2dfd40f3c56515132c174eb755bdb5a2d7c4a87b>
- Fahrenberg, J., Myrtek, M., Pawlik, K., & Perrez, M. (2007). Ambulantes Assessment – Verhalten im Alltagskontext erfassen. Eine verhaltenswissenschaftliche Herausforderung an die Psychologie. *Psychologische Rundschau*, 58(1), 12–23. <https://doi.org/10.1026/0033-3042.58.1.12>
- Fydrich, T. (2012). Diagnostik in der Klinischen Psychologie. In L. Schmidt-Atzert & M. Amelang, *Psychologische Diagnostik* (pp. 503-535). Springer.
- Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgements based on offices and bedrooms. *Journal of Personality and Social Psychology*, 82(3), 379–398. <https://doi.org/10.1037/0022-3514.82.3.379>
- Greenwald, A. G., McGhee, D. E., & Schwarz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal*

- of Personality and Social Psychology*, 74(6), 1464–1480.
<https://doi.org/10.1037/0022-3514.74.6.1464>
- Guilford, J. P. (1954). *Psychometric Methods*. McGraw-Hill.
- Hathaway, S. R., & McKinley, J. C. (1943). *Manual for the Minnesota Multiphasic Personality Inventory*. Psychological Corporation.
- Hossiep, R., & Wottawa, H. (1993). Diagnostik. In A. Schorr (Hrsg.), *Handwörterbuch der Angewandten Psychologie* (pp. 131–136). Deutscher Psychologen-Verlag.
- Huber, H. P., & Pflieger, U. (1994). Zeitliche Stabilität von individualspezifischen kardiovaskulären Reaktionsmustern. In D. Bartussek & M. Amelang (Hrsg.), *Fortschritte der differentiellen Psychologie und psychologischen Diagnostik. Festschrift zum 60. Geburtstag von Prof. Dr. Kurt Pawlik* (pp. 161–170). Hogrefe.
- Jäger, R. S. (2003). Intervention. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselselbegriffe der Psychologischen Diagnostik* (pp. 230–231). Beltz/PVU.
- Jäger, R. S. (2006). Diagnostischer Prozess. In F. Petermann & M. Eid (Hrsg.), *Handbuch der Psychologischen Diagnostik* (pp. 89–96). Hogrefe.
- Jäger, R. S., & Petermann, F. (1999). *Psychologische Diagnostik. Ein Lehrbuch*. Beltz/PVU.
- Keßler, B. H. (1999). Daten aus dem Interview. In R. S. Jäger & F. Petermann (Hrsg.), *Psychologische Diagnostik. Ein Lehrbuch* (pp. 429–438). Beltz/PVU.
- Kici, G., & Westhoff, K. (2000). Anforderungen an psychologisch-diagnostische Interviews in der Praxis. *Report Psychologie*, 25, 428–436.
- Kristof, W. (1958). Statistische Prüfverfahren zur Beurteilung von Testprofilen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 5, 520–533.
- Krohne, H. W., & Hock, M. (2015). *Psychologische Diagnostik. Grundlagen und Anwendungsfelder*. Kohlhammer.
- Kubinger, K. D., & Ebenhöf, J. (1996). *Arbeitshaltungen - Kurze Testbatterie: Anspruchsniveau, Frustrationstoleranz, Leistungsmotivation, Impulsivität/Reflexivität*. [Software und Manual]. Swets.
- Kubinger, K. D. (2006). *Psychologische Diagnostik*. Hogrefe.
- Langer, I., & Schulz von Thun, F. (1974). *Messung komplexer Merkmale in Psychologie und Pädagogik*. Reinhardt.
- Marcus, B. (2003). Persönlichkeitstests in der Personalauswahl: Sind „sozial erwünschte“ Antworten wirklich nicht wünschenswert? *Zeitschrift für Psychologie*, 211(3), 138–148. <https://doi.org/10.1026/0044-3409.211.3.138>
- McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, 96(4), 690–702.
<https://doi.org/10.1037/0033-295X.96.4.690>

- Meehl, P. E. (1954). *Clinical vs. statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press.
- Mees, U. (1977). Einführung in die systematische Verhaltensbeobachtung. In U. Mees & H. Selg (Hrsg.), *Verhaltensbeobachtung und Verhaltensmodifikation* (pp. 14–32). Klett.
- Mehl, M. R., & Pennebaker, J. W. (2003). The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology*, 84(4), 857–870.
<https://doi.org/10.1037/0022-3514.84.4.857>
- Mummendey, H. D. (2003). *Die Fragebogen-Methode* (4., unveränderte Aufl.). Hogrefe.
- Ortner, T. M., Proyer, R. T., & Kubinger, K. D. (2007) (Hrsg.). *Theorie und Praxis objektiver Persönlichkeitstests*. Huber.
- Pawlik, K. (2006). Psychologische Diagnostik I: Methodische Grundlagen. In K. Pawlik, (Hrsg.), *Handbuch Psychologie. Wissenschaft – Anwendung - Berufsfelder* (pp. 555–562). Springer.
- Pawlik, K. (1976). *Diagnose der Diagnostik*. Klett.
- Reinecker, H. (1986). Methoden der Verhaltenstherapie. In Deutsche Gesellschaft für Verhaltenstherapie (Hrsg.): *Verhaltenstherapie - Theorien und Methoden - Forum 11* (S. 64–179). DGVT.
- Renner, K.-H. (2002). *Selbstinterpretation und Self-Modeling bei Redeängstlichkeit*. Göttingen: Hogrefe.
- Renner, K.-H. (2005). Verhaltensbeobachtung. In H. Weber & T. Rammsayer (Hrsg.), *Handbuch der Persönlichkeitspsychologie und Differentiellen Psychologie* (pp. 149–157). Hogrefe.
- Rogosa, D., Brandt D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin*, 92(3), 726–748.
<https://doi.org/10.1037/0033-2909.92.3.726>
- Rorschach, H. (1921). *Psychodiagnostik*. Bircher.
- Schmidt-Atzert, L., & Amelang, M. (2012). *Psychologische Diagnostik* (5. Aufl.). Springer.
- Schmitt, M., & Gschwendner, T. (2006). Regeln der Datenintegration. In F. Petermann & M. Eid (Hrsg.), *Handbuch der Psychologischen Diagnostik* (pp. 383–395). Hogrefe.
- Schultheiss, O. C., & Pang, J. S. (2007). Measuring implicit motives. In R. W. Robins, R. C. Fraley & R. Krueger (Eds.), *Handbook of Research Methods in Personality Psychology* (pp. 322–344). Guilford.

- Shedler, J., & Westen, D. (1998). Refining the measurement of Axis II: A Q-Sort-Procedure for Assessing Personality Pathology. *Assessment*, 5(4), 333–353. <https://doi.org/10.1177/107319119800500403>
- Stemmler, G., Hagemann, D., Amelang, M. & Spinath, F. (2016). *Differentielle Psychologie und Persönlichkeitsforschung* (8. Aufl.). Kohlhammer.
- Willet, J. B. (1989). Questions and answers in the measurement of change. *Review of Research in Education*, 15(1), 345–422. <https://doi.org/10.3102/0091732X015001345>
- Winter, D. G. (1973). *The power motive*. Free Press.
- Woodworth, R. S. (1918). *Personal data sheet*. Chicago: Stoelting.
- Wottawa, H., Kirbach, C., Montel, C., & Oenning, S. (2005). Psychologiegestütztes Recruiting im Internet: Stand und Perspektiven. *Report Psychologie*, 30, 160-166.