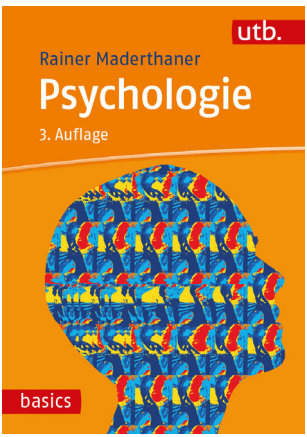
Lernunterlage zum Aufnahmeverfahren

Psychologie

Studienjahr 22/23

Die Kapitel zur Vorbereitung auf die Aufnahmeprüfung für das Studium der Psychologie werden Ihnen durch eine Initiative des Verfassers, der Universitäten Wien, Salzburg, Graz und Innsbruck sowie des facultas Universitätsverlages zur Verfügung gestellt und sind aus folgendem Buch entnommen:



Rainer Maderthaner
Psychologie
3., vollständig überarbeitete und ergänzte Auflage
utb:facultas 2021, 482 Seiten
ISBN: 978-3-8252-5540-4

Definition, Ziele und Positionen der Psychologie

2

Inhalt

2.1 Definitionen von Psychologie

2.2 Allgemeine Zielsetzungen wissenschaftlicher Psychologie
Beschreiben
Erklären
Vorhersagen
Verändern

2.3 Kontroversielle Grundannahmen der Psychologie
Leib – Seele
Anlage – Umwelt
Vergangenheit – Gegenwart
Freier Wille – Determiniertheit
Bewusst – unbewusst
Allgemeingültigkeit – Einzigartigkeit
Wertfreiheit – Wertbekenntnis
Objektivität – Subjektivität
Zergliederung – Ganzheitlichkeit
Statik – Dynamik
Quantitativ – qualitativ

2.4 Gegenwärtige Forschungsorientierungen der Psychologie

Definitionen von Psychologie

2.1

Das Wort **Psychologie** bedeutet, wie erwähnt, „Seelenkunde“ oder „Seelenlehre“ (griech. „psyche“: Hauch, Leben, Seele; griech. „logos“: Wort, Begriff). Die Auffassungen darüber, was unter Seele verstanden wird, unterscheiden sich jedoch ziemlich. Nachfolgend

sollen einige innerhalb des Wissenschaftsfaches Psychologie verbreitete Definitionen und Umschreibungen für „Psychologie“ präsentiert werden.

„Die Psychologie ist eine empirische Wissenschaft. [...] Ihr Gegenstand ist das (zumeist menschliche) Erleben und Verhalten, ihr Ziel ist es, allgemeingültige Aussagen über diesen Gegenstand zu machen – ihn zu beschreiben, beobachtbare Regelmäßigkeiten und Zusammenhänge aufzudecken, diese zu erklären, und womöglich Vorhersagen zu machen“ (Hofstätter & Wendt, 1974, 1). In ähnlicher Weise versteht Traxel (1974, 15) die Psychologie als Erfahrungswissenschaft, die als ein „System methodisch gewonnener Aussagen über einen bestimmten Gegenstand“ zu definieren ist.

Merksatz

Psychologie untersucht die Zustände und Veränderungen des Verhaltens, des Erlebens und des Bewusstseins.

Als zentral für die Definition von Psychologie wird oft die Angabe des Forschungsgegenstands angesehen, mit dem sich das Fach zu beschäftigen hat. Bourne und Ekstrand (1992, 2) formulieren: „Die Psychologie ist die wissenschaftliche Erforschung von Verhalten.“ Bei dieser breiten Definition könnte das Missverständnis entstehen, es sei nur das „äußere“ (beobachtbare) Verhalten gemeint. In Rohrachers international viel beachtetem Werk „Einführung in die Psychologie“ gelten dagegen die bewussten Prozesse mit ihren Auslösern und Effekten als Hauptcharakteristikum des Forschungsfelds der Psychologie: „Psychologie ist die Wissenschaft, welche die bewußten Vorgänge und Zustände sowie ihre Ursachen und Wirkungen untersucht“ (Rohracher, 1965, 7). Hier werden die zahlreichen unbewussten, automatisch ablaufenden psychischen Vorgänge noch vernachlässigt, zumindest aber ergibt sich eine Abgrenzung zu anderen Humanwissenschaften.

Zimbardo und Gerrig (1999, 2) definieren: „Gegenstand der Psychologie sind Verhalten, Erleben und Bewusstsein des Menschen, deren Entwicklung über die Lebensspanne und deren innere (im Individuum angesiedelte) und äußere (in der Umwelt lokalisierte) Bedingungen und Ursachen.“ Diese Definition ist bereits spezifischer. Die Bedeutung „innerer“ (introspektiver) Prozesse für die psychologische Forschung – der europäischen Tradition entsprechend – wird ebenso angesprochen wie der Aspekt des „Interaktionismus“ mit Einflüssen seitens der Umwelt.

Mandler (1979, 32) dagegen formuliert: **Psyche** ist ein komplexes, einem Individuum zugeschriebenes Informationsverarbeitungssystem, „das Input verarbeitet (einschließlich dem Input aus seinen eigenen Handlungen und Erfahrungen) und Output an die verschiedenen Subsysteme und die Außenwelt abgibt." In dieser Umschreibung des Forschungsfeldes der Psychologie wird Mandler sowohl den unbewussten als auch den bewussten Prozessen gerecht, indem er die *Psyche* als komplexes Regulationssystem definiert, innerhalb dessen dem *Bewusstsein* nur eine „Lupenfunktion" zukommt (s. unten).

Dörner und Selg (1996, 20) definieren im Sinne der Kybernetik: Psychologie ist die „Wissenschaft von den offenen oder variablen **Regulationen**" (Bischof, 2016). Als „offen" werden Regulationen dann bezeichnet, wenn sie „nicht genau durch genetische Vorprogrammierungen" festgelegt sind (Dörner & Selg, 1996, 20). Gemeint sind kybernetische Regelsysteme, die sich plastisch entwickeln können (z.B. Lern- und Denkvorgänge) und nicht genetisch fixiert sind (z.B. Reflexe oder *Erbkoordinationen*). Dass die Unterscheidung zwischen variablen und stabilen Regulationen auf empirischer Basis – zumindest bis heute – noch äußerst schwerfällt, erschwert allerdings die Anwendung dieser Definition.

Regulation ist eine Steuerung, welche die Stabilität eines dynamischen Systems aufrechterhält.

Merksatz

Psychologie ist eine Erfahrungswissenschaft, die in möglichst erschöpfender Breite und mit möglichst großer Realitätsnähe die Psyche bzw. ihre „Produkte" erforscht, nämlich das Verhalten, Erleben und Bewusstsein von Lebewesen.

Dörner und Selg (1996, 24) formulieren weiter: „Gegenstand der Psychologie kann alles werden, was erlebbar ist und / oder sich im Verhalten äußert [...]". Übereinstimmend mit einigen vorigen Definitionen werden hier introspektives Erleben und beobachtbares Verhalten als gleichwertige Datenquellen der Psychologie verstanden. Vorteilhaft an dieser breiten, aber pragmatischen Definition erscheint außerdem ihre Orientierung in Richtung **Interdisziplinarität** und **Transdisziplinarität**, ohne die eine erschöpfende und realitätsnahe Erklärung psychischer Phänomene kaum möglich ist.

Interdisziplinarität ist die Zusammenarbeit verschiedener Wissenschaftsdisziplinen zur Lösung eines Problems. Transdisziplinarität erfordert den Einbezug von Praktikerinnen und Praktikern in den wissenschaftlichen Diskurs.

2.2 Allgemeine Zielsetzungen wissenschaftlicher Psychologie

In verbreiteten Einführungswerken der Psychologie (vgl. etwa Bourne & Ekstrand, 2005; Gerrig & Zimbardo, 2008; Ulich, 2000)

Box 2.1 | Häufige Artefakte bei Befragungen

- Unklarheiten in der Formulierung von Fragen (z.B. Mehrdeutigkeit, zu komplizierte Sätze)
- Fehlinterpretationen von Anweisungen („Instruktionen“)
- *Sequenzeffekte* (Ermüdung, „Trainingseffekte“)
- *Hawthorne-Effekt* (sich beobachtet oder analysiert zu fühlen, erhöht zumeist die Leistungsbereitschaft)
- Mangelnde Bereitschaft zur Selbstenthüllung (bei privaten Inhalten)
- Motive zur *Selbstdarstellung,* Effekt der *sozialen Erwünschtheit* (bei Interviewpartnerinnen und -partnern einer Befragung einen bestimmten Eindruck zu hinterlassen, sich nicht zu blamieren etc.)
- Befürchtung negativer Konsequenzen (Zweifel an anonymer Verarbeitung der Daten)
- *Sponsorship-Bias* (Vermutungen über die Absichten der Auftraggeberinnen und -geber von Befragungen)
- *Kontext-Effekte* (z.B. Einfluss von Stimmungen)
- *Urteilsheuristiken* (pragmatische, zeitsparende und oft unlogische Art der Schlussfolgerungen)
- *Anwesenheitseffekte* (Beeinflussung des Antwortverhaltens durch anwesende Personen)

In Anlehnung an Bortz & Döring (1995)

finden sich – gut vergleichbar mit anderen empirischen Sozial- und Humanwissenschaften (wie etwa der Soziologie, der Ökonomie oder der Medizin) – vier Hauptziele für die Wissenschaftsdisziplin Psychologie:

2.2.1 | Beschreiben

Darunter versteht man das (möglichst) präzise, systematische und theoriegeleitete Erfassen von Informationen (*Daten*) über die zu untersuchenden psychischen Phänomene. Häufig verwendete Erhebungsverfahren sind *Selbst- und Fremdbeobachtungen*, *Befragungen*

(*Interviews*), *Experimente*, *Tests*, *nichtreaktive Verfahren* (z.B. Archive, Abnützungsgrad von Böden oder Gebrauchsgegenständen), *Textanalysen* (z.B. Tagebücher), *Inhaltsanalysen* (Häufigkeit und Bedeutung verwendeter Begriffe), *Skalierungen* (*Semantisches Differential* bzw. *Polaritätsprofil*), *Simulationen* (z.B. Computermodelle, Szenarien), *hirnelektrische Ableitungen* (z.B. *EEG*), *Messungen* (z.B. Reaktionszeiten) oder *Labordaten* (z.B. blutchemische Werte). Die Auswahl der Beschreibungsmittel von psychologischen Phänomenen richtet sich primär nach der wissenschaftlichen Grundorientierung der forschenden Person, nach der Art des Phänomens, und bei quantitativen Daten auch nach deren statistischer Verwertbarkeit.

Als **Objektivitätsproblem** bezeichnet man die Schwierigkeit, Daten unverfälscht zu erfassen (Box 2.1). Bei diagnostischen Verfahren zur Beschreibung von Störungsbildern oder Personenmerkmalen werden hohe *Gütekriterien* gefordert, die sinngemäß für alle psychologischen Datenerhebungen gelten (s. 3.6):

1. *Objektivität:* Sie ist umso größer, je ähnlicher die Daten bei unterschiedlichen datenerhebenden Personen sind.
2. *Reliabilität*: Die sogenannte „Zuverlässigkeit" von Daten ist umso größer, mit je weniger Erhebungsfehlern sie überlagert sind.
3. *Validität*: Die „Gültigkeit" von Daten nimmt in dem Maße zu, in dem sie tatsächlich jene Eigenschaft beschreiben, die registriert werden soll (z.B. Intelligenz und nicht auch Konzentration oder Bildung).

Daneben sollten jedoch noch weitere Qualitätsanforderungen an psychologische Daten gestellt werden, nämlich bezüglich der *Skalierung* (Wiedergabe korrekter Quantitäten), der *Normierung* (Normen bzw. Bezugssysteme für Ergebnisse sollen vorhanden sein), der *Fairness* (Daten über verschiedene soziale Gruppen dürfen nicht systematisch verfälscht sein), der *Ökonomie* (der Aufwand der Datenerhebung soll vertretbar sein), der *Zumutbarkeit* (Konsequenzen für Probanden sowie deren Akzeptanz sind zu berücksichtigen), der *Unverfälschbarkeit* (Ergebnisse sollen nicht manipulierbar sein) und der *Nützlichkeit* (Daten sollen zweckentsprechend sein).

Merksatz

Die Beschreibung von Forschungsphänomenen in der Psychologie (Datenerhebung) geschieht hauptsächlich über Selbst- und Fremdbeobachtung, Befragung, Messung, Experiment, Test, Textanalyse, Inhaltsanalyse, Skalierung, Simulation oder Fallstudien, wobei einer verfälschungsfreien Erfassung der Daten besondere Beachtung geschenkt wird (Gütekriterien).

2.2.2 | Erklären

Eine zweite wichtige Zielsetzung der Psychologie ist die Erklärung der beobachteten oder gemessenen Phänomene. Dies geschieht durch **Gesetze** oder durch deren Zusammenfassungen, die **Theorien**. Diese werden durch Ableitung von **Hypothesen** über zu erwartende Ergebnisse in empirischen Untersuchungen getestet. Die Resultate dieser Befragungen, Experimente oder Beobachtungen werden inhaltlich interpretierend (*qualitativ*) oder statistisch (*quantitativ*) auf Gesetzlichkeiten überprüft und mit den hypothetisch postulierten Zusammenhängen verglichen. Stimmen die empirisch gefundenen Zusammenhänge mit den erwarteten überein, dann spricht man von einer **Verifikation** der Hypothesen, im gegenteiligen Fall von deren **Falsifikation**. Eine solche Hypothesentestung setzt die Formulierung einer Theorie oder zumindest die Vorannahme einer Gesetzlichkeit voraus. In diesem Falle spricht man von einer **konfirmativen** (bestätigenden) Vorgangsweise, im Gegensatz zu einem **explorativen** Verfahren, wenn es darum geht, an einem Pool gewonnener Daten unbekannte Zusammenhänge erst zu finden.

Merksatz

Hypothesen sind wissenschaftlich begründete Annahmen (Wenn-dann-Aussagen) über Zusammenhänge von Ereignissen. Bestätigte Hypothesen nennt man Gesetze. Als Theorie bezeichnet man zumeist ein System von Gesetzen.

Gesetze und Hypothesen sind zumeist in Form von „Wenn-dann-Aussagen" formuliert und beziehen sich auf vermutete Kausalzusammenhänge in der Realität. Die „Wenn-Komponente" von Hypothesen beschreibt jeweils die Ursachen, Bedingungen oder Auslöser von Wirkungen, während die Effekte oder ausgelösten Veränderungen in der „Dann-Komponente" formuliert werden (Box 2.2; Westermann, 2000). Ein Beispiel eines Gesetzes aus der Kognitionsforschung (*Yerkes-Dodson-Gesetz*): Eine zu hohe oder zu niedrige psychophysiologische Aktivierung (Wenn-Komponente) verringert die Konzentrations-, Denk- und Gedächtnisleistungen (Dann-Komponente).

Merksatz

Wichtige Qualitätskriterien für Gesetze und Theorien sind ihr Grad an Repräsentativität, ihr Realitätsbezug sowie ihre zeitliche und situative Stabilität.

Ein grundlegendes Problem bei der Interpretation von Untersuchungsergebnissen, das sogenannte **Repräsentativitätsproblem**, ist die Frage nach der Verallgemeinerbarkeit, nämlich danach, wie gut von den jeweils beobachteten Daten – den

Fällen der *Stichprobe* – auf die Grundgesamtheit bzw. *Population* zu schließen ist (s. auch Replikationsproblem; Open Science Collaboration, 2015).

Eine andere Unsicherheit besteht darin, ob die abstrakt formulierten Theorien eine inhaltliche Entsprechung in den empirisch ausgewählten Untersuchungsverfahren finden: Die Rede ist vom **Operationalisierungsproblem** bzw. **Validitätsproblem**. Hier geht es etwa darum, ob die *Intelligenz* eines Menschen (d.h. die abstrakte Annahme über die geistige Leistungsfähigkeit einer Person) tatsächlich durch spezielle Intelligenzaufgaben eines Tests erfassbar bzw. ob die theoretische Vorstellung über Intelligenz anhand von anschaulich-konkreten Daten überprüfbar ist.

Weiters ist im sozialwissenschaftlichen Bereich kaum davon auszugehen, dass eine einmal gefundene Gesetzmäßigkeit an allen möglichen Orten, zu allen möglichen Zeiten und unter allen möglichen Umständen gilt, was als **Reliabilitätsproblem** bezeichnet wird (Bortz & Döring, 1995; Schnell, Hill & Esser, 2005). Zur Überprüfung der *Reliabilität* von Ergebnissen bedient man sich verschiedener statistisch gestützter Methoden, bei denen zum Beispiel ein Test für eine psychische Eigenschaft bei gleichen Personen wiederholt eingesetzt wird („Retest-Reliabilität") oder die Ergebnisse verschiedener Tests zur gleichen Eigenschaft miteinander verglichen werden („Paralleltest-Reliabilität").

Reliabilität = bedingungsunabhängige Verlässlichkeit einer Datenerhebung

Eine für die Verlässlichkeit von Forschungsergebnissen wichtige Bedingung ist deren Replizierbarkeit bzw. Reproduzierbarkeit, welche in letzter Zeit in mehreren Disziplinen nicht zufriedenstellend ausfiel (Baker, 2016), sodass Verbesserungen im Forschungsprozess vorgeschlagen wurden (Erdfelder & Ulrich, 2018; Fiedler, 2018).

Vorhersagen

2.2.3

Die Formulierung von Gesetzen dient auch zur Erstellung von Prognosen. Wenn zum Beispiel über einen spezifischen Sachverhalt Informationen gegeben sind, dann können unter Verwendung der psychologischen Gesetze Rückschlüsse auf weitere nicht bekannte Merkmale des Sachverhalts gezogen werden (Box 2.2). In der Fachliteratur ist der Fundus an psychologischen Vorhersagen unüberschaubar groß, und die Prognosegüte für zahlreiche Praxissituationen ist vielversprechend (Frey, Hoyos & Stahlberg, 1988;

Box 2.2 | Prognosen durch Gesetze

Bekannte **Vorinformationen** (Prämissen):
Person X ist sprachbegabt.
Person X ist lernmotiviert.
Person X hat gute Lernbedingungen.

Gesetze (Prämissen):
Wenn eine Person sprachbegabt und lernmotiviert ist sowie gute Lernbedingungen vorfindet, dann erzielt sie höchstwahrscheinlich gute Lernleistungen in Fremdsprachen.

Schlussfolgerung (Konklusion): Person X wird sehr wahrscheinlich gute Lernleistungen in Fremdsprachen erbringen.

Baumann & Perrez, 1990, 1991; Schwarzer, 1997; Hellbrück & Fischer, 1999; Süss & Negri, 2019 usw.): Welche Erziehungsmaßnahmen fördern eine gesunde Entwicklung von Kindern und Jugendlichen? Welche häuslichen Bedingungen sind Voraussetzungen für gute Schulleistungen? Welche Kommunikationsformen erleichtern die berufliche Kooperation? Welche Einflüsse hat die Lebensumwelt auf das Wohlbefinden? Wie kann man am besten Ängsten und Depressionen begegnen? Wie lernt man am schnellsten große Stoffmengen?

Merksatz

Aus psychologischen Gesetzen können vielfältige Vorhersagen über psychische Strukturen oder Abläufe und über deren Abhängigkeit von Umweltbedingungen abgeleitet werden.

Grundsätzlich können Vorhersagen über die Struktur von psychischen Phänomenen (z.B. Intelligenzstruktur, Persönlichkeitsstruktur, Einstellungsprofil) und über deren Dynamik (z.B. Reifungsprozesse, geistige Entwicklung, Entstehung psychischer Störungen) getroffen werden. Ähnlich wie bei politischen Wahlprognosen hängt auch im psychischen Bereich der Erfolg der Vorhersagen wesentlich von der Güte der verwendeten Theorien und der mathematisch-statistischen Prognoseverfahren ab.

Verändern

2.2.4

Eine dauerhafte Veränderung bzw. Optimierung menschlichen Erlebens und Verhaltens (die Veränderung von Gefühlen, Einstellungen, Motiven, Entscheidungen etc.) lässt sich in den meisten Fällen nicht allein durch Vermittlung von Einsichten (z.B. über Kindheitstraumen), durch Anwendung „psychologischer Tricks“ (z.B. *paradoxe Intervention*) oder durch einzelne suggestive Maßnahmen (z.B. *Hypnose*) bewerkstelligen, sondern es sind sehr oft komplexe Vorgangsweisen nötig. Dabei müssen nicht nur die Klientinnen und Klienten, sondern auch deren soziale und physische Umfelder einbezogen werden. Der beratenden, pädagogischen oder therapeutischen Anwendung solcher Veränderungsprogramme gehen oft umfangreiche Studien an hunderten Versuchspersonen voraus, um den Erfolg unter möglichst vielen Bedingungen sicherzustellen.

| Box 2.3

Arten psychologischer Intervention

- Beobachtungen und Befragungen (haben an sich schon indirekte Auswirkungen, z.B. durch Reflektieren des eigenen Verhaltens oder durch Problematisieren von Befragungsinhalten)
- Kommunikationsstil (kann meinungsbildend, kommunikationsfördernd und konfliktlösend wirken, z.B. durch Maßnahmen der Moderation oder *Mediation*)
- Aufklärung und Bildung (vermittelt psychologisches Wissen und Können, z.B. über optimales Lernen, Möglichkeiten der Stressbewältigung)
- Beratung (Schulberatung, Berufsberatung, Erziehungsberatung, Coaching etc.)
- Training (z.B. Entspannungstechniken, Lerntechniken, Kommunikations- und Kooperationstraining, Elterntraining)
- Therapie (z.B. Kognitive Verhaltenstherapie, Gesprächstherapie)
- Umweltgestaltung und *Partizipation* (z.B. Mitwirkung bei Planungen für menschengerechtes Wohnen und Siedeln, für eine humane Arbeitsplatzgestaltung oder für eine zukunftsfähige Mobilität)

Die Liste möglicher psychologischer Einflussnahmen ist relativ groß und beginnt schon damit, dass Personen sich anders verhalten, wenn sie sich beobachtet fühlen. Ein gutes Beispiel dafür ist der sogenannte **Hawthorne-Effekt**, nach den amerikanischen „Western Electric Hawthorne Works" in Chicago benannt, einer Fabrik, in der in den Jahren 1924 bis 1927 Elton Mayo den Einfluss von Arbeitsbedingungen auf die Produktivität testete: Er kam zum Schluss, dass man mehr zu leisten bereit ist, wenn man sich (z.B. im Zuge einer wissenschaftlichen Untersuchung) beobachtet fühlt (Flick et al., 1995).

Der Kommunikationsstil nimmt in mehrfacher Weise Einfluss: Kommunikationspsychologisch geschulte Moderatorinnen und Moderatoren können in Diskussionsrunden den Meinungsaustausch und die Konfliktbewältigung wesentlich dadurch fördern, dass sie eine partnerschaftliche Atmosphäre mit fairen Regeln für die Argumentation schaffen (Montada & Kals, 2013). Durch Kommunikationstechniken – wie dem „Partnerzentrierten Gespräch", dem „Kontrollierten Dialog" und der „Themenzentrierten Interaktion" – sinkt in Partnerschaften, Arbeitsgemeinschaften und Firmen die Streithäufigkeit, während die Kooperationsfähigkeit steigt.

Bereits wesentlich aufwendiger gestaltet sich der Einsatz psychologischen Wissens für Beratungstätigkeiten im Bildungs-, Arbeits- und Therapiebereich. Ähnlich wie in der Medizin werden hier vorerst die jeweiligen Ausgangsbedingungen erhoben (*Anamnese*) und die Probleme und Störungen festgestellt bzw. analysiert (*Diagnose*). In manchen Fällen sind umfangreiche Testungen, wie etwa zur Feststellung der Begabungsorientierung, der Interessenausrichtung, des Motivationsprofils oder der allgemeinen Problemsituation der Klientinnen und Klienten, nötig.

Besonders spezialisiert und auf die Art und Bedingtheit der behandelten Störung maßgeschneidert (s. Reinecker, 2003a) sind die in der empirischen Psychologie entwickelten Therapieverfahren (insbesondere Verhaltenstherapie, Gesprächstherapie, Familientherapie). Sie sind im Rahmen eines Psychologiestudiums aufgrund des hohen Übungsbedarfs und der nötigen Supervision nicht ausreichend lern- und trainierbar und müssen daher in anspruchsvollen

Merksatz

Auf Basis psychologischer Gesetze und Theorien konnte eine große Vielfalt von Maßnahmen (Interventionen) zur Veränderung problematischen Verhaltens, Erlebens und Bewusstseins entwickelt werden.

Zusatzausbildungen nach dem Studium vermittelt werden (z.B. in Österreich die postgraduale Ausbildung für „Klinische und Gesundheitspsychologie“ und / oder für „Psychotherapie“).

Weitere nicht unwichtige Einflussmöglichkeiten der Psychologie liegen im Bereich der Evaluation (Wottawa & Thierau, 2003) und Intervention im Wohn-, Wirtschafts-, Arbeits- und Bildungsbereich. Über die sogenannte „User Needs Analysis“ (UNA), „Post Occupancy Evaluation“ (POE) oder „Environmental Impact Analysis“ (EIA) lassen sich zum Beispiel wichtige Lebensbedürfnisse des Menschen ermitteln und Vorschläge für deren Befriedigung erarbeiten sowie eine allgemeine Verbesserung der Lebensumstände schaffen (Harloff, 1993). In neuerer Zeit werden immer mehr moderne Technologien auf ihre psychologische Nutzbarkeit hin untersucht (EDV-Arbeitsplätze, Internet-Aktivitäten, E-Learning, Teleworking etc.).

Kontroversielle Grundannahmen der Psychologie | 2.3

Seit etwa 1960 hat sich an den universitären Psychologieinstituten im deutschen Sprachraum eine „Mainstream-Psychologie“ durchgesetzt, nämlich jene mit naturwissenschaftlicher, empirisch-statistischer Orientierung. In den Achtzigerjahren meinte Hofstätter (1984, 103): „Die Konflikte zwischen den Richtungen und Schulen gehören fast überall in der Psychologie der Vergangenheit an“, und begründete dies damit, dass kaum mehr der Anspruch erhoben werde, mit „gleichen Prinzipien die verschiedenen Problemfelder“ der Psychologie aufzuklären. Tatsächlich ist seit den späten Sechzigerjahren die Heftigkeit der Auseinandersetzung zwischen den verschiedenen wissenschaftlichen Strömungen innerhalb der Psychologie erheblich zurückgegangen, was aber nicht gleichzeitig bedeutet, dass die wechselseitige Akzeptanz zugenommen hätte. Immer noch bestehen zwischen Angehörigen des Faches – vor allem aber zwischen Praktikern und Forschern – erhebliche Meinungsunterschiede darüber, was unter wissenschaftlichem Vorgehen zu verstehen ist, welche Themen als forschungswürdig anzusehen sind und was als allgemein verbindlicher Wissensbestand der Psychologie zu betrachten ist.

Auch wenn sich Wissenschaftlerinnen und Wissenschaftler zumeist nicht explizit zu ihren inhaltlichen oder methodischen Grundannahmen (Forschungsaxiomen) bekennen, so lassen sich

3.1 | Wissenschaftlichkeit

Wissenschaftliches Handeln sollte sich an logisch begründeten, explizit formulierten und verbindlichen Kriterien orientieren. Nach Wohlgenannt (1969) sowie Konegen und Sondergeld (1985) sind dies folgende:

- Es sollen nur Aussagen über Sachverhalte gemacht werden, die wirklich vorhanden sind (Beobachtbarkeit bzw. Erlebbarkeit).
- Die Aussagen sollen ein System bilden und nach expliziten (wissenschaftsspezifischen) Regeln zustande kommen.
- Es müssen Regeln zur Definition von Fachausdrücken (Termini) vorhanden sein.
- Für das gegebene System von Aussagen müssen Ableitungsregeln gelten („induktive“ und „deduktive“ Schlussregeln).
- Das Aussagensystem muss widerspruchsfrei sein.
- Aussagensysteme mit empirischem Bezug (faktische Aussagen) dürfen sich nicht auf die Aufzählung von Fakten beschränken, sondern müssen auch Verallgemeinerungen enthalten.
- Faktische Aussagen müssen intersubjektiv prüfbar sein.

Merksatz

Wissenschaftliches Vorgehen will für Tatsachen (Fakten) ein möglichst widerspruchsfreies System von mehr oder weniger abstrakten, logisch verknüpften und intersubjektiv prüfbaren Aussagen bilden.

In ähnlicher Weise charakterisieren Bortz und Döring (1995, 7) aus Sicht der Psychologie wissenschaftliche Aussagen: „Wissenschaftliche Hypothesen sind Annahmen über reale Sachverhalte (empirischer Gehalt, empirische Untersuchbarkeit) in Form von Konditionalsätzen. Sie weisen über den Einzelfall hinaus (Generalisierbarkeit, Allgemeinheitsgrad) und sind durch Erfahrungsdaten widerlegbar (Falsifizierbarkeit).“

3.2 | Von der Empirie zur Theorie

Merksatz

Die Methoden einer wissenschaftlichen Disziplin sollen die korrekte und zweckmäßige „Abbildung“ eines empirischen (konkreten) Systems in einem theoretischen (abstrakten) System erlauben.

Die Human- und Sozialwissenschaft Psychologie unterscheidet sich insofern grundlegend von den Naturwissenschaften, als hier die wissenschaftlichen Phänomene nicht direkt zugänglich sind, sondern oft indirekt erschlossen werden müssen. So sind etwa Persönlichkeit, Intelligenz oder Einstellungen theoretische Be-

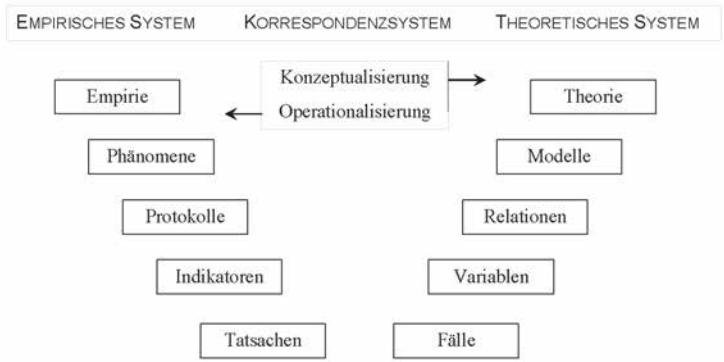


Abb 3.1

Die theoretische Beschreibung der Realität kann als deren abstrakte Abbildung in einem Symbolsystem (Sprache, Vorstellung, Programme ...) aufgefasst werden. Dabei wird ein vermittelndes, transformierendes Korrespondenzsystem benötigt, wodurch wissenschaftscharakteristische Vorschriften zur Gewinnung, Beschreibung, Erklärung und Interpretation der jeweiligen Systemelemente zur Verfügung gestellt werden.

griffe, die nur über Verhaltenstendenzen, Fertigkeiten oder Gefühlsreaktionen erfasst werden können. Der Weg von der Empirie zur Theorie ist daher in der Psychologie oft weit und erfordert viele Zwischenschritte. In der psychologischen Methodenlehre unterscheidet man zumeist ein **empirisches System,** das die Forschungsdaten liefert, und ein **theoretisches System**, das die Gesetze und Erklärungen zu formulieren gestattet, und bezeichnet die Vorgangsweisen, Methoden und Instrumente, die zwischen beiden eine Verbindung herstellen, als **Korrespondenzsystem**.

Die Schritte vom empirischen zum theoretischen System (Abb. 3.1) lassen sich wie folgt charakterisieren: Vorerst werden aus einer Vielzahl von Strukturen und Abläufen in der psychischen oder sozialen Realität – der sogenannten **Empirie** – jene Phänomene identifiziert, die Gegenstand der wissenschaftlichen Untersuchungen werden sollen. Die jeweils zu erforschenden **Phänomene** (z.B. Denkprozesse, Lernformen, Stressverarbeitung) müssen exakt beschrieben werden, was sich in verbalen, bildlichen oder symbolischen Datenmengen bzw.

Merksatz

Im empirischen System werden Phänomene menschlicher Erfahrung ausgewählt und in ihren konkreten Erscheinungsweisen (Tatsachen) verbal oder symbolisch (über Indikatoren) protokolliert.

Protokollen niederschlägt. Insbesondere in der quantitativen Forschung versucht man die Datenmenge auf jene Informationseinheiten zu begrenzen, die zur Beschreibung der Gesetzmäßigkeiten der Phänomene relevant erscheinen. Die Gesamtheit aller Ausprägungen von **Indikatoren** (zu einem untersuchten Phänomen) zu einer bestimmten Zeit und an einem bestimmten Ort sind jene einzelnen **Tatsachen**, die der empirischen Forschung als Grundeinheiten zur Gewinnung oder Überprüfung von (statistischen) *Hypothesen* zur Verfügung stehen.

Zu einem **Fall** des theoretischen Systems wird eine Tatsache dann, wenn die erfassten Eigenschaften der Indikatoren in Ausprägungen von **Variablen** umgewandelt werden. Ein Fall ist somit durch eine bestimmte Konfiguration von (empirischen) Variablen, genauer durch deren jeweilige Ausprägungen, definiert. Mittels statistischer Auswertungsverfahren werden auf Basis der zur Verfügung stehenden Fälle zwischen den Variablen entweder hypothetische **Relationen** (Funktionen, Beziehungen etc.) geprüft oder unbekannte Relationen gesucht. **Hypothesen** sind Annahmen über *Relationen* zwischen mindestens zwei (empirischen) Variablen. Wenn eine hypothetische Relation zufriedenstellend oft in verschiedenen, wissenschaftlich seriösen Untersuchungen empirisch bestätigt wurde, spricht man von einem **Gesetz**. Mehrere Gesetze (oder Hypothesen), die ein logisch konsistentes Erklärungsgerüst für ein bestimmtes Phänomen darstellen, werden zusammenfassend als **Modell** bezeichnet (z.B. Wahrnehmungs-, Lern-, Gedächtnis- und Handlungsmodelle). Der Übergang von Modellen zu Theorien ist fließend. Eine **Theorie** ist ein System von zusammenhängenden Gesetzen, die maximal abstrakt formuliert sind.

Merksatz

Im theoretischen System werden anhand von Fällen Relationen zwischen Variablen gesucht, funktional zusammenhängende Relationen zu Modellen zusammengefasst und thematisch verwandte Modelle zu einer Theorie integriert.

Merksatz

Konstrukte sind speziell definierte, nicht direkt beobachtbare Begriffe einer psychologischen Theorie (z.B. Intelligenz, Motivation, Aggression), für die Operationalisierungen vorhanden sind oder entwickelt werden müssen.

Um die Realität in Form von Gesetzen oder Theorien abbilden zu können, müssen *Begriffe* (*Konzepte*) zur Klassifikation empirischer Phänomene entweder vorhanden sein (Alltagsbegriffe) oder neu entwickelt werden (Fachbegriffe bzw. Ter-

mini). Diese **Konzeptionalisierung** („Konzeptspezifikation"; Schnell, Hill & Esser, 1993) der Wahrnehmungs- oder Erlebenswelt darf weder zu fein noch zu grob ausfallen, damit ein adäquater Auflösungsgrad für die untersuchten Phänomene gegeben ist. Für neu eingeführte theoretische Fachbegriffe, sogenannte **Konstrukte** (d.h. theoretische Konstruktionen), ist die konkrete Bedeutung in der Welt unserer Erfahrungen mittels **Operationalisierungen** klarzulegen. Als solche Interpretationshilfen für theoretische Fachbegriffe können spezielle Beobachtungen, Testverfahren, Teile von Fragebögen oder sonstige Datenerfassungsverfahren herangezogen werden. Mögliche Operationalisierungen von „Angst" sind etwa bei einem Versuchstier der körperliche Zustand in Erwartung elektrischer Schläge, die gemessene Herzfrequenz oder die motorische Unruhe. „Intelligenz" kann durch die Leistungen in einem bestimmten Intelligenztest, und „Glück" durch die Beantwortung von Fragen in einem Befindlichkeitstest operationalisiert werden. Das Korrespondenzsystem mit einschlägigen Konzeptualisierungen und Operationalisierungen ist Bestandteil des jeweiligen *wissenschaftlichen Paradigmas* (s. auch Maderthaner, 2003).

Merksatz

Bedeutungsinterpretierende Zuordnung beobachtbarer Sachverhalte zu einem theoretischen Begriff bezeichnet man als dessen Operationalisierung.

Fälle und Variablen 3.3

Als **empirische Einheiten** kommen in der Psychologie beliebige statische oder dynamische Systeme infrage (z.B. Personen, Gruppen, Situationen, Abläufe), in denen sich psychische Gesetzmäßigkeiten äußern. Wie bereits erwähnt, wird die Beschreibung (*Protokoll*) eines *Phänomens* auf gesetzesrelevante Merkmale (*Indikatoren*) reduziert, sodass zuletzt nur mehr ein sogenannter „Fall" mit phänomencharakteristischen Variablen übrig bleibt. *Fälle* sind also die – im Sinne einer wissenschaftlichen Fragestellung – maximal informationsreduzierten empirischen Einheiten, anhand derer Gesetze verifiziert oder falsifiziert werden sollen.

Da es in der Psychologie nur selten möglich ist, die gesamte **Population** bzw.

Merksatz

In sozialwissenschaftlichen Untersuchungen werden anhand von Fällen (Stichprobe) Gesetze gewonnen, welche auf ähnliche Sachverhalte (Population, Geltungsbereich der Gesetze) hin verallgemeinert werden.

Grundgesamtheit empirischer Einheiten zu erfassen, für die ein Gesetz gelten soll, beschränkt man sich in der Forschung auf eine **Stichprobe** (engl. sample), deren Zusammensetzung in den gesetzesrelevanten Eigenschaften jener der Population möglichst ähnlich sein sollte, damit die auf Basis der Stichprobe gewonnenen Erkenntnisse berechtigt verallgemeinert werden können. Der Schluss von der Stichprobe auf die Population ist am ehesten dann gerechtfertigt, wenn die Stichprobe nach dem Zufallsprinzip aus der Grundgesamtheit ausgewählt wird (*Randomisierung*) und die Stichprobe entsprechend groß ist (s. auch Schnell et al., 1993).

Abb 3.2

Ein Beispiel für eine einfache, aber prägnante Charakterisierung von Personen (Fällen) ist jene nach Persönlichkeitsfaktoren (Variablen). Das Profil in der Abbildung kennzeichnet eine Person in den sogenannten „Big-Five-Faktoren" („NEO Five-Factor Inventory" von Costa & McGrae, 1992; Becker, 2004).

Wenn die Ausprägungen relevanter Untersuchungsvariablen in einer Stichprobe mit jenen der Population annähernd übereinstimmen, darf von **Repräsentativität** der Stichprobe gesprochen werden. Im Forschungsalltag ist Repräsentativität aufgrund verschiedenster Forschungshemmnisse nur selten vollständig erreichbar (Kostenbegrenzung, Unerreichbarkeit von Personen, Teilnahmeverweigerung etc.), sodass häufig nur **Gelegenheitsstichproben** (z.B. Studierendensamples) zur Verfügung stehen oder die Stichprobenselektion eher mittels **Quotaverfahren** (Vergleichbarkeit der Stichprobe

mit der Population hinsichtlich der Verteilung einiger wichtiger Merkmale wie Geschlecht, Bildung, Beruf usw.), mittels **Schneeballverfahren** (Probandinnen und Probanden vermitteln selbst wieder weitere Probandinnen und Probanden) oder mittels **Klumpenverfahren** erfolgt („cluster sampling": Cluster von Fällen, z.B. Unternehmen, Organisationen, Branchen, werden zufällig ausgewählt und hierin alle Mitglieder untersucht). Leider erhöhen die letztgenannten Auswahlverfahren die Fehleranfälligkeit und mindern den Grad an Verallgemeinerbarkeit.

In der Mathematik sind **Variablen** („Platzhalter", „Leerstellen") jene Zeichen in Formeln, die für einzelne Elemente aus einer Menge möglicher Zahlen oder Symbole stehen. Die verschiedenen Belegungen von Variablen nennt man ihre **Ausprägungen** oder – wenn diese aus Zahlen bestehen – ihre **Werte**. Als **Wertebereich** einer Variablen bezeichnet man alle Zahlen vom Minimal- bis zum Maximalwert. Variablen charakterisieren Fälle hinsichtlich ihrer untersuchungsrelevanten Merkmale. In psychologischen Untersuchungen können diese äußerst vielfältig sein und schließen Beschreibungsmerkmale, Testergebnisse, Prozentschätzungen, physiologische Messwerte und andere Aspekte mit ein (Abb. 3.2).

Während in der Mathematik Zahlen definitionsgemäß eine quantitative Bedeutung haben, das heißt, dass bestimmte Rechenoperationen mit ihnen durchgeführt werden können (Addition, Multiplikation, Potenzierung etc.), kann dies bei Variablenwerten der psychologischen Empirie nicht vorausgesetzt werden. Hier können Zahlen zum Beispiel für Benennungen herangezogen werden (z.B. Abzählung von Personen in einer Gruppe), sie können eine Rangordnung symbolisieren (z.B. der 1., 2. oder 3. in einem Wettkampf) oder sie können ein Vielfaches von Grundeinheiten darstellen (z.B. Häufigkeiten). Aus diesem Grund werden die Ausprägungen von Variablen in der Psychologie hinsichtlich ihrer sogenannten *Skalenqualität* unterschieden, wovon insbesondere die Anwendbarkeit statistischer Auswertungsverfahren abhängt.

Faktoren, denen innerhalb von Phänomenen ein Einfluss zugeschrieben wird, heißen in den empirischen Sozialwissenschaften (so wie in der Mathematik bei Funktionsgleichungen) **unabhängige Variablen (UV),** während jene Faktoren, welche die Auswirkungen des Einflusses symbo-

Merksatz

Hypothetische Ursachen werden in empirischen Untersuchungen mittels unabhängiger Variablen charakterisiert und hypothetische Wirkungen mittels abhängiger Variablen.

lisieren, als **abhängige Variablen (AV)** bezeichnet werden. In einer wissenschaftlichen *Kausalhypothese* (s. 3.4) stellt der Wenn-Teil die Ausprägungen der unabhängigen Variablen und der Dann-Teil die vorhergesagten Ausprägungen der abhängigen Variablen dar (Box 3.1). Auf diese Unterscheidung verzichtet man, wenn die Einflussrichtung zwischen den Variablen nicht spezifiziert ist oder als wechselseitig angenommen wird (z.B. bei *Korrelationsstudien*).

Von den eigentlichen Wirkvariablen unterscheidet man sogenannte **Moderatorvariablen**, denen ein modifizierender Einfluss auf die funktionalen Beziehungen zwischen unabhängigen und abhängigen Variablen zugeschrieben wird (Box 3.1).

Da selbst bei bestens geplanten und genau kontrollierten Experimenten Einflüsse wirksam werden, die nicht erwünscht sind, existieren in allen empirischen Untersuchungen auch **Störvariablen**. Je mehr Störeinflüsse in einer Untersuchung vorhanden sind, desto vager und unschärfer werden die wissenschaftlichen Resultate. Von **Konfundierung** spricht man, wenn unabhängige Variablen mit dem Effekt anderer Variablen vermischt sind. Gebräuchliche Maßnahmen gegen eine Verfälschung durch konfundierende Variablen oder Störvariablen sind deren

lat. confundere: zusammengießen, vermischen, vermengen, verwirren

- „Elimination“ (d.h. Versuch ihrer Ausschaltung),
- „Matching“ (d.h. Gleichhaltung ihres Effektes bei den Ausprägungen der unabhängigen und abhängigen Variablen) sowie
- *Randomisierung* (d.h. zufällige Aufteilung ihrer Quellen, wie etwa der Auswahl der *Probandinnen und Probanden*).

Insbesondere bei modernen statistischen Modellen findet man häufig die Unterscheidung in **manifeste** und **latente Variablen**. Als manifest gelten alle durch direkte Erhebung (als Ergebnis der empirischen Datenerhebung) zustande gekommenen Variablen, während latente Variablen theoretisch begründet sind und zur Erklärung der empirischen Resultate herangezogen werden. So etwa kann das *Konstrukt* Intelligenz durch eine latente Variable beschrieben werden, wenn diese als Summe aller gelösten Intelligenzaufgaben definiert wird. Die Ausprägungen latenter Variablen werden in der Forschungspraxis mittels mehr oder weniger komplexer mathematischer Prozeduren (z.B. über *Mittelwertsbildungen, lineare Funktionen, probabilistische Schätzungen*) aus den ihnen über die Operationalisierung zugeordneten manifesten Variablen errechnet.

probabilistisch: wahrscheinlichkeitstheoretisch berechnet

Beispiel für eine Variablentypisierung | Box 3.1

Wenn etwa in einem Experiment der Einfluss des Alkoholkonsums auf die Fahrleistung in einem Fahrsimulator untersucht werden soll, dann könnte die Hypothese lauten: Wenn Verkehrsteilnehmer Alkohol trinken, dann begehen sie überdurchschnittlich viele Fehler im Simulator. Als *unabhängige Variable* fungiert der Alkoholgehalt des Blutes, welcher zumindest in zwei Ausprägungen vorliegen muss (z.B. 0,0 Promille Blutalkoholgehalt – 0,5 Promille Blutalkoholgehalt). Als *abhängige Variable* könnte in einer normierten Fahrleistungsprüfung die Anzahl an Fahrfehlern herangezogen werden. Als *Moderatorvariablen*, welche die Beziehung zwischen Alkoholisierungsgrad und Fahrleistung verändern könnten, wären die Fahrpraxis, die Alkoholtoleranz oder die Trinkgeschwindigkeit der Versuchspersonen einzubeziehen. Als *Störvariablen* können Messfehler bei der Blutalkoholbestimmung, Konzentrationsschwankungen der Probandinnen und Probanden oder Ablenkungen in der Versuchssituation angenommen werden.

Kausalität und Wahrscheinlichkeit | 3.4

Die Annahme, dass Ereignisse der Realität einander gesetzmäßig beeinflussen, d.h. in einem Kausalzusammenhang zueinander stehen, wird implizit in jeder Wissenschaft vorausgesetzt. Würde die Welt nicht deterministischen oder zumindest probabilistischen Gesetzen (wie z.B. in der Quantenphysik) unterliegen, hätte das Betreiben von Wissenschaft keinen Sinn. Das **Kausal(itäts)prinzip**, nämlich die Annahme, dass jedes Ereignis eine oder mehrere Ursachen hat, ist eine grundsätzlich unbeweisbare These, die aber sowohl im Alltag als auch in der Wissenschaft dazu motiviert, immer wieder nach Ursachen und Wirkungen zu fragen. Im Vergleich zur oft trivial vereinfachten Kausalanalyse des täglichen Lebens (z.B.: Wer ist schuld an einer Scheidung? Was ist die Ursache eines Unfalls?) unterscheidet man in der Wissenschaft mehrere Arten von Kausalbeziehungen.

Box 3.3 | Beispiel für eine multivariate aussagenlogische Beschreibung von Variablenzusammenhängen (mit Multikausalität und bedingter Kausalität)

Inhaltliche Aussage (aus der Sozialpsychologie):

Kinder, die früher an Modellpersonen beobachten konnten, dass sich Aggression „lohnt“ (L), oder solche, die gerade von einem anderen Kind frustriert (F) wurden, tendieren diesem gegenüber zu aggressivem Verhalten (A), wenn dieses eher als wehrlos (W) empfunden wird, wenn ihm gegenüber keine moralischen Hemmungen (M) bestehen und wenn im Moment keine Strafdrohung (S) von Aufsichtspersonen für aggressive Reaktionen zu erwarten ist.

Aussagenlogische Form: $((L \vee F) \wedge W \wedge \neg M \wedge \neg S) \rightarrow A$

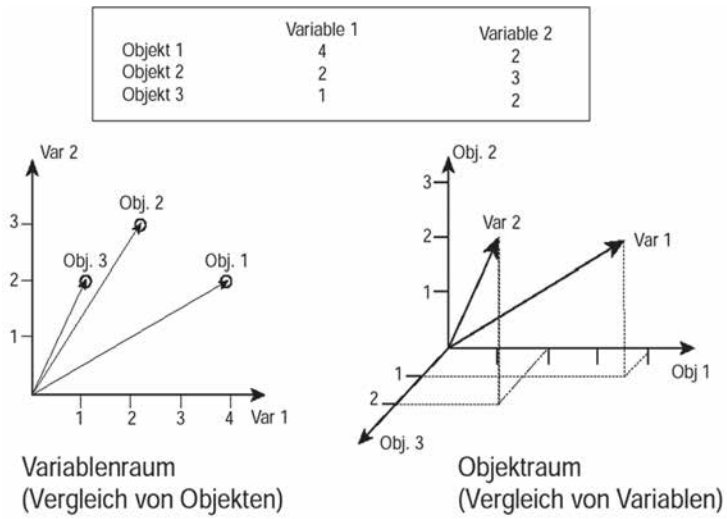
(Zur Bedeutung der Symbole siehe Abb. 3.6; $\neg$: Negation)

Als eines der größten methodischen Probleme der gegenwärtigen psychologischen Forschung kann gelten, dass die meisten Erklärungsmodelle und Hypothesen weder Multikausalität noch bedingte Kausalität einbeziehen, und daher zu geringe Prognosesicherheit erreichen (Maderthaner, in Vorbereitung).

3.6 | Beschreibende und hypothesenprüfende Statistik

Die **Statistik** (engl. statistics) fungiert innerhalb der Psychologie als Hilfswissenschaft zur Auffindung und Beschreibung von nichtdeterministischen Gesetzen (Relationen). “Statistics is a set of concepts, rules and methods for (1) collecting data, (2) analyzing data, and (3) drawing conclusions from data” (Iversen & Gergen, 1997, 4). Wenn in der Nachrichtentechnik damit gerechnet wird, dass elektronische Signale von Störungen überlagert werden, dann sendet man gleiche Signale mehrmals hintereinander, um beim Empfang auf Basis ihres *Mittelwerts* (Durchschnitt) auf das ursprüngliche „wahre“ Signal schließen zu können. Nach dem gleichen Prinzip

Abb 3.7



Ähnlich wie physikalische Objekte in einem (euklidischen) Raum Positionen einnehmen können, lassen sich auch Fälle als „Datenobjekte" auffassen und in einen (multidimensionalen) Variablenraum projizieren. Die Datenobjekte sind einander umso näher, je ähnlicher ihre Variablenausprägungen sind. Ebenso lässt sich die Ähnlichkeit von Variablen in Objekträumen abbilden. Im Beispiel sind drei Objekte (z.B. Personen) im Zweivariablenraum (z.B. Gewicht und Größe) dargestellt und daneben die gleichen Variablen im Dreiobjekteraum.

werden in der Psychologie wiederholt oder simultan Daten über psychische Abläufe, Einstellungen oder Fähigkeiten gesammelt, um daraus Schätzungen über die untersuchten Phänomene ableiten zu können. Die (klassische) **Testtheorie** postuliert diesem Prinzip gemäß, dass sich jeder **Messwert** (z.B. eine physiologische Ableitung, eine Fragebogenantwort, eine Prozentschätzung) aus einem **wahren Wert** und einem zufälligen **Fehlerwert** zusammensetzt und dass sich Fehlerüberlagerungen durch Heranziehung mehrerer Messwerte des gleichen Ereignisses „ausmitteln" lassen.

Merksatz

Mittels der Statistik als Hilfswissenschaft werden in der Psychologie verfügbare Daten beschrieben und auf vorhandene Gesetzmäßigkeiten untersucht.

3.6.1 | Deskriptivstatistik – beschreibende Statistik

Die deskriptive Statistik bietet charakteristische, formelhafte Beschreibungen oder grafische Darstellungen für eine große Zahl von Fällen, Variablen oder Variablenrelationen an. Wie schon erwähnt (Abb. 3.7), werden Fälle als Punkte in einem multidimensionalen Raum von Variablen gedacht, und ihre Verteilung wird durch **statistische Kennwerte** bzw. **Statistiken** näher charakterisiert. Eine Voraussetzung für eine solche Darstellung von Variablen (sowie ihres Einbezugs in komplexe statistische Auswertungsverfahren) ist, dass sie quantitativ interpretierbar sind, d.h., dass ihre Ausprägungen unterschiedliche Quantitäten einer Eigenschaft oder eines Merkmals von Fällen kennzeichnen (Backhaus et al., 2003).

Um die quantitative Bedeutung der Ausprägungen von Variablen einzustufen, werden diese hinsichtlich ihrer **Skalenqualität**, d.h. nach **Skalenniveau** bzw. **Messniveau**, differenziert:

1. **Nominalskala**: Wenn eine Variable nur dieses Skalenniveau zugeschrieben bekommt, sind ihre Ausprägungen (Zahlenwerte) im Sinne von Klassifikationen zu verstehen. Es handelt sich also um Variablen, die („qualitative") Eigenschaften, wie etwa Geschlecht, Beruf, Nationalität oder Haarfarbe, kennzeichnen.

Merksatz

Die quantitative Interpretierbarkeit von empirischen Variablen bzw. der durch sie beschriebenen Indikatorausprägungen wird durch die ihnen zugeschriebene Skalenqualität (Messniveau) charakterisiert.

2. **Ordinalskala (Rangskala)**: Die Werte von Variablen mit ordinaler Skalenqualität gestatten nicht mehr bloß die Unterscheidung zwischen gleich- und ungleichartig, sondern erlauben zusätzlich die Erstellung einer quantitativ begründeten Rangreihe der Variablenausprägungen. Typische Ordinalvariablen sind Listen von Schulnoten oder Rangreihungen bei Wettbewerben.
3. **Intervallskala**: Für Variablen dieses Typs wird angenommen, dass ihre aufeinanderfolgenden Zahlenwerte die Zunahme einer variablenspezifischen Eigenschaftsquantität immer um den gleichen Betrag symbolisieren (der Quantitätszuwachs von 1 auf 2 ist der gleiche wie etwa von 4 auf 5). In der Psychologie erwartet man zumindest Intervallskalenniveau von all jenen Variablen, die quantitative Abstufungen von individuellen Leistungspotenzialen (z.B. Konzentration, Intelligenz) oder von psychischen Dis-

positionen (z.B. Einstellungen, Persönlichkeitsdimensionen) zum Ausdruck bringen wollen.

4. **Verhältnisskala (Rationalskala)**: Variablen dieser Art sind gewissermaßen Intervallskalen mit einem fixen Nullpunkt. In der Psychologie gehören Verhältnisschätzungen für Wahrnehmungsreize diesem Skalentyp an oder bestimmte probabilistische Testkennwerte („Item-Response-Modelle").
5. **Absolutskala**: So bezeichnet man Variablentypen, die ebenfalls einen fixen Nullpunkt haben, bei denen aber auch „echte" Einheiten gegeben sind. Zu diesem Typus zählen alle Variablen, die Häufigkeiten bzw. Frequenzen zum Ausdruck bringen (z.B. Schätzungen der Anzahl von Objekten oder Personen).

Als quantitative Variablen im engeren Sinne zählen für die Statistik nur solche, die als Intervall-, Verhältnis- oder Absolutskalen zu interpretieren sind.

Eine empirische Variable hat noch eine weitere, für komplexe statistische Auswertungen wichtige Eigenschaft: die **Verteilung** ihrer Ausprägungen.

Die Betrachtung der Verteilung empirischer Variablen ist aus verschiedenen Gründen wichtig:

1. Aus ihr geht hervor, welche Zahlenwerte mit welcher Wahrscheinlichkeit in einer Population zu erwarten sind (z.B. Mess- oder Testergebnisse).
2. Sie kann Hinweise darüber geben, ob der Wertebereich einer Variablen für die Beschreibung eines empirischen Prozesses optimal gewählt wurde (z.B. nicht optimal bei „schiefen" Verteilungen, wenn sich die Werte bei den Minimal- oder Maximalwerten der Variablen häufen).
3. Ein weiterer Grund für die **Verteilungsprüfung** von Variablen liegt in der notwendigen Prüfung von Verteilungsvoraussetzungen (z.B. dem Erfordernis der Normalverteilung von Fehlerkomponenten) für bestimmte multivariate statistische Auswertungsmethoden (z.B. der „Regressionsanalyse").

Die besondere Bedeutung der **Normalverteilung** (oder „Gauß'schen Glockenkurve") und der (mit ihr verwandten) **Binomialverteilung** in der Statistik ist darauf zurückzuführen, dass beide als Idealformen zufallsbedingter Verteilungsprozesse angesehen werden. Wie bereits erwähnt, wird bei empirischen Variablen angenommen, dass sich ihre Werte aus einer wahren Komponente und einer zufälligen Fehlerkomponente zusammensetzen.

Um den „Schwerpunkt“ von mehreren Variablenwerten zu bestimmen, werden Maße der „Zentraltendenz“ („Lageorientierung“) herangezogen, wie etwa der **Mittelwert** (m) bzw. „Durchschnittswert“, nämlich die Summe (Σ) aller Werte (x) dividiert durch die Anzahl der Werte (n):

$$m = \frac{1}{n} \cdot \sum x$$

Ebenfalls als Maß der Zentraltendenz gebräuchlich ist der **Median** (jener Wert, von dem aus etwa 50 % aller Werte größer oder kleiner sind).

Eine zweite wichtige Kennzeichnung von Variablen sind statistische Kennwerte, die den „Streubereich“ („Dispersion“) der Ausprägungen von Variablen aufzeigen. Maße dafür sind etwa die **Streubreite** (Bereich vom maximalen Wert bis zum minimalen Wert), die **Varianz**

$$v = \frac{1}{n} \cdot \sum (x - m)^2$$

oder die **Standardabweichung**:

$$s = \sqrt{v}$$

Die Varianz (v) ist als Durchschnittswert für die Abweichungsquadrate aller Werte (x) vom Mittelwert (m) definiert. Die Standardabweichung (s) als Wurzel der Varianz bezeichnet jene Abweichungen vom Zentrum der Normalverteilung, innerhalb derer etwa 68 % aller Werte liegen.

Um nun Variablen (mit verschiedenen Mittelwerten und Streuungen) besser miteinander vergleichen und auf wechselseitigen Zusammenhang (*Korrelation*) überprüfen zu können, werden sie oft durch einen einfachen Rechenvorgang in sogenannte **Standardvariablen** mit jeweils einem Mittelwert von 0,0 und einer Streuung von 1,0 umgewandelt (*Standardisierung*). Dies wird erreicht, indem alle Werte einer Variablen um ihren Mittelwert reduziert und durch die Streuung dividiert werden („lineare Transformation“, s. auch Abb. 8.19):

$$z = \frac{x - m}{s}$$

Um das Ausmaß der linearen „Ko-Relation“ zwischen zwei Variablen abschätzen zu können, bedient man sich seit etwa hundert Jahren des Pearson’schen **Korrelationskoeffizienten** (r), der bei einem maximal positiven Zusammenhang zwischen den zwei Variablen den Wert +1,0 annimmt (wenn beide gleichsinnig zu- oder abnehmen). Bei Fehlen einer linearen Beziehung wird er 0,0, bei einem maximal gegensätzlichen Zusammenhang dagegen erhält man −1,0. Dieser in den Sozialwissenschaften häufig verwendete statistische Kennwert wird auch als „Produkt-Moment-Korrelation“ bezeichnet und lässt sich für zwei Standardvariablen sehr einfach, nämlich als mittleres Produkt der z-Werte, bestimmen (gebräuchliche Formeln zur Berechnung einer **Korrelation** findet man in statistischen Lehrbüchern oder im Internet):

$$r = \frac{1}{n} \cdot \sum (z_y \cdot z_x)$$

(Σ = Summenzeichen, n = Anzahl der Fälle, z_y und z_x = Standardwerte der Variablen Y und X)

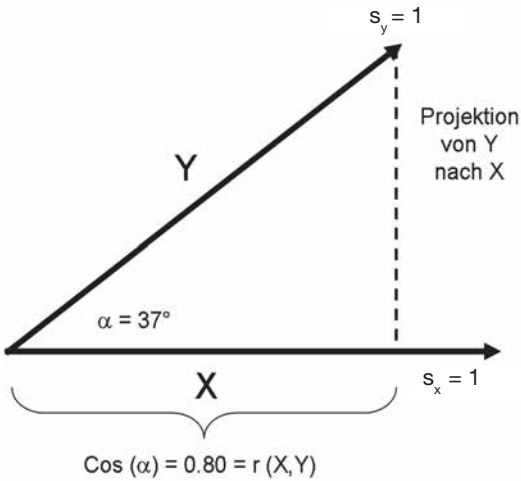
Eine praxisrelevante Nutzanwendung dieser Statistik besteht in ihrer Vorhersagefunktion für die Ausprägungen einer Variablen (Y), wenn die Werte einer anderen, mit ihr (linear) korrelierenden Variablen (X) bekannt sind:

$$z_y = r \cdot z_x$$

Da mittels dieser (linearen) Funktion die Werte einer Variablen auf jene einer anderen Variablen zurückgeführt werden können (Abb. 3.8), nennt man diese Bezugsherstellung **Regression** und das statistische Verfahren **Regressionsrechnung**. Mittels der Korrelation lässt sich somit der vermutete Einfluss einer Variablen auf eine andere Variable abschätzen. Das Ausmaß des statistischen Effektes einer Variablen auf eine oder mehrere andere Variablen wird als **Effektstärke** bezeichnet. So etwa kann man aus dem Quadrat des Korrelationswertes die Stärke des vermuteten Einflusses einer Variablen auf eine andere abschätzen (z.B. r = 0,5, r^2 = 0,25, d.h. 25 % Prädiktion), wenn es sich um eine bidirektionale (notwendige und hinreichende) Beziehung handelt (s. Abb. 3.3), was bei psychologischen Effekten eher selten der Fall ist. Neben dem Korrelationskoeffizienten existieren noch weitere Kennwerte für Effektstärken (s. Bortz & Döring, 1995; Westermann, 2000).

lat. regredere: zurückgehen, zurückführen

Abb 3.8

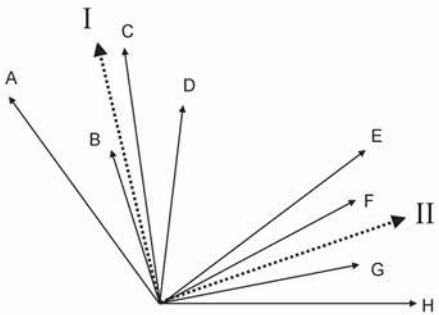


Wenn zwei Variablen (X, Y) durch zwei Einheitsvektoren symbolisiert werden (d.h. als Standardvariablen mit Standardabweichung von s = 1), und die Variablen miteinander im Ausmaß von r = 0,80 korrelieren, dann kann diese Relation durch einen Winkel von 37° zwischen den Vektoren $\vec{Y}$ und $\vec{X}$ im Variablenraum dargestellt werden: r = 0,80 = Cos (37°). Der Wert r entspricht somit der Abbildung einer Variablen auf eine andere.

Eine Besonderheit der geometrischen Betrachtungsweise von Variablen besteht darin, dass das Ausmaß ihrer linearen Beziehung (Korrelation) durch den Winkel ihrer Vektordarstellungen im Variablen- bzw. Merkmalsraum dargestellt werden kann (Andres, 1996; Abb. 3.8). Da jeder Vektor eine variablenspezifische Eigenschaft symbolisiert und gleichgerichtete Bündelungen von Vektoren somit auf Eigenschaftsüberlappungen der entsprechenden Variablen hinweisen, können für solche Variablencluster gewissermaßen „Schwerpunktvektoren" berechnet werden, die man **Faktoren** nennt und die als oberbegriffliche Beschreibungen der durch die Variablen symbolisierten Eigenschaften aufzufassen sind (Abb. 3.9).

Mittels solcher faktorieller Beschreibungen kann man nicht nur komplexe Variablensysteme auf ihre „Hauptkomponenten" reduzieren, sondern auch den korrelativen Zusammenhang zwischen verschiedenen Gruppen von Variablen (mit ähnlicher Eigenschaftsbedeutung) bestimmen. (Statistische Verfahren, die auf diesem Prinzip basieren, sind etwa die „Faktorenanalyse", die „Multivaria-

Abb 3.9



Das Prinzip der „Faktorenanalyse“: Wenn zwischen je zwei dieser acht Variablen der Korrelationskoeffizient berechnet wird und die Variablen in den entsprechenden Winkeln zueinander grafisch dargestellt werden, können Bündel davon durch sogenannte Faktoren (I, II) charakterisiert werden. Die vorliegenden acht Variablen lassen sich relativ gut in nur zwei Dimensionen darstellen, wobei die Länge der Variablenvektoren das Ausmaß ihrer Charakterisierbarkeit durch die beiden senkrecht zueinander stehenden Faktoren widerspiegelt. Im Beispiel könnten die vier Variablen A, B, C und D etwa die Eigenschaften schön, vielfältig, harmonisch und heiter von architektonischen Objekten symbolisieren und aufgrund ihrer vektoriellen Bündelung einen Faktor (I) beschreiben, den man ästhetischer Eindruck nennen könnte.

te Varianzanalyse“, die „Kanonische Korrelation“ oder die „Diskriminanzanalyse“.)

Eine Erweiterung dieser Verfahren ist die sogenannte „topologische Datenanalyse“ (Wasserman, 2018; Morris, 2015), bei der Daten an empirische Formen oder Strukturen angepasst werden (z.B. Protein-Strukturen, Kommunikationsnetze).

Inferenzstatistik – schließende und prüfende Statistik

3.6.2

Wie mehrfach erwähnt, müssen in der Psychologie Schlussfolgerungen über die allgemeine Gültigkeit von Gesetzen auf Basis von *Stichproben* gezogen werden. Dies geschieht zumeist unter Verwendung der **Wahrscheinlichkeitstheorie**, mittels derer man zu bestimmen versucht, ob die in den Daten festgestellten Variablenrelationen nur zufällig oder doch durch Einwirkung eines Gesetzes zustande gekommen sind.

Vereinfacht, aber sehr prägnant kann das Bestreben empirischer Sozialforschung anhand des mathematischen **Bayes-Theorems** illustriert werden:

$$p(H/D) = \frac{p(D/H) \cdot p(H)}{p(D)}$$

lat. a posteriori: von dem, was nachher kommt
lat. a priori: von vornherein, ohne Einbezug von Erfahrungen

In empirischen Wissenschaften geht es um die Einschätzung der Wahrscheinlichkeit p(H|D) für die Gültigkeit einer Hypothese (H) unter der Bedingung, dass hypothesenbestätigende (oder widerlegende) empirische Daten (D) berücksichtigt werden. Die „Aposteriori-Wahrscheinlichkeit" p(H|D) für eine Hypothese (d.h. nach Einbezug der Daten) nimmt zu, wenn die „Apriori-Wahrscheinlichkeit" für die Hypothese p(H) größer wird und/oder wenn die Wahrscheinlichkeit p(D|H) für das Auftreten hypothesenbestätigender Daten ebenfalls zunimmt. Sie nimmt hingegen ab, wenn die hypothesenrelevanten Daten auch unabhängig von der Hypothese häufiger auftreten, das heißt, wenn p(D) größer wird.

Die Plausibilität dieses Ansatzes kann am Beispiel einer medizinischen Diagnose über das Vorliegen einer Covid-19-Infektion illustriert werden: Die Annahme, dass eine Person an Covid-19 (C) erkrankt ist, wenn sie Fieber hat (p(C/F)), stimmt umso eher, (1) je größer p(C) ist, das heißt, je mehr Personen bereits an Covid-19 erkrankt sind (z.B. bei einer Epidemie), (2) je größer p(F|C), die Wahrscheinlichkeit von Fieber bei dieser Viruserkrankung, ist und (3) je kleiner p(F) ist, nämlich die Erwartung des Auftretens von Fieber im Allgemeinen (s. auch 8.5.3; Tschirk, 2019).

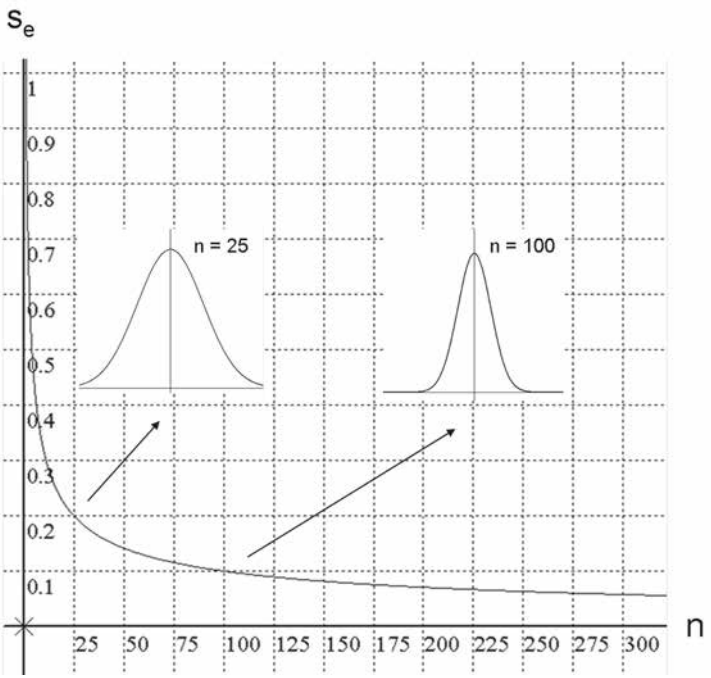
Merksatz

Die möglichst stabile Kennzeichnung von Personen oder Personengruppen hinsichtlich wichtiger Eigenschaften, Einstellungen oder Handlungsweisen („Punktschätzungen") ist eine zentrale sozialwissenschaftliche Zielsetzung.

Eine zentrale sozialwissenschaftliche Zielsetzung besteht in der möglichst stabilen Kennzeichnung von Personen oder Personengruppen hinsichtlich wichtiger Eigenschaften, Einstellungen oder Handlungsweisen („Punktschätzungen"). Da solche Kennwerte immer fehlerbehaftet sind, wird mittels statistischer Techniken ein **Vertrauensintervall** bzw. **Konfidenzintervall** für sie bestimmt, innerhalb dessen mit 95%iger (99%iger) Wahrscheinlichkeit der „wahre" Kennwert vermutet wird.

Es ist leicht einzusehen, dass der **Schätzfehler** für einen statistischen Kennwert mit zunehmender Größe der Stichprobe immer kleiner wird und schließlich gegen Null geht, wenn alle möglichen Fälle in die Berechnung einbezogen sind (Abb. 3.10).

Abb 3.10



Der Schätzfehler (s_e) für die Bestimmung des Mittelwertes einer Population von Fällen aufgrund einer Stichprobe ist eine Funktion der Stichprobenstreuung (s) und des Stichprobenumfanges (n): $s_e = s/\sqrt{n}$. Je mehr Fälle für eine Schätzung zur Verfügung stehen, desto genauer wird die Vorhersage. Wenn etwa geschätzt werden sollte, wie viel Zeit Arbeiterinnen und Arbeiter durchschnittlich für einen bestimmten Arbeitsgang in einem Produktionsprozess benötigen, dann wird die Schätzung des Mittelwertes anhand einer Stichprobe von 100 Personen eine nur halb so große Fehlerstreuung aufweisen (in Einheiten der Standardabweichung) wie jene auf Basis einer Stichprobe von 25 Personen.

Die mathematisch begründeten Methoden der **Inferenzstatistik** sollen also eine Einschätzung erlauben, ob überhaupt und in wel-

lat. inferre: hineintragen

chem Ausmaß statistische Resultate von Stichproben auf die jeweilige Population übertragbar sind.

Wenn die Wahrscheinlichkeit dafür, dass bestimmte Variablenrelationen zufällig zu erklären sind, einen vereinbarten Wert unterschreitet (z.B. p = 0,05, p = 0,01 oder p = 0,001), dann spricht man von **statistischer Signifikanz** des Ergebnisses. Bortz und Döring (1995, 27) definieren statistische Signifikanz als ein „per Konvention festgelegtes Entscheidungskriterium für die vorläufige Annahme von statistischen Populationshypothesen". Wenn also ein statistisches Ergebnis nur mehr zu 5 % (oder weniger) durch Zufallsprozesse erklärt werden kann, wird es als statistisch signifikant angesehen („überzufällig" oder „unterzufällig"). Die restliche, für eine Zufallserklärung verbleibende Unsicherheit von 5 % (oder weniger) nennt man **Irrtumswahrscheinlichkeit** („Fehler 1. Art", „Alpha-Fehler"), die dazugehörige den Zufallsprozess charakterisierende Annahme (über die Datenverteilung) heißt **Nullhypothese**.

Merksatz

Inferenzstatistische Verfahren zielen darauf ab, den Grad der Allgemeingültigkeit von Gesetzmäßigkeiten zu prüfen, die auf Basis von Stichproben gewonnen werden.

Da die praktische Bedeutsamkeit eines signifikanten Ergebnisses aber auch von dessen *Effektstärke* abhängt, müssen abgesehen von der Nullhypothese auch **Alternativhypothesen** statistisch getestet werden. Das Ausmaß, in dem die Datenverteilungen mit den Vorhersagen einer Alternativhypothese übereinstimmen, wird als **Teststärke** (engl. power) bezeichnet. Um diese berechnen zu können, ist es nötig, die jeweilige Alternativhypothese zu spezifizieren, indem man die erwartete *Effektgröße* präzisiert, d.h. schätzt, wie stark die jeweilige unabhängige Variable auf die abhängige Variable einwirken dürfte. Der Vorteil einer solchen Vorgangsweise besteht vor allem darin, dass man nicht nur vage auf „Über- oder Unterzufälligkeit" von statistischen Ergebnissen schließt, sondern sogar die Wahrscheinlichkeit bestimmen kann, mit der die Daten für die Alternativhypothese sprechen.

Forschungsmethoden der Psychologie 3.7

Laborexperiment 3.7.1

Mittels eines Experiments ist es möglich, hypothetische Wirkfaktoren gezielt zu manipulieren, um ihre Auswirkungen unter verschiedenen Bedingungen zu analysieren. Experimente werden bevorzugt zur Prüfung von Kausalhypothesen eingesetzt (Stapf, 1987). Im Experiment wird eine künstliche Realität konstruiert, um die vermuteten Einflussfaktoren in ihrer Wirksamkeit unter Abschirmung von möglichen Störeinflüssen zu untersuchen.

Häufig wird in psychologischen Experimenten der (den) **Experimentalgruppe(n)** (Versuchsbedingungen) eine **Kontrollgruppe** (Kontrollbedingung) gegenübergestellt. Den Fällen der Experimentalgruppen sind solche Ausprägungen der unabhängigen Variablen (Ursachenvariablen) zugeordnet, von denen ein Effekt auf die abhängigen Variablen (Wirkungsvariablen) erwartet wird, während den Fällen der Kontrollgruppe Ausprägungen der unabhängigen Variablen zugeteilt sind, denen kein systematischer Effekt zugeschrieben wird. Diese Gruppe dient somit nur dazu, Veränderungen zu erfassen, die entweder auf natürliche Weise auftreten (Zeiteffekte, Gewöhnungsprozesse etc.) oder durch die experimentellen Umstände selbst zustande kommen, nämlich durch die künstliche Situation oder den Eindruck, beobachtet zu werden.

Die künstliche Realität des Experiments ist einerseits ein Vorteil, weil durch die Beseitigung von Störeinflüssen der Zusammenhang zwischen unabhängigen und abhängigen Variablen klarer erkannt werden kann (hohe „interne *Validität*"), andererseits aber auch ein Nachteil, weil die Ergebnisse nur mit Vorsicht auf den Alltag übertragbar sind (geringe „externe" bzw. „ökologische *Validität*").

Ein wesentliches Merkmal psychologischer Experimente ist die **Randomisierung**. Durch die Randomisierung sollen sich Störeffekte ausmitteln, die eventuell durch unausgewogene Stichproben zustande kommen. In der zuvor erwähnten Studie über die Wirkung des Alkohols auf das Fahrverhalten (Box 3.1) würden zum Beispiel die sich meldenden Versuchspersonen per Zufall den Gruppen mit unterschiedlicher Alkoholaufnahme zugewiesen werden.

Randomisierung meint die zufällige Zuordnung von Personen (oder Gruppen) zu den jeweiligen Ausprägungen der unabhängigen Variablen.

Die **Störeffekte** in psychologischen Experimenten haben im Wesentlichen drei verschiedene Quellen (Gniech, 1976):

1. **Versuchssituation**: Der sogenannte „Aufforderungscharakter" eines experimentellen Umfelds, nämlich die Art der Information über den Zweck der Untersuchung, die Rahmenbedingungen, die Art der Instruktion, die gestellten Fragen und Ähnliches hinterlassen bei den Versuchspersonen Eindrücke, die ihr experimentell induziertes Verhalten beeinflussen können.
2. **Versuchspersonen**: Eine unüberlegte, nicht randomisierte Auswahl der Stichprobe kann Verfälschungen in den Ergebnissen bewirken. Ein nicht zu unterschätzendes Problem bei der Interpretation von Untersuchungsergebnissen ist zum Beispiel die oft notwendige Beschränkung der Teilnahme auf Freiwilligkeit und das Ausscheiden von Teilnehmenden aus dem Experiment („drop out"), wodurch natürlich die erwünschte Zufallsauswahl einer Stichprobe beeinträchtigt ist. Personen, die sich freiwillig für ein Experiment melden, sind im Allgemeinen besser gebildet, haben einen höheren gesellschaftlichen Status, sind stärker sozial orientiert und haben ein stärkeres Bedürfnis nach Anerkennung (Rosenthal & Rosnow, 1975; zit. nach Gniech, 1976). Natürlich wirken sich auch Einstellungsunterschiede der Teilnehmerinnen und Teilnehmer gegenüber der Untersuchung aus, je nachdem, ob es sich um kooperierende, sabotierende oder neutrale Versuchspersonen handelt.
3. **Versuchsleitung**: Vonseiten der Forscherinnen und Forscher sollten **beobachterabhängige Urteilsverzerrungen** (engl.: observer bias) beachtet werden, die durch persönliche Motive und Erwartungen entstehen. Besonders störend sind unbewusste Einflussnahmen (z.B. über nonverbale Kommunikation) im Sinne eigener theoretischer Vorstellungen („Erwartungseffekte", „Rückschaufehler", *Selbsterfüllende Prophezeiung*). In Experimenten mit Volksschulkindern (Box 3.4) konnte etwa nachgewiesen werden, dass Lehrpersonen gegenüber fremden Kindern, die ihnen aufgrund von Testresultaten als angeblich begabt ausgewiesen wurden (als „Spätentwickler"), sich sympathischer, förderlicher und entgegenkommender verhielten, sodass sie mit ihrem Verhalten de facto dazu beitrugen, die Fähigkeiten der Kinder zu steigern (Rosenthal & Jacobson, 1968).

Dass Erwartungshaltungen, zum Beispiel hinsichtlich der Wirksamkeit eines Medikamentes, beachtliche Auswirkungen haben können, ist seit Langem aus der Medizin unter der Bezeichnung **Placebo-Effekt** bekannt. Darunter versteht man die positive Wirkung

lat. placebo: „Ich werde gefallen."

Selbsterfüllende Prophezeiung | Box 3.4

Die Selbsterfüllende Prophezeiung wird auch Pygmalion-Effekt genannt, nach dem Bildhauer der griechischen Mythologie, der die Statue einer perfekten Frau schuf („Galatea“) und sie durch seinen festen Glauben und seine Sehnsucht nach ihr zum Leben erweckte (Göttin Aphrodite soll allerdings mitgeholfen haben).

In einem Experiment der Sechzigerjahre waren 18 Klassen einer Volksschule einbezogen. Bei allen Schülerinnen und Schülern wurde ein nonverbaler Intelligenztest durchgeführt, den man als Indikator für eine zu erwartende intellektuelle Entwicklung der Kinder in den nächsten acht Monaten ausgab. Aus allen Klassen wurden 20 % der Kinder zufällig (!) ausgewählt, die den Lehrpersonen mit dem Hinweis genannt wurden, dass von diesen Kindern aufgrund des durchgeführten Tests in der nächsten Zeit ein intellektueller Fortschritt zu erwarten sei. Nach acht Monaten zeigten die mit dem positiven Vorurteil bedachten Kinder im Intelligenztest deutliche Verbesserungen! Der gleiche Effekt konnte auch in Tierexperimenten nachgewiesen werden (Rosenthal, 2002).

auf Befinden oder Gesundheit ausgehend von medizinisch unwirksamen Substanzen – sogenannten „Placebos“ (z.B. Milchzucker, Stärke, Salzlösungen) –, allein durch Herstellung einer Erwartung von Wirksamkeit.

Das Gegenteil vom Placebo-Effekt ist der **Nocebo-Effekt**, nämlich die durch Erwartung hervorgerufene negative Auswirkung eines eigentlich unwirksamen Medikaments oder einer Behandlung.

Um die genannten Artefakte in Experimenten zu reduzieren, werden **Doppel-blind-Verfahren** eingesetzt, bei denen weder die Versuchspersonen noch die unmittelbar das Experiment betreuenden Forscherinnen und Forscher über die Art der experimentell gesetzten Einwirkungen Bescheid wissen dürfen. Da natürlich auch in Blindstudien die Probandinnen und Probanden über die Intention einer Studie Vermutungen entwickeln, müssen in der psychologischen Forschung manchmal auch **Täuschungen** eingesetzt werden. Selbstverständlich sind diese nach Ende des Experiments aufzuklären.

Experimente sowie andere Forschungsdesigns können sowohl als **Querschnittuntersuchung** (engl.: cross sectional study) als auch als **Längsschnittuntersuchung** (engl.: longitudinal study) durchgeführt werden. Bei der häufig eingesetzten Querschnittstudie werden an einzelnen Fällen (Personen, Gruppen, Situationen etc.) die interessierenden Variablen nur einmalig erhoben, sodass strukturelle Gesetze von Variablen analysiert werden können, während bei einer Längsschnittstudie zwei oder mehr Datenerhebungen zu den gleichen Variablen stattfinden und somit auch deren zeitliche Dynamik erfassbar ist. Ein großer Vorteil der Längsschnittmethode liegt auch darin, dass intraindividuelle Veränderungen beobachtet werden können und Verfälschungen durch unausgewogene Stichproben, wie sie bei der interindividuellen Querschnittmethode vorkommen, reduziert sind (Daumenlang, 1987). Nachteilig ist hingegen über einen längeren Zeitraum der Schwund an Versuchspersonen und die Problematik der mehrmaligen Anwendung gleichartiger Testverfahren (Gefahr von „Serialeffekten“).

Merksatz

In einem Experiment werden unter abgeschirmten Bedingungen die Effekte (abhängige Variablen) systematisch variierter Wirkfaktoren (unabhängige Variablen) registriert, wobei durch eine zufällige Zuteilung der Fälle zu den Bedingungen der Wirkfaktoren etwaige Verfälschungen der Ergebnisse minimiert werden sollen.

3.7.2 | Quasiexperiment

Diese Untersuchungstechnik gleicht vom Design her dem Experiment, nur verzichtet man auf eine zufällige Zuordnung der Versuchspersonen zu den Versuchs- bzw. Kontrollbedingungen und nimmt das Risiko von Stichprobeneffekten in Kauf. In manchen Forschungsbereichen ist eine zufällige Zuteilung zu den verschiedenen Bedingungen entweder nicht realisierbar oder ethisch nicht zu rechtfertigen; so etwa die zufällige Zuordnung von Schülerinnen und Schülern zu Schultypen, von Mitarbeiterinnen und Mitarbeitern zu Betrieben oder von Patientinnen und Patienten zu Behandlungsverfahren. Um die aus dem Verzicht einer Randomisierung resultierenden **Artefakte** zu kompensieren, werden in solchen

Merksatz

Ein Quasiexperiment gleicht vom Aufbau her einem Experiment – mit dem Unterschied, dass die Fallzuordnung zu den Bedingungen nicht zufällig erfolgt.

Artefakt: In Psychologie und Nachrichtentechnik steht dieser Ausdruck für verfälschte Ergebnisse.

Untersuchungen einerseits größere Probandengruppen angestrebt und andererseits zusätzliche Personenmerkmale erhoben, denen ein direkter oder indirekter Einfluss auf die abhängigen Variablen zugeschrieben werden kann. Zu diesen gehören die **soziodemografischen Merkmale** (Geschlecht, Alter, Schulbildung, Beruf ...), aber auch andere individuelle Charakteristika, die aufgrund ihrer Ungleichverteilung in den Bedingungen der unabhängigen Variablen zu systematischen Verfälschungen von Ergebnissen führen könnten. Mittels statistischer Korrekturverfahren lassen sich einige solcher Verfälschungen kompensieren bzw. aus den Ergebnissen herausrechnen („auspartialisieren").

Feldforschung

3.7.3

Im Gegensatz zum Experiment versucht man in der Feldforschung, Phänomene unter möglichst natürlichen Bedingungen zu beobachten und zu erklären. Dem Vorteil der Natürlichkeit steht hier der Nachteil gegenüber, dass Störvariablen weniger gut kontrolliert werden können. Da Forschungsphänomene „im Feld" wesentlich komplexer in Erscheinung treten als im Labor, kommt bei der Feldforschung der Entwicklung von genauen und effizienten Beschreibungsmethoden sowie der Ausarbeitung von Verhaltensregeln zur optimalen Datengewinnung besondere Bedeutung zu (s. Flick et al., 1995).

Merksatz

Methoden der Feldforschung bezwecken eine Untersuchung von Phänomenen unter natürlichen Rahmenbedingungen bzw. unter minimierter versuchsbedingter Beeinflussung.

Sogenannte **Fallstudien** („single case studies") sind häufig erste Erfahrungsquellen und als solche nur Anregungen für weitere Forschungstätigkeiten. Obwohl Forschungsphänomene durch Fallstudien hervorragend konkretisiert und plastisch vorstellbar gemacht werden können, mangelt es ihren Ergebnissen logischerweise an Verallgemeinerbarkeit.

Einen Katalog möglicher Verhaltensweisen in natürlichen Umweltbedingungen nennt man in der Verhaltensforschung **Ethogramm**, innerhalb dessen ein „behavior mapping" charakterisiert, wer was wo tut (Hellbrück & Fischer, 1999).

Non-reaktive Verfahren bezwecken eine Analyse psychologischer Problemstellungen, ohne dass die untersuchten Personen bemer-

ken, dass sie untersucht werden, was insbesondere bei **Inhaltsanalysen** von schriftlichen Dokumenten (z.B. Tagebüchern, Archiven), bei Auszählungen von Unfall-, Krankheits- und Fehlzeitstatistiken, Verkaufsstatistiken oder Abnützungen von Bodenbelägen, Pfaden oder Gebrauchsgegenständen („Spurenanalyse“) gut gelingt.

In der Feldforschung werden häufig, aber nicht ausschließlich, qualitative Methoden verwendet, weil diese flexibler auf die Eigenarten von Personen oder von Situationen anzupassen sind.

3.7.4 | Test und Rating

Merksatz

Eine Standardisierung besteht aus Maßnahmen, die eine Vergleichbarkeit von verschiedenen Personen, Objekten, Situationen oder von Variablenwerten ermöglichen.

Ein Test ist ein wissenschaftlich normiertes und standardisiertes Prüfverfahren, welches stabile Eigenschaften eines komplexen Systems (Person, Gegenstand, Organisation) ermitteln soll. Unter **Standardisierung** versteht man Maßnahmen, aufgrund derer Situationen, Aktionen oder Objekte unter Bezugnahme auf Normen oder Regeln miteinander verglichen werden können. So etwa müssen in Experimenten Instruktionen und Rahmenbedingungen der Durchführung für alle Versuchspersonen standardisiert, d.h. als maximal ähnlich aufgefasst werden; Gleiches gilt für die Diagnostik, wo Tests verschiedenen Kandidatinnen oder Kandidaten vorgegeben werden. Bei standardisierten Fragebögen müssen die Fragen immer den gleichen Wortlaut haben, auch die möglichen Antworten sind fix vorgegeben (z.B. in Form von Antwortalternativen). Bei statistischen Auswertungen gelten Variablen dann als standardisiert, wenn ihre Werte als Differenz zu ihrem *Mittelwert* und in Einheiten ihrer *Streuung* dargestellt werden (s. 3.6.1), wodurch auch inhaltlich sehr verschiedenartige Variablen miteinander in Relation gesetzt werden können. Bei Leistungs-, Intelligenz- oder Persönlichkeitstests bedeutet eine Standardisierung, dass die Ergebnisse der Probandinnen und Probanden auf die Verteilungen von sogenannten **Referenz-** oder **Normstichproben** bezogen sind.

Merksatz

Eine Skala soll Ausprägungen einer spezifischen Eigenschaft eines empirischen Sachverhaltes exakt (anhand von Zahlen) charakterisieren.

Da die in einem Test zu erfassenden *Konstrukte* aus Teilaspekten bzw. verschiedenen Inhaltskomponenten bestehen, setzen sich Tests aus entsprechend vielen **Subtests** bzw. **Skalen** zusammen, die jeweils ein homogenes Merkmal feststellen oder „messen" sollen. Eine Skala ordnet somit empirischen Objekten (z.B. Personen) Zahlen zu, ähnlich wie dies bei der Längenmessung physikalischer Objekte anhand einer Meterskala geschieht (Niederée & Narens, 1996). Die Prüfung, ob zur Vermessung eines empirischen Objekts eine quantitative Skala akzeptiert werden kann, erfolgt auf Basis mathematisch-statistischer **Messtheorien** (s. auch 3.6.1, *Skalenqualität*).

engl. scale: Maßstab, Anzeige, Skala; ital. scala: Maßstab, Treppe, Leiter, Skala

Subtests oder Skalen bestehen selbst meist wieder aus zwei, drei oder mehr Elementen, den **Items**. Je nach inhaltlicher Orientierung des Tests oder der Skala können sich die Items aus Leistungsaufgaben, Fragen mit Antwortalternativen oder aus *Skalierungen*, nämlich Einschätzungen von Merkmalen anhand von Zahlen, zusammensetzen. Die Art der Reaktion einer Person auf ein Item wird über (Zahlen-)Symbole kodiert, welche unter Verwendung mathematisch-statistischer Modelle zu quantitativen Werten (z.B. Mittelwert über die einzelnen Items) für die einzelnen Skalen verrechnet werden. Je mehr Items für eine Skala zur Verfügung stehen, d.h., je mehr unabhängige elementare „Messinstrumente" für eine Eigenschaft vorliegen, desto größer ist im Allgemeinen die Zuverlässigkeit der entsprechenden Skala.

engl. item: Datenelement, Einheit, Einzelheit, Element, Punkt, Nummer

Eine Aufzählung nach Bühner (2010) soll illustrieren, in welch verschiedenen Bereichen psychologische Tests eingesetzt werden: psychische und psychosomatische Störungen, Befindlichkeitsstörungen, Therapie- und Heilungsverlauf, Familien-, Ehe- und Erziehungsberatung, Berufsberatung und Personalauslese, Verkehrseignung (TÜV), Strafvollzug (Haftentlassung), Entwicklungsstörungen, Schulreife, Leistungsstörungen, Hochschuleignung, Produktbeurteilung, Einstellungs- und Motivationsmessung (Arbeitszufriedenheit, Leistungsmotivation) usw. „Die Auswahl und Interpretation von Test- und Fragebogenergebnissen zählen zu den Routineaufgaben in der späteren Berufspraxis" von Psychologinnen und Psychologen (Bühner, 2010, 11).

Die Genauigkeit, die Aussagekraft und der Vorhersageerfolg von Testergebnissen hängen von sogenannten **Gütekriterien** der Tests ab. Allgemeine und unverzichtbare Gütekriterien von Datenerhebungsinstrumenten sind Objektivität, Reliabilität sowie Validität.

Die **Objektivität** eines Tests kennzeichnet die Unabhängigkeit seines Ergebnisses von der Person, die den Test durchführt. Sie ist besonders hoch, wenn verschiedene Testanwender zu gleichen Testergebnissen kommen. Dafür ist es allerdings nötig, dass die Anwenderinnen und Anwender fundierte testpsychologische Grundkenntnisse und Fertigkeiten besitzen (s. DIN-Norm 33430 für „Berufsbezogene Eignungsdiagnostik", Hornke & Winterfeld, 2004; Bühner, 2010). Objektivitätsmängel können sowohl durch fehlende Sorgfalt bei der Testdurchführung als auch durch Unterschiede bei der Auswertung oder Interpretation entstehen.

Reliabilität bedeutet Zuverlässigkeit und Genauigkeit eines Tests und ist gegeben, wenn bei wiederholter Anwendung des Tests bei gleichen Probanden auch weitgehend gleiche Ergebnisse zustande kommen. Hinweise auf die Zuverlässigkeit von Tests bekommt man, indem man (1) einen Test (falls möglich) wiederholt vorgibt und dessen Ergebnisse auf Übereinstimmung prüft („Retest-Methode"), oder indem man (2) sogenannte Paralleltests, nämlich Tests mit gleicher Aussagekraft, entwickelt und deren Übereinstimmung bei ein und derselben Personengruppe kontrolliert („Paralleltest-Methode), oder indem man (3) die Teile eines Tests auf Homogenität, d.h. auf inhaltliche Ähnlichkeit prüft („Konsistenzmethode").

Die **Validität** (Gültigkeit), das wichtigste Gütekriterium eines Tests, gibt an, wie gut er in der Lage ist, das zu messen, was er zu messen vorgibt (z.B. Intelligenz, Motivation, Persönlichkeitsmerkmale). „Inhaltliche Validität" oder „Augenscheinvalidität" besitzt ein Test dann, wenn es aufgrund der Art der Testung (Fragen, Leistungen usw.) offensichtlich ist, welcher Aspekt sich im Testergebnis hauptsächlich niederschlägt (z.B. Additionstest für Rechenfertigkeit, Bildermerktest für Vorstellungsfähigkeit). Die empirische Validitätsprüfung eines Tests geschieht hauptsächlich durch Berechnung des statistischen Zusammenhanges (*Korrelation*) seiner Werte mit einem plausiblen Kennwert („Kriteriumsvalidität") oder mit einem anderen Test, der den gleichen Aspekt zu messen vorgibt („Konstruktvalidität"). Beispielsweise könnte bei Schülerinnen und Schülern für einen Test über rechnerische Intelligenz die Mathematiknote als Validitätskrite-

Merksatz

Ein Test ist ein wissenschaftlich begründetes, normiertes und bestimmten Gütekriterien unterworfenes Verfahren mit dem Ziel einer quantitativen Erfassung von Merkmalen.

rium oder ein ebenfalls auf Rechenleistungen bezogener anderer Test als Validitätskonstrukt herangezogen werden.

Zwischen den genannten drei Gütekriterien besteht allerdings eine Implikationsbeziehung: Wenn ein Test nicht objektiv ist, kann er nicht reliabel sein, und wenn er nicht reliabel ist, ist er nicht valide. Wenn nämlich bereits die Datenerhebung stark fehlerbehaftet ist, können bei wiederholten Messungen keine gleichen Resultate auftreten, und wenn Letzteres nicht gesichert ist, kann auch die zu messende empirische Eigenschaft nicht befriedigend von anderen Eigenschaften unterschieden werden.

Insbesondere bei der Konstruktion von Tests werden neben Objektivität, Reliabilität und Validität noch weitere, ebenfalls wichtige Gütekriterien überprüft (s. Kubinger, 2003): **Skalierung** (quantitative Interpretierbarkeit der Testwerte), **Normierung** (Vergleichsmöglichkeit mit Bevölkerungsgruppen), **Fairness** (Chancengleichheit für alle Bevölkerungsgruppen), **Ökonomie** (Minimum an Ressourcenverbrauch), **Zumutbarkeit** (Minimum an zeitlicher, emotionaler und psychischer Belastung der Probandinnen und Probanden) und **Unverfälschbarkeit** (geringe Möglichkeit zur willkürlichen Beeinflussung der Testergebnisse durch die Testpersonen).

Tab 3.1

	sehr	ziemlich	eher	weder noch	eher	ziemlich	sehr	
SCHÖN	3	2	1	0	1	2	3	HÄSSLICH
RUHIG	3	2	1	0	1	2	3	LAUT
FREMD	3	2	1	0	1	2	3	VERTRAUT
TEUER	3	2	1	0	1	2	3	BILLIG
DUNKEL	3	2	1	0	1	2	3	HELL
GUT	3	2	1	0	1	2	3	SCHLECHT
VIELFÄLTIG	3	2	1	0	1	2	3	MONOTON
GROSS	3	2	1	0	1	2	3	KLEIN

Das Polaritätsprofil ist eine in der Psychologie sehr verbreitete Methode zur Erfassung einstellungsbezogener oder gefühlsmäßiger Reaktionen auf Objekte, Personen oder Situationen. Dabei wird von den Versuchspersonen eine Reihe von Eigenschaften oder Eigenschaftspaaren (ca. 5–25) hinsichtlich ihres Zutreffens zahlenmäßig eingestuft.

In der Philosophie den Tests sehr ähnlich und ebenfalls sehr verbreitet sind **Ratingverfahren**, mittels derer Eigenschaften von Personen, Objekten oder Situationen (z.B. Wahrnehmungen, Ausdruckswirkungen oder Einstellungsintensitäten) anhand von Zahlenzuordnungen quantitativ eingestuft werden. Ein Beispiel dafür ist das **Polaritätsprofil** („Semantisches Differential"; Tab. 3.1). In anderen Ratings bzw. **Skalierungen** wird etwa der Grad an Zustimmung zu Meinungen in Prozentpunkten, die Bewertung von Objekten oder Aspekten in Schulnoten oder eine Präferenzentscheidung mittels Punktesystem angegeben.

3.7.5 | Beobachtung

Die **Selbst-** und **Fremdbeobachtung** zählt zu den ältesten Forschungsinstrumenten der Psychologie. Die wissenschaftliche Beobachtung unterscheidet sich von jener des Alltags durch ihre Theoriegeleitetheit und Systematik. „Unter Beobachtung versteht man das systematische Erfassen von wahrnehmbaren Verhaltensweisen, Handlungen oder Interaktionen einer Person oder Personengruppe zum Zeitpunkt ihres Auftretens" (Ebster & Stalzer, 2003, 221). Grundsätzlich sollte die Beobachtung als Mittel der Informationsgewinnung in allen Untersuchungen zumindest begleitend eingesetzt werden, und auch die beschriebenen *Gütekriterien* von Tests sollten eigentlich für alle Datengewinnungsverfahren in der Psychologie gelten. So sind auch Beobachtungen einer Objektivitätsprüfung zu unterziehen, indem die Übereinstimmung verschiedener, unabhängiger Beobachterinnen oder Beobachter festgestellt wird.

Merksatz

Wissenschaftliche Formen der Selbst- und Fremdbeobachtung sind theoriegeleitet, systematisiert und den allgemeinen Gütekriterien der Datenerhebung unterworfen.

Die **teilnehmende Beobachtung** ist ein Verfahren, bei dem die forschende Person selbst am untersuchten Geschehen teilnimmt und von den erforschten Personengruppen wie ihresgleichen behandelt werden möchte. Man erwartet sich dadurch sowohl eine lebensnähere Perspektive des beforschten Phänomens als auch tiefere Einblicke in die jeweilige Problematik. Bei **nichtteilnehmender Beobachtung** ist man als Forscher den Untersuchungspersonen gegenüber distanzierter eingestellt und an objektiven Ergebnissen interessiert. **Verdeckte Be-**

Lernen und Anpassung

6

Inhalt

6.1 Umwelt und Verhalten

6.2 Aktivierung und Lernen

6.3 Speicherstrukturen des Gehirns

6.4 Neuronale Netzwerkmodelle

6.5 Habituation

6.6 Prägungsartiges Lernen

6.7 Klassische Konditionierung – Signallernen

6.8 Instrumentelles Konditionieren – Erfolgslernen

6.9 Fertigkeiten – Motorisches Lernen

6.10 Kognitives Lernen – Kategorien, Begriffe und Schemata

6.11 Imitationslernen – Beobachtungslernen – Modelllernen

Lernen ist eine erfahrungsbedingte, dauerhafte, aber modifizierbare Anpassung von Wahrnehmungen, Vorstellungen, Denkprozessen, Gefühlen, Motivationen oder Verhaltensweisen an Lebensbedingungen. Diese Anpassung ist also nicht auf Ermüdung, Reifung oder andere nicht erfahrungsbedingte Prozesse zurückzuführen. Lernen ist zudem stets mit einer Funktionsveränderung neuronaler Strukturen verbunden.

Empirische Hinweise über lernbedingte Strukturveränderungen im Zentralnervensystem (ZNS) können anhand von Hirnverletzungen („Hirnläsionen“), Hirnreizungen („Hirnstimulationen“) und mittels Aktivitätsmessungen am Gehirn gewonnen werden (z.B.

Merksatz

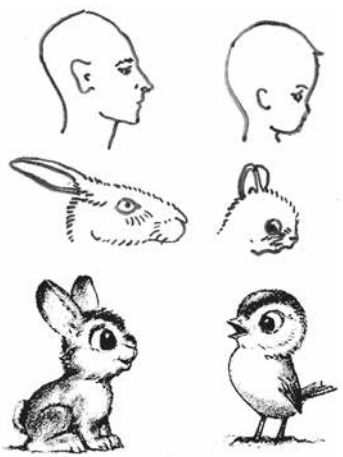
Lernen ist eine durch Erfahrung hervorgerufene dauerhafte, aber modifizierbare Anpassung von psychischen Strukturen und Reaktionen (Wahrnehmung, Denken, Verhalten, ...).

Elektroenzephalogramm/EEG, Positronen-Emmissions-Tomografie/PET, funktionale Magnet-Resonanz-Tomografie/fMRT). Die Ergebnisse der meisten Lernprozesse lassen sind allerdings nicht im neuronalen Substrat direkt nachweisen, sondern müssen indirekt an Veränderungen der Wahrnehmung, des Denkens oder des Verhaltens abgelesen werden.

6.1 | Umwelt und Verhalten

Die Evolutionstheorie (Darwin, 1859) zeigt auf, wie sich Organismen über Jahrmillionen durch Mutation und Selektion an wechselnde Umwelten angepasst haben (Buss, 2017). Die **Verhaltensforschung** bzw. **Ethologie** konnte nachweisen, dass diese Umweltadaptation nicht nur das Aussehen (Morphologie), sondern auch das regelhafte Verhalten der Lebewesen verändert hat. Insekten, Käfer und Spinnen verfügen bereits über Reflexe (Reiz-Reaktions-Koppelungen) und Reaktionsketten sowie über ein Instinktrepertoire (z.B. Brutpflege- und Fortpflanzungsinstinkte). Dadurch gelingen erstaunlich komplexe Anpassungen ihrer Lebensabläufe an die Um-

Abb 6.1 |



Erbkoordinationen (Angeborene Auslösermechanismen/AAM) sind reizabhängige Verhaltensprogramme, die nicht durch individuelle Erfahrungen erworben werden, sondern in einem bestimmten Alter des Lebewesens ausreifen. Als angeborener Auslöser darf das sogenannte Kindchenschema gelten (Lorenz, 1943), auf welches der Mensch (und andere Säuger) gefühlsmäßig reagiert („süß", „herzig", „niedlich") und mit „Pflegereaktionen" antwortet (z.B. streicheln, liebkosen). In Cartoons und bei Puppen werden die typischen Merkmale (relativ großer Kopf, große Augen, pummelige Gliedmaßen) zur Steigerung des Effektes oft übertrieben dargestellt.

welterfordernisse, insbesondere bezüglich Nahrungsaufnahme, Schlaf, Brutpflege und Sexualität. Lebewesen mit höher entwickeltem Zentralnervensystem, wie etwa Vögel und Säugetiere, steuern ihr Verhalten nicht mehr nur mittels solcher relativ starrer angeborener Mechanismen (wie zum Beispiel durch *Erbkoordinationen* und *Ritualisierungen*, Abb. 6.1 und 6.2), sondern auch durch erworbene Verhaltensanpassungen wie Prägung (s. 6.6, Abb. 6.14) und Lernen. Beim Menschen geschieht die Anpassung überwiegend durch Lernprozesse und Denken.

Kybernetik: Wissenschaft von der Struktur komplexer Systeme

Menschliches **Verhalten** kann aus kybernetischer Sicht als doppelt-hierarchische Organisation bezeichnet werden, bestehend aus einer bewusst kontrollierten Willkürmotorik („Top-down-Kontrolle“) und den Automatismen („Bottom-up-Kontrolle“), nämlich vor allem den Reflexen, den Gewohnheiten, den Triebreaktionen und der Stütz- und Gangmotorik (Pritzel et al., 2003). In jedem Verhaltensablauf sind beide Kontrollinstanzen integriert, das heißt, sowohl die willkürlich gesteuerten Aktivitäten werden der Körperhaltung, dem Körpertonus und den reflektorischen Aktionen angepasst als auch umgekehrt die unwillkürlich ablaufenden Regulationen den intendierten Handlungsmustern. Angeborene Reflexe (z.B. Kniesehnen- oder Speichelflussreflex, Husten-, Nies- oder Schluckreflex), erworbene Automatismen (z.B. erlernte Angstreaktionen, Sprechen, Schreiben, Spielen eines Musikinstruments) oder Willkürverhalten werden zwar auf unterschiedlich komplexen Regulationsebenen gesteuert, sind aber bei einem voll funktionstüchtigen Organismus gut aufeinander abgestimmt. Viele unwillkürliche Verhaltensweisen (Fertigkeiten, Gewohnheiten, Routinetätigkeiten) werden zunächst schrittweise und willkürlich kontrolliert aufgebaut, um dann später weitgehend ohne bewusste Kontrolle ausgeführt zu werden.

Merksatz

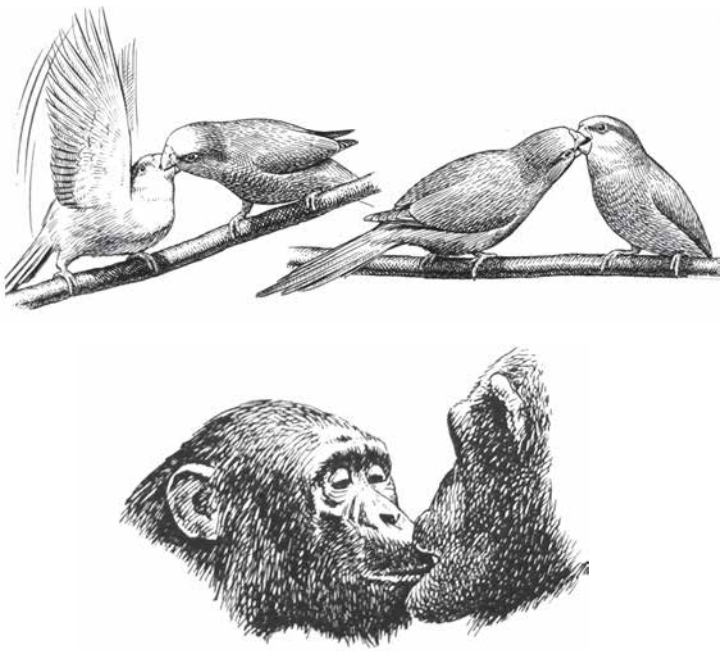
Menschliches Verhalten ist das komplexe Resultat einer unwillkürlichen Steuerung (durch Reflexe, Erbkoordinationen, Automatismen) und einer willkürlichen Steuerung (durch bewusste Handlungsentwürfe).

6.2 Aktivierung und Lernen

Im Laufe der Evolution des Lebens war die Fähigkeit zu lernen eine der wichtigsten phylogenetischen Umweltanpassungsleistungen der Lebewesen. Lernen als Instrument der Lebensbewältigung musste aber effizient sein, das heißt, der Nutzen eines Lernprozesses musste in einem vernünftigen Verhältnis zu seinen Kosten stehen, damit biologische Ressourcen (z.B. Speicherkapazität des Zentralnervensystems) nicht vergeudet wurden. Lernfähige Lebewesen benötigten daher ein verlässliches Kriterium zur Auswahl einprägenswerter Erlebnisse bzw. korrespondierender Umweltereignisse.

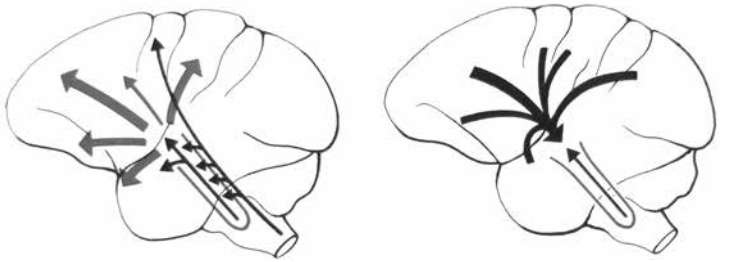
Als ein solcher biologischer Indikator für die Bedeutsamkeit eines Erlebnisses dürfte sich bereits früh die Erregungszunahme im Zentralnervensystem herausgebildet haben. Tatsächlich ist es so, dass immer dann, wenn organismusinterne oder -externe Reize

Abb 6.2



Ritualisierungen sind zweckentfremdete Erbkoordinationen, die in der phylogenetischen Entwicklung ursprünglich einem biologischen Zweck und danach der Sicherung des sozialen Zusammenhaltes innerhalb der Art dienten. So etwa entwickelte sich aus der Fütterung des Jungtieres das Schnäbeln oder Küssen erwachsener Tiere.

Abb 6.3

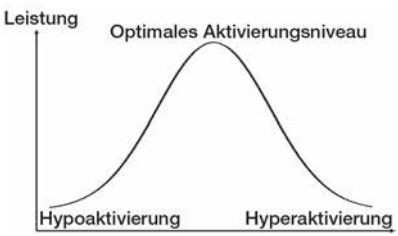


Aufsteigendes retikuläres Aktivierungssystem (ARAS): Stark schematisierte Darstellung der von der Formatio reticularis (netzartiges Gebiet im Hirnstamm) zum Kortex aufsteigenden und der zu ihr absteigenden Nervenbahnen, durch welche eine Steuerung des (tonischen) Aktivierungsniveaus erfolgt (am Beispiel des Affenhirns).

auftreten, die vermutlich oder tatsächlich Lebensrelevanz besitzen (z.B. Umweltveränderungen, Anblick von Nahrung, Schmerzen, Bedrohungen), sich das Niveau der allgemeinen **Aktivierung** (engl. arousal) im Zentralnervensystem erhöht und damit einhergehend die Intensität der neuronalen Informationsverarbeitung. Einen experimentellen Hinweis auf die lernfördernde Wirkung eines erhöhten neuronalen Aktivierungsniveaus liefert das „Brain-Trigger-Design" (Guttmann & Bauer, 1984), bei welchem durch die Vorgabe von Lernmaterialien in Phasen erhöhter neuronaler Aktivierung um 25 % bessere Merkleistungen nachweisbar waren.

Bei den meisten höher entwickelten Lebewesen befindet sich im Stammhirn ein entsprechendes **Aktivierungssystem** (engl. activating system; Abb. 6.3), welches sowohl auf Veränderungen in der Umwelt (Wahrnehmungssituation) als auch auf Änderungen im Organismus (Bedürfnislage) anspricht und eine Steigerung der Aktivierungslage im Zentralnervensystem auslöst („Orientierungsreak-

Abb 6.4



Das Yerkes-Dodson-Gesetz (Yerkes & Dodson, 1908) besagt, dass weder ein zu niedriges noch ein zu hohes, sondern ein mittleres Aktivierungsniveau die beste Voraussetzung für geistige Leistungen ist (s. auch Hebb, 1955). Daher werden im Allgemeinen schwere (anspannende) Aufgaben bei relativ niedriger Aktivierung und leichte Aufgaben bei höherer Aktivierung besser bewältigt.

tion“). Eine elektrische Stimulation dieses Aktivierungszentrums bewirkt eine Verschärfung der Wahrnehmung, Steigerung der Aufmerksamkeit, Beschleunigung der Informationsverarbeitung und eine Erhöhung der Reaktionsgeschwindigkeit, während eine Schädigung (z.B. als Folge von Encephalitis lethargica) oder Zerstörung (z.B. durch Verletzung) befristete bzw. dauerhafte Schlafzustände auslöst (Guttmann, 1972). Eine besondere Rolle bei der Aktivierung von Lernprozessen dürfte den Amygdalae zukommen, deren elektrische oder chemische Aktivierung in Experimenten die Einprägung von (emotionalen) Inhalten wesentlich verbessern kann (McGaugh, 2015; Roessler & McGaugh, 2019).

Encephalitis lethargica: „Europäische Schlafkrankheit“, Form der Gehirnentzündung

Merksatz

Überraschende und vital bedeutsame Ereignisse lösen Aktivierungsveränderungen im Zentralnervensystem aus und erhöhen damit die Wahrscheinlichkeit ihrer Speicherung.

Der Aktivierungszustand des Nervensystems verändert sich im Laufe des Wachzustandes sowohl tonisch (im Grundpegel) als auch phasisch (in Fluktuationen). Eine sehr hohe, aber auch eine sehr niedrige tonische Aktivierung (Hyperaktivierung – Hypoaktivierung) ist für kognitive Leistungen kontraproduktiv (Abb. 6.4). Optimal ist ein mittleres Aktivierungsniveau, im Bereich dessen vor allem das Ausmaß des Aktivierungsabfalls (Erleichterung, Entspannung, Befriedigung) darüber entscheidet, ob und in welchem Ausmaß die zuvor erlebten Eindrücke und Handlungen eingespeichert werden (Abb. 6.5, Abb. 6.6).

In einigen praxisnahen Untersuchungen hat Guttmann (1986) herausfinden können, dass ein Leistungsabfall (z.B. im Sport) zwischen Training und Wettbewerbssituation bei sogenannten **Trainingsweltmeistern** häufig auf übermäßige Aktivierung zurückführbar ist. In welchem Ausmaß bei Aktivierung Leistungssteigerungen oder Leistungshemmungen auftreten, hängt auch von der Persönlichkeit (Extraversion) und von der Aufgabenart ab (Storbeck & Clore, 2008).

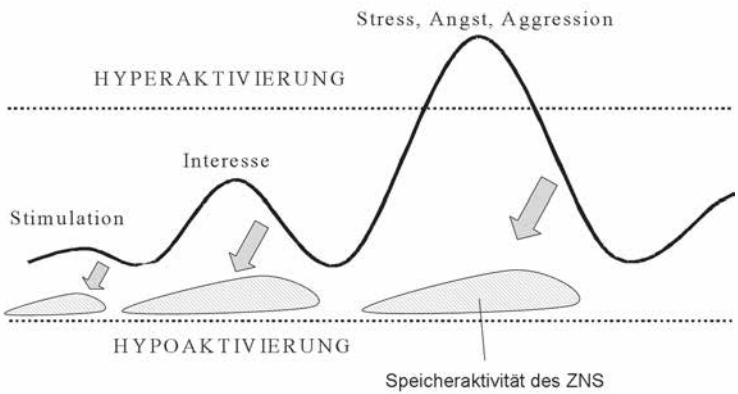
Das Aktivierungssystem stellt allerdings nur einen groben Regulationsmechanismus zur biologischen Bewertung von Lebensumständen dar, sodass sich in der Phylogenese komplexer Lebensformen (Säugetiere) bald auch ein differenzierteres zentralnervöses **Bewertungssystem**, nämlich das *Limbische System* (Amygdala, Abb. 9.5), herausbildete. Dieses nimmt laufend einen Vergleich zwischen Ist- und Sollwerten im biologischen und psychischen Bereich vor und

stellt fest, ob die gegebene Situation grundsätzlich eher als günstig oder als ungünstig einzuschätzen ist (Critchley & Garfinkel, 2018). Führt dieser Vergleich zu einem positiven Ergebnis, dann manifestiert sich dies subjektiv in einem **positiven Gefühl** (Zufriedenheit, Freude, Glück, ...), verbunden mit der Tendenz, den vorhandenen Zustand aufrechtzuerhalten und die gerade ausgeübte Tätigkeit fortzusetzen oder in Zukunft zu wiederholen (z.B. Essen, wenn etwas schmeckt). Weichen jedoch die Ist- von den Sollwerten zu stark ab, dann kommt es zu einem **negativen Gefühl**, wie etwa Unruhe, Angst oder Aggression, verbunden mit der Tendenz, den vorhandenen Zustand zu verändern und in Zukunft zu vermeiden. In die Bewertung der Situation fließt auch die Wahrnehmung des eigenen Aktivierungsniveaus und der körperlichen Empfindungen mit ein (Critchley & Harrison, 2013),

Merksatz

Sowohl ein zu hoher als auch ein zu niedriger zentralnervöser Aktivierungsgrad sind für kognitive Leistungen (Wahrnehmen, Einprägen, Erinnern, Problemlösen, Urteilen, ...) nachteilig. Bewusstseinsinhalte haben eine umso größere Chance auf Einprägung, je deutlicher sie zunächst von einem zentralnervösen Aktivierungsanstieg und darauffolgend von einem Aktivierungsabfall begleitet sind.

Abb 6.5

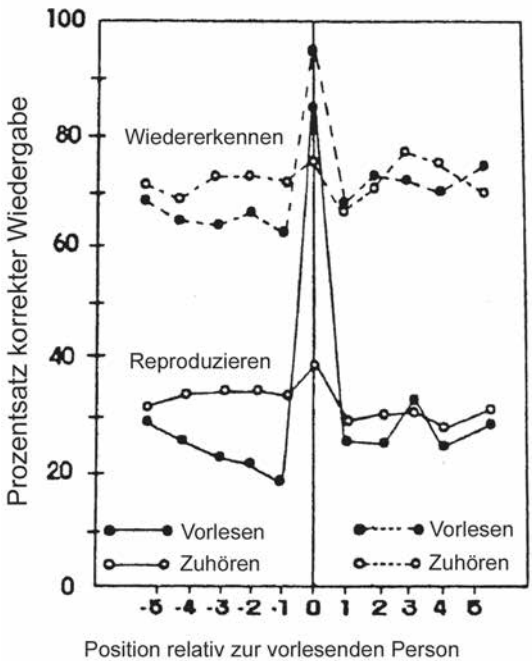


Bei höher entwickelten Lebewesen führen überraschende oder bedeutsame Ereignisse zu einem Anstieg des Aktivierungspegels im Zentralnervensystem. Wenn sich danach die Situation als ungefährlich herausstellt, sie in ihrem Bedeutungsgehalt durchschaut wird oder wenn sie auf beliebige Weise bewältigt wird, sinkt das Aktivierungsniveau wieder. Die Zunahme zentralnervöser Aktivierung und deren nachfolgende Reduktion gelten somit als Indikatoren dafür, dass die vorangegangenen Erlebnisse einerseits potenziell wichtig waren und andererseits erfolgreich gemeistert wurden, sodass sie aus biologischer Sicht einspeicherungswürdig sind.

wobei sowohl sehr niedrige Aktivierung (Langeweile) als auch sehr hohe Aktivierung (Stress) negativ erlebt werden, während sich bei mittlerem Aktivierungsniveau oft positive Emotionen und Stimmungen einstellen (z.B. Interesse, Leistungsmotivation, Unternehmungslust).

Routtenberg (1968) kommt aufgrund neuropsychologischer Befunde zu dem Schluss, dass positive Erlebnisse jeglicher Art (z.B.

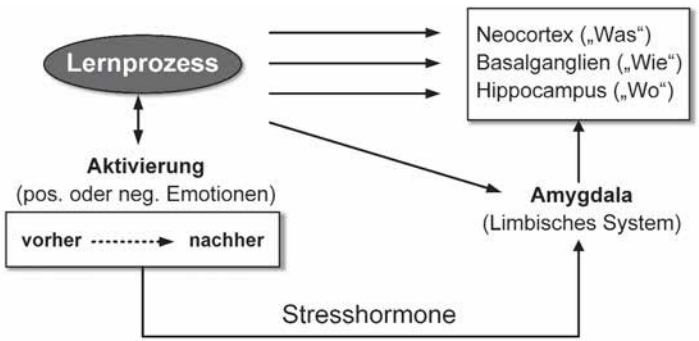
Abb 6.6



In einem von Brenner (1973) eingeführten experimentellen Design zur Gedächtnisprüfung sitzen im Kreis 11 Paare von Probanden, von denen jeweils einer ein Drei- bis Zwölfbuchstabenwort vorzulesen hat und der zweite nur zuhören muss. Die Versuchspersonen haben die Aufgabe, sich alle vorgelesenen Wörter (d.h. auch jene der anderen) zu merken. Wie erwartet, lag beim Wiedererkennen die durchschnittliche Erinnerungsleistung höher (ca. 75 %) als beim Reproduzieren (ca. 30 %). Die Wörter, die von den vorangehenden Personen (Positionen −5 bis −1) vorgelesen worden waren, bevor die Person selbst an die Reihe kam (Position 0), wurden schlechter gemerkt („next-in-line effect"). Eine enorme Einprägungswirkung ergab sich jedoch sogar beim Reproduzieren der Wörter, die die Probandinnen und Probanden selbst vorgelesen hatten (ca. 85 bzw. 95 % beim Wiedererkennen). Erklärbar wäre dies durch die Aktivierung beim Vorlesen und die nachfolgende Entspannung, die der passive Partner kaum miterlebt.

Triebbefriedigungen, Erfolgserlebnisse, „Aha-Erlebnisse“) die Belohnungszentren im *Limbischen System* ansprechen lassen, welche ihrerseits das Aktivierungssystem hemmen und auf diese Weise zu einer Reduktion unspezifischer Erregung und gleichzeitig zu einer Verstärkung lernspezifischer Neuronenverbindungen im Zentralnervensystem führen (s. Birbaumer, 1975). Bei diesem aktivierungsbezogenen Speichermechanismus könnte es sich um ein biologisches Grundprinzip des Lernens handeln, welches den meisten Lernformen zugrunde liegt (z.B. *Signallernen*, *Verstärkungslernen*, *kognitives Lernen*; s. auch „Brain Trigger Design“, Guttmann & Bauer, 1984).

Abb 6.7



Eine Reihe von Untersuchungen zum Lernen befasste sich mit dem Einfluss von Körpersignalen (z.B. Herzschlag), Stress (z.B. leichte elektrische Schocks) und positiven oder negativen Emotionen (z.B. Filmeeffekte) auf die Einprägung. Die dabei gewonnenen Erkenntnisse zeigten, dass jegliche Formen der Aktivierung einer Person vor dem Lernprozess (ca. 15 min), währenddessen und danach (ca. 15 min) Einprägungssteigerungen nach sich zogen (s. Schwabe et al., 2008; Critchley & Garfinkel, 2018; McGaugh, 2015, 2017; Roesler & McGaugh, 2019). Als wichtige Vermittler für den Einprägungsprozess gelten Stresshormone (Adrenalin, Cortisol), welche die Amygdala aktivieren, die ihrerseits Lernstrukturen des ZNS (Hippocampus, Basalganglien, Neocortex) positiv beeinflusst.

Speicherstrukturen des Gehirns 6.3

Sehr frühe Erkenntnisse über die Auswirkungen von Gehirnprozessen auf seelische Abläufe hatten bereits die Ägypter (ca. 2500 v.

Chr.). Griechische Gelehrte entdeckten ca. 500 v. Chr. die kreuzweise Zuständigkeit der Hirnhälften für die muskuläre Kontrolle der Körperhälften (Finger, 1994). Im 19. Jahrhundert wurden elektrische Gehirnreizungen im Tierexperiment eingesetzt, erste Funktionskartierungen der Kortexoberfläche von Hunden erarbeitet und Sprachstörungen als Folge von Gehirnschädigungen in bestimmten Kortexbereichen aufgedeckt (Broca'sches und Wernicke-Sprachzentrum). Nachdem 1906 Camillo Golgi und Santiago Ramón y Cajal für ihre Arbeiten über Nervenzellen und deren weitläufige Vernetzungen („Neuronendoktrin") den Nobelpreis bekommen hatten, intensivierte sich im 20. Jahrhundert die neurologische und neuropsychologische Forschung, nicht zuletzt durch den breiten Einsatz von Computern, sodass heute zumindest grobe Modelle über die Funktionsweise des Gehirns vorliegen (Abb. 6.8).

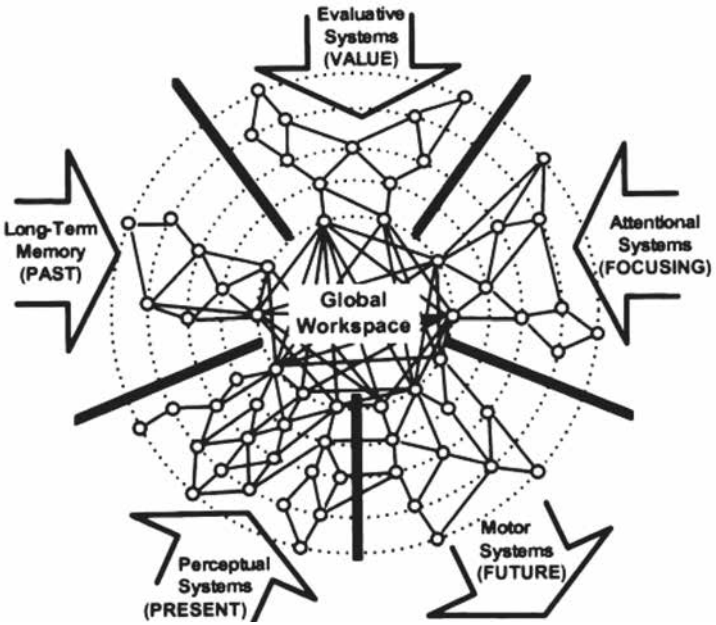
Einfache regulatorische Anpassungsprozesse finden pausenlos im zentralen und periphären Nervensystem statt (z.B. Ausdifferenzierung von Nervenzellfortsätzen, Produktion von „Neurotransmittern"). Komplexe Anpassungen hingegen, wie wir sie beim Lernen vermuten können, sind im Wesentlichen an drei miteinander in enger Beziehung stehende Systemkomplexe des Gehirns gebunden (Abb. 6.9): **Hirnstamm** (Aktivierungssystem), **Limbisches System** (Bewertungssystem) und **Großhirn** (Speichersystem). Den drei genannten Gehirnregionen kommen im Wesentlichen folgende Funktionen zu (Birbaumer & Schmidt, 2006):

Merksatz

Zu den wichtigsten Gehirnstrukturen für Lernprozesse zählen: 1. Hirnstamm (Aktivierungsregulation, vegetative Reaktionen), 2. Limbisches System (emotionale Reaktionen, Konsolidierung), 3. Großhirn (Speicherung und Klassifikation von Lerninhalten, Aufmerksamkeitssteuerung, Abrufung von Speicherinhalten).

Hirnstamm (Mittelhirn, Hinterhirn, Nachhirn): Diese Region kontrolliert das allgemeine Aktivierungsniveau und viele vegetative Basisfunktionen (Herzschlag, Atmung, Blutdruck, Verdauung, Wärmeregulation, Schlaf) und ist zudem der Ursprung von zehn „Hirnnerven" (sensorische und motorische Kontrolle des Kopfes und des Halses: Auge, Ohr, Mund, Kehlkopf, ...) einschließlich des Nervus vagus („Parasympaticus"). Ebenso sind dem Hirnstamm Instinktregelungen der Selbsterhaltung zuzuschreiben, wie etwa das Auslösen von Kampf- oder Fluchtreaktionen, Impulse zum Territorialverhalten, soziales Dominanzverhalten. Die vom Hirnstamm ausge-

Abb 6.8

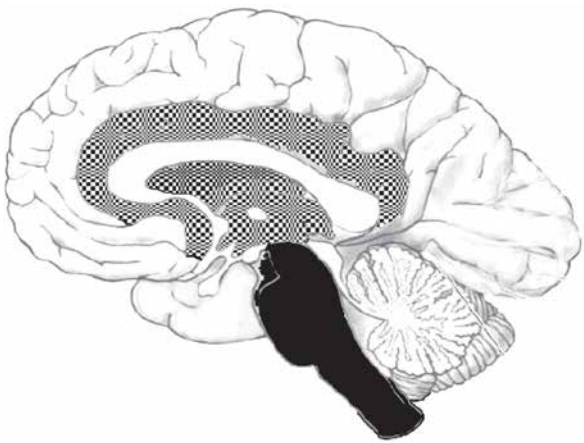


Das Modell des globalen Arbeitsspeichers gibt die heutige Sicht des Gehirns als Gruppierung stark vernetzter, miteinander simultan interagierender Neuronensysteme wieder. Teile des Netzwerkes (Module) erfüllen Spezialleistungen (Wahrnehmung, Erinnerung, Bewertung, Konzentration, Verhalten) und werden durch ein Verbindungsnetzwerk („Global Workspace") in ihren Aktivitäten gesteuert (s. auch Abb. 4.2).

henden Verhaltensprogramme sind allerdings noch weitgehend autonom und entsprechend starr.

Limbisches System (Zwischen- und Endhirnanteile): Diese Gehirnareale sind untereinander sowie mit den übrigen Gehirnbereichen stark vernetzt und werden auch als „emotionales Gehirn" bezeichnet, weil sie wesentlich an der Entstehung von Gefühlen, der Bewertung von Lebenssituationen und der Verstärkung von Verhaltenstendenzen beteiligt sind. Die Funktionstüchtigkeit des Hippocampus, einer Teilstruktur des Limbischen Systems, ist Voraussetzung für das dauerhafte Einprägen von Erlebnissen und Wissensinhalten (*Konsolidierung*).

Abb 6.9

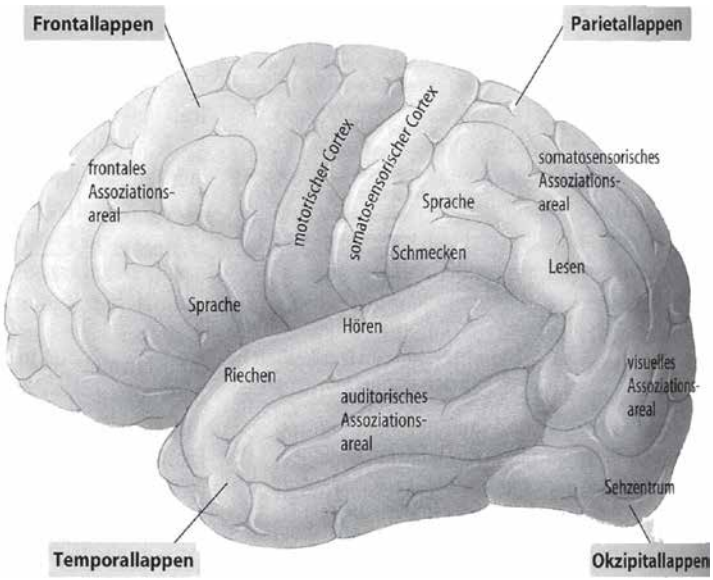


Aktivierungssystem (Hirnstamm – schwarz), Bewertungssystem (Limbisches System – kariert) und Speichersystem (Großhirn, Kleinhirn – weiß).

Großhirn bzw. **Neocortex**: Dieser Teil des Gehirns hat beim Menschen und bei anderen Primaten das größte Volumen. Vor etwa 100 Jahren entstanden die ersten Neocortex-Kartierungen, die den vier **Gehirnlappen** (Stirnlappen/Frontallappen, Scheitellappen/Parietallappen, Schläfenlappen/Temporallappen, Hinterhauptslappen/Okzipitallappen) und ihren Windungen spezielle Funktionen und Aufgaben der Informationsverarbeitung zuschrieben, was inzwischen bestätigt wurde (Abb. 6.10). Allerdings darf nicht mehr davon ausgegangen werden, dass Merkinhalte oder Fertigkeiten ausschließlich an einer Stelle des Großhirns lokalisiert sind. Neurologische Befunde bei Gehirnverletzungen (Gehirnläsionen) lassen deutlich erkennen, dass viele Einprägungen an verschiedenen Stellen des Neocortex gespeichert sind und auch die Kontrolle psychischer Prozesse nur selten von einem Gehirnareal allein ausgeht. Interessanterweise dürften auch geschlechtsspezifische Unterschiede vorhanden sein (Cahil, 2006).

Ebenso jedoch, wie sich für die Gehirnlappen eine Funktionsspezialisierung erkennen lässt, gibt es eine solche auch für die linke und rechte **Gehirnhemisphäre** (Gehirnhälfte; s. etwa Dehaene, 2009). Während in der sogenannten „dominanten" Hemisphäre schwerpunktmäßig Sprachverständnis, Sprachproduktion und sprachgebundenes logisches Denken gesteuert werden, ist die

Abb 6.10



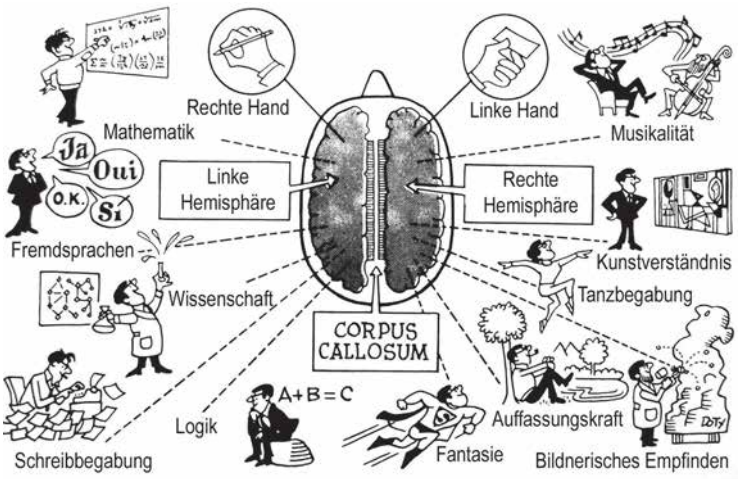
Die Orte der Speicherung des Faktenwissens (deklaratives Gedächtnis) liegen je nach Wahrnehmungsmodus (Tastsinn, Geschmack, Gehör, Geruch, Sehsinn) im Scheitel-, Schläfen- oder Hinterhauptslappen. Das Handlungswissen (prozedurales Gedächtnis), Handlungsorientierungen (Werthaltungen) und die Handlungskontrolle (Aufmerksamkeit, Denkprozesse) sind überwiegend im Vorderlappen des Großhirns lokalisiert.

nichtdominante Hemisphäre vorwiegend für Vorstellungsleistungen, musisches Verständnis und vorstellungsgebundenes logisches Denken verantwortlich (Abb. 6.11). Bei etwa 95 % aller rechtshändigen Personen ist die linke Hemisphäre dominant, die linksseitige **Hemisphärendominanz** gilt jedoch auch bei ca. 75 % aller Linkshänder (Springer & Deutsch, 1987). Der rechten Hemisphäre werden insbesondere auch Leistungen der Einsicht (Problemlösen) zugeschrieben (Kounios & Beeman, 2014).

Merksatz

Die linke Großhirnhälfte kontrolliert überwiegend die Speicherung, Verarbeitung und Produktion sprachlicher Signale, während die rechte Großhirnhälfte zumeist stärker auf die Speicherung, Verarbeitung und Produktion von Vorstellungsinhalten spezialisiert ist.

Abb 6.11

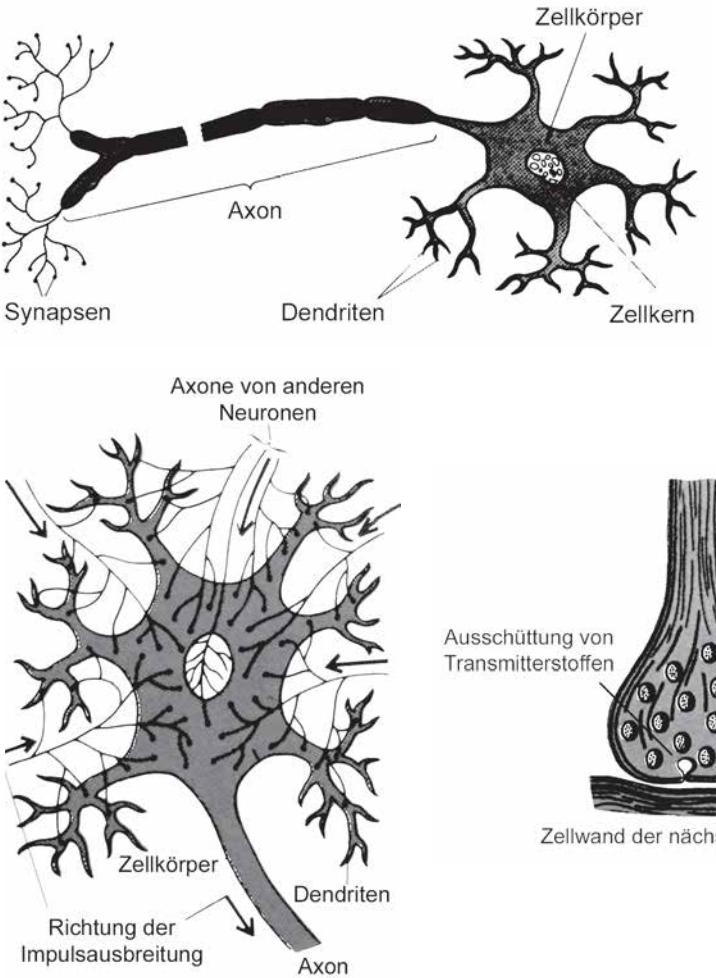


Eine stark pointierte Darstellung funktionaler Unterschiede zwischen linker und rechter Gehirnhemisphäre, die sich aufgrund von klinischen Erfahrungen (Gehirnverletzungen), neurologischen Studien (z.B. Split-Brain-Forschung) und neuropsychologischen Experimenten (z.B. dichotisches Hören) vermuten lassen (Springer & Deutsch, 1987). Trotz solch tendenzieller Lateralisierungen von Funktionen werden zweifellos die meisten Gehirnfunktionen von beiden Hemisphären beeinflusst und kontrolliert.

6.4 Neuronale Netzwerkmodelle

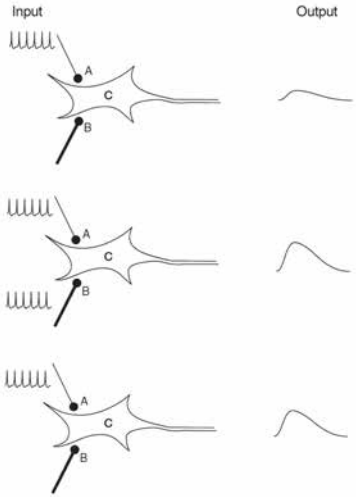
Die vermutlich 25 Milliarden Nervenzellen des Menschen (Birbaumer & Schmidt, 1991; Abb. 6.12) zeichnen sich außer durch ihre herausragende elektrochemische Erregbarkeit (zur Impulsweiterleitung) auch durch ihr Potenzial an Verbindungsfähigkeit aus. Eine einzige Nervenzelle kann bis zu 10.000 Kontakte zu anderen Nervenzellen entwickeln, und diese Kontakte können unterschiedlich stark ausgeprägt sein. Das heißt, es ergeben sich fast unendlich viele Konfigurationen von unterschiedlichen Vernetzungen als Ergebnis von Lernprozessen oder sonstigen Anpassungsleistungen. Nervenzellen modifizieren ihre Kontakte zueinander sowohl durch Vermehrung oder Verbreiterung der Kontaktstellen (Dendriten, Synapsen) als auch durch Änderung der Impulsübertragung (Ausschüttung von *Transmitterstoffen*).

Abb 6.12



Die Grundbausteine des Nervensystems, die Nervenzellen, bestehen aus einem Zellkörper und einem Zellkern, welche beide für den Stoffwechsel zuständig sind, sowie aus zahlreichen, manchmal bis zu einem Meter langen Nervenzellfortsätzen, von denen die Dendriten auf die Aufnahme und die Axone („Neuriten") auf die Abgabe von Information spezialisiert sind. Wenn eine Nervenzelle an ihren Dendriten oder am Zellkörper mechanisch, elektrisch oder chemisch erregt wird, dann pflanzt sich der elektrische Impuls bis zum Ende des Axons fort, wo an dessen Endknöpfchen, den Synapsen, Transmitterstoffe mit erregender oder hemmender Wirkung für die folgende Nervenzelle ausgeschüttet werden.

Abb 6.13

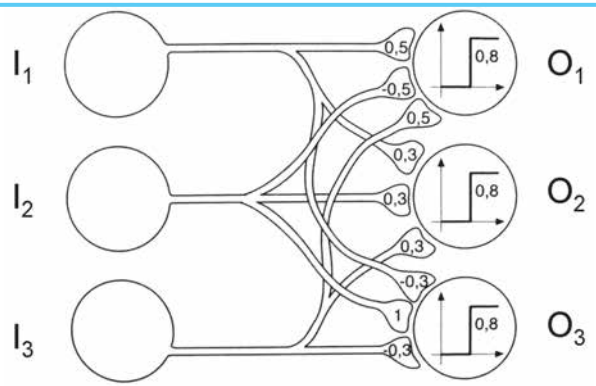


Das Phänomen der Langzeitpotenzierung (LTP) besagt, dass die Verbindung zwischen Nervenzellen dann verstärkt wird, wenn beide gleichzeitig erregt sind („Hebb'sche Lernregel"): Wenn beispielsweise ursprünglich eine Synapse der Nervenzelle A bei der Nervenzelle C einen geringen Output auslöst, diese anschließend gleichzeitig durch eine zweite Nervenzelle B stärker erregt wurde, dann löst später die Reizung A bei C ebenfalls einen stärkeren Impuls aus.

Ein wichtiger Mechanismus der Modulation von Verbindungen zwischen Nervenzellen ist die **Langzeitpotenzierung** (Abb. 6.13). Darunter versteht man die Verstärkung der Impulsübertragung zwischen Nervenzellen, die zur gleichen Zeit erregt sind. Hebb (1949) vertrat die Auffassung, dass die zahlreichen Querverbindungen zwischen Neuronen kreisförmige Erregungen auslösen, welche zur dauerhaften Festigung (*Konsolidierung*) der entsprechenden Erregungskonstellationen und damit zur langzeitigen Gedächtniseinprägung führen.

Die Einblicke in die Funktionsweise von Nervennetzen bei Lernprozessen haben zur Entwicklung von Theorien und technischen Modellen geführt, die als neuronale **Netzwerkmodelle** bekannt wurden und als Theorienrichtung unter dem Oberbegriff **Konnektionismus** zusammengefasst werden. Man versteht darunter die idealtypische Simulation von Netzwerken, die dem Nervensystem nachempfunden sind und bestimmte Lernleistungen (z.B. Mustererkennung, Wissensspeicherung) selbstorganisiert bewältigen. Die Mo-

Abb 6.14



Am Beispiel eines einfachen künstlichen neuronalen Netzwerkes (Spitzer, 2002) mit nur drei Input- und drei Outputneuronen kann selbstorganisiertes Lernen, das prinzipiell auch im Zentralnervensystem stattfindet, demonstriert werden. Vor dem Lernprozess sind die Kontaktstellen („Synapsen") zwischen den Input- und Outputneuronen des Netzwerkes nur mit kleinen positiven oder negativen Zufallswerten für die Verbindungsstärke besetzt. Wenn dann in wiederholten Durchgängen etwa immer wieder die gleichen drei Muster von Inputelementen aktiviert werden (z.B. $I_1 + I_3$ oder $I_1 + I_2 + I_3$ oder I_2) und das Outputneuron für das 1. Muster die Bezeichnung O_1, jenes für das 2. O_2 und jenes für das 3. O_3 bekommt, wird die Differenz zwischen der jeweils zustande kommenden Gewichtssumme an den Kontaktstellen der Outputneuronen und ihrer Idealaktivierung (= 1,0) errechnet. Gemäß diesem „Fehler" werden die Gewichtswerte jeweils um einen bestimmten kleinen Betrag angepasst („Backpropagation"). Nach 300–500 Lerndurchgängen erfolgt zumeist eine zufriedenstellende Unterscheidung der drei Muster. In der Abbildung sind die nach einem derartigen Lernprozess erhaltenen Gewichtswerte an den Kontaktstellen zu den Outputneuronen eingetragen, die erst ab einem Schwellenwert von 0,8 ansprechen:
O_1 reagiert nur noch bei Aktivität von I_1 und I_3:
$I_1(0,5) + I_3(0,5) = O_1(1,0)$
O_2 nur noch bei Aktivität von allen drei Inputelementen:
$I_1(0,3) + I_2(0,3) + I_3(0,3) = O_2(0,9)$
O_3 nur noch bei Aktivität von I_2:
$I_2(1,0) = O_3(1,0)$

delle bestehen aus Mengen von Funktionseinheiten („Neuronen"), die zumeist in Schichten aufgeteilt (z.B. Input-, Zwischen- und Outputschicht) und vielfältig miteinander in Beziehung gesetzt sind. Durch Anwendung spezieller Rechenregeln („Lernalgorithmen") werden die Verknüpfungen zwischen den Einheiten so verstärkt oder geschwächt, dass sich die angestrebten Lernleistungen optimal ausbilden (Abb. 6.14).

Komplexere neuronale Netzwerke (z.B. solche mit hunderten Neuronen und mehreren Zwischenschichten) vollbringen heute be-

reits aufwendige Adaptations- und Klassifikationsleistungen, die im industriellen Bereich (Steuerungsanlagen, Expertensysteme) und in der Forschung (Simulation kognitiver Leistungen, statistische Prognosen) erfolgreich und effizient eingesetzt werden (s. etwa „IBM SPSS Neuronal Networks"; Kriegeskorte, 2015).

6.5 | Habituation

Eine der ältesten Formen der Anpassung von Lebewesen an die Umwelt ist die Habituation, nämlich die Gewöhnung an diejenigen häufig auftretenden Umweltreize, die mit der Zeit als irrelevant eingestuft werden. Neue, unerwartete Reize lösen üblicherweise eine Aktivierungssteigerung aus („Orientierungsreaktion"), die mit einer Aufmerksamkeitszuwendung verbunden ist. Wenn dem Reiz danach aber weder positive noch negative Konsequenzen regelmäßig folgen, kommt es zu einer Habituation.

Merksatz

Wenn Lebewesen auf Reize, die normalerweise eine Aktivierungssteigerung oder eine Aufmerksamkeitszuwendung bewirken, weniger oder gar nicht mehr reagieren, spricht man von Gewöhnung oder Habituation.

Das zunehmende Nichtreagieren auf irrelevante Umgebungsmerkmale entlastet die Informationsverarbeitung und hält die geistigen und körperlichen Ressourcen frei für wichtigere Erfahrungen. Adaptationen dieser Art finden sich bereits bei vielen einfachen Lebewesen, wie Schnecken, Würmern oder Insekten. Habituationsreaktionen beim Menschen weisen folgende Charakteristika auf (Mazur, 2004):

- Je anhaltender oder häufiger ein Reiz auftritt, desto schneller gewöhnt man sich an ihn.
- Der Gewöhnungseffekt ist anfangs am größten und nimmt dann kontinuierlich ab.
- Wenn ein habituierter Reiz längere Zeit nicht auftritt, löst er später wieder eine Orientierungsreaktion aus.
- Wenn eine Gewöhnung erworben und danach „verlernt" wurde, kann sie später schneller wieder reaktiviert werden.
- An intensive Reize gewöhnt man sich langsamer als an schwache Reize.
- Die Gewöhnung an einen bestimmten Reiz überträgt sich auch auf andere ähnliche Reize.

Die Habituationsforschung leistete einen wichtigen Beitrag bei der Entwicklung moderner Angsttherapien (*Verhaltenstherapie*), bei denen die Prinzipien des „Konfrontierens“ und des „Floodings“ (massierte Konfrontation mit Angstreizen) eine wichtige Rolle spielen (Box 12.6).

Prägungsartiges Lernen

6.6

In der Verhaltensforschung („Ethologie“) wird seit Lorenz (1935) jener Vorgang als **Prägung** bezeichnet, durch den sich in der frühen Entwicklung vor allem bei „Nestflüchtern“ (Hühner, Enten, Gänse) bestimmte Verhaltensweisen irreversibel ausbilden. Prägungen können sich auf Nachfolgereaktionen (**Nachlaufprägung**),

Merksatz

Prägung ist die irreversible Verknüpfung eines zweckgerichteten Verhaltens mit bestimmten Reizen in frühen Lebensphasen (Nachfolgeprägung, Gesangsprägung, Heimatprägung usw.) bestimmter Tierarten.

Abb 6.15

Der österreichische Nobelpreisträger Konrad Lorenz (1903–1989), Mitbegründer der vergleichenden Verhaltensforschung, demonstriert die Nachlaufprägung bei Entenküken, die nach dem Schlüpfen nicht die Entenmutter, sondern ihn gesehen hatten.

den Gesang („Gesangsprägung"), auf Sexualpartnercharakteristiken oder auf Eigenschaften des Territoriums („Heimatprägung") beziehen (Abb. 6.15). Charakteristisch für Prägungen sind folgende Merkmale (s. Eibl-Eibesfeldt, 1999; Hess, 1972):

- Prägungen können nur in einer sensiblen Periode, zumeist in frühen Entwicklungsstadien des Lebewesens, stattfinden (z.B. wenige Stunden nach dem Schlüpfen von Küken).
- Prägungen sind irreversibel bzw. „therapieresistent" (z.B. sexuelle Prägungen, wie etwa die „Menschenperversität" von Gänsen).
- Prägungen beziehen sich nicht auf individuelle Reize, sondern auf Reizklassen (z.B. Menschen, Tierarten, Gesangsart, Lebensräume).
- Geprägt wird nur eine bestimmte Reaktion bzw. ein Rollenverhalten (z.B. Eltern- oder Geschlechtsrolle).
- Prägung kann schon stattfinden, wenn sich das zu prägende Verhalten noch gar nicht zeigt (z.B. Gesang, Sexualverhalten).

Die Frage, ob es beim Menschen (als „Nesthocker") ebenfalls zu Prägungsphänomenen kommen kann (z.B. Prägungen auf Sexualpräferenzen), wird von den meisten Psychologen verneint. Klix (1971) vertritt allerdings die Ansicht, dass sich beim Menschen zumindest prägungsartiges (d.h. fast irreversibles) Verhalten herausbilden könnte, wenn im Stadium eines hohen Trieb- oder Aktivierungszustandes (*Stress*, sexuelle Erregung) bestimmte Verhaltensweisen (Fluchtreaktionen, sexuelle Praktiken) zu einem massiven Entspannungszustand führen und somit stabil eingeprägt werden. Denkbar wäre etwa, dass in einem jugendlichen Entwicklungsstadium ein extrem erhöhter sexueller Triebzustand durch eine besondere Art der Befriedigung zu einer sexuellen Fixierung (z.B. „Fetischismus") führen könnte. Die meisten Lernforscherinnen und -forscher sind jedoch überzeugt, dass es sich hierbei nicht um Prägung, sondern nur um eine intensive *Konditionierung* handelt, die durch konkurrierende neue Erfahrungen oder innerhalb eines psychotherapeutischen Prozesses wieder verändert werden können.

6.7 | Klassische Konditionierung – Signallernen

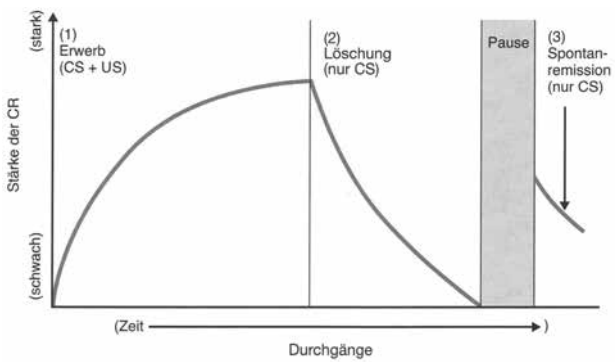
Der russische Physiologe Iwan Pawlow (1849–1936) entdeckte das Phänomen sogenannter „psychischer Sekretionen" als Zufallsbe-

fund bei seinen Untersuchungen zum Verdauungsprozess von Hunden, die ihm später den Nobelpreis (1904) einbrachten. Er beobachtete bei seinen Versuchstieren, dass der durch Futterdarbietung ausgelöste Speichelreflex auch auf andere Auslösereize übertragbar ist, sofern diese dem Anblick von Futter vorausgingen (z.B. Öffnen der Labortüre, Hereinkommen des Tierpflegers). Diese Form des Lernens manifestiert sich auch in zahlreichen Alltagssituationen des Menschen immer dann, wenn Reize oder Situationen das nachfolgende Auftreten lebensrelevanter Erlebnisse signalisieren (z.B. Reflexverhalten, Triebbefriedigung, Schmerz, Angst). Angesichts der assoziativen Koppelung von ursprünglich neutralen mit bedeutungsvollen Reizen wird diese Lernform auch als **Signallernen** oder „Lernen vom Typ S" (Reiz: engl. stimulus) bezeichnet.

Allgemein wird die klassische Konditionierung durch folgende Begriffe charakterisiert: Ein unkonditionierter Stimulus US löst eine angeborene Reaktion UR aus, zum Beispiel: Futterdarbietung → Speichelfluss, Lichtblitz → Pupillenkontraktion, Pistolenschuss → Schreckreaktion, Luftstoß → Lidschlussreflex. Geht nun dem US regelmäßig ein anderer Reiz voraus und kündigt ihn damit an, dann löst dieser später auch allein als CS eine ähnliche Reaktion wie die UR aus, welche als CR bezeichnet wird (Abb. 6.16).

US: engl. unconditioned stimulus
UR: engl. unconditioned response
CS: engl. conditioned stimulus
CR: engl. conditioned response

Abb 6.16



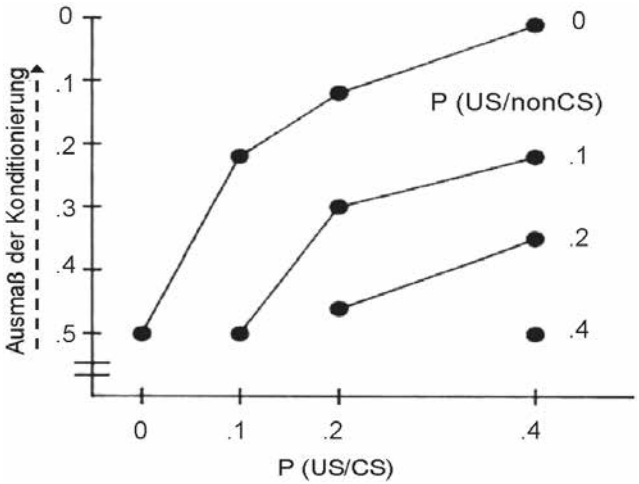
Phasen klassischer Konditionierung:
Erwerb: Mit der Anzahl an Koppelungen zwischen CS und US wird die konditionierte Reaktion (CR) stärker.
Löschung: Wenn später nur der CS (ohne nachfolgenden US) dargeboten wird, kommt es zu einer Abschwächung der CR.
Spontanerholung: Das Wiederauftreten einer gelöschten konditionierten Reaktion nach einer Pause.

Für die Lernform der klassischen Konditionierung sind folgende Aspekte bedeutsam (vgl. Anderson, 2000):

- **Koppelungsanzahl** und **Zeitdifferenz**: Bei den meisten Konditionierungen reichen 10–50 Koppelungen zwischen CS und US aus, um eine deutliche CR hervorzurufen. Je größer die Zeitdifferenz ist, mit der ein CS einem US vorausgeht, desto schwieriger ist der Aufbau einer Konditionierung auf diesen CS, wobei das wirksame Zeitintervall zwischen CS und US von der Geschwindigkeit der Reaktion abhängt (z.B. 0,3 bis 1,0 Sek. beim Lidschlag, 10 Sek. bei Angst, bis zu mehreren Stunden bei Geschmackskonditionierungen).
- **Kontingenz**: Nicht nur raumzeitlich benachbartes Auftreten („Kontiguität"), sondern vor allem zeitliches Nacheinander („Kontingenz") zwischen CS und US ist entscheidend für das Konditionierungslernen: Je besser ein Reiz einen anderen vorherzusagen gestattet, desto stärker wirkt er als Signalreiz bei einer Konditionierung (Abb. 6.17).
- **Reizintensität** und *Auffälligkeit*: Die Intensität und die Auffälligkeit eines CS erhöhen seine Wirksamkeit als Signal.
- **Informativität**: Wenn bereits ein CS_1 (z.B. Ton) als Signal für einen US aufgebaut wurde und ein weiterer, zur Vorhersage des US unnötiger Reiz CS_2 (gleichzeitig) hinzukommt, wird dieser ignoriert und löst später keine CR aus („Blockierung").
- **Generalisation** und **Diskrimination**: Die Auslösequalität eines CS für eine CR überträgt sich auch auf andere Reize im Ausmaß ihrer Ähnlichkeit mit dem ursprünglichen CS. Diese Ausweitung der Signalwirkung eines Reizes kann aber durch ein Unterscheidungstraining eingeschränkt werden.
- **Konditionierung höherer Ordnung**: Für einen bereits konditionierten Reiz CS_1 kann vorher oder nachher ein anderer Reiz CS_2 als Signal erlernt werden („preconditioning", „second-order conditioning").

Alltagsbeispiele für klassische Konditionierungen sind äußerst zahlreich und beziehen im Grunde alle Reaktionsbildungen mit ein, bei denen Reflexe, physiologische Prozesse, Gefühle oder Triebauslöser durch Signale angekündigt werden. Wenn uns beim Sprechen über Lieblingsspeisen „das Wasser im Mund zusammenläuft", wenn wir uns beim Eintritt in ein uns bekanntes Prüfungszimmer plötzlich unbehaglich fühlen oder wenn wir beim Anblick bestimmter Gesten plötzlich – für andere unerklärlich – ärgerlich werden,

Abb 6.17



Nur wenn der konditionierte Reiz (CS) den unkonditionierten Reiz (US) verlässlich vorherzusagen imstande ist, kann Konditionierung stattfinden. Rescorla (1968) testete die Wirksamkeit unterschiedlicher bedingter Wahrscheinlichkeiten p(US|CS) sowie p(US|nonCS) auf die Stärke der Konditionierung (Abzisse: p für US, wenn vorher CS auftrat; Funktionen: p für US, wenn CS nicht auftrat). Der Effekt war umso größer, je eindeutiger vom Signal auf den US geschlossen werden konnte, z. B. wenn die Wahrscheinlichkeit für den US bei Auftreten des CS 0,4 und bei Nichtauftreten des CS nur 0,0 war.

haben wir es mit emotionalen Konditionierungen zu tun („conditional emotional response", CER), die manchmal bis in unsere Kindheit nachzuverfolgen sind.

Auf dem Gebiet der Angstkonditionierung ist die Entdeckung von Seligman (1970, 1971) bedeutsam, dass nicht alle Reize gleichermaßen einfach und nachhaltig (löschungsresistent) angstkonditionierbar sind, sondern dass im Laufe der menschlichen Evolution bestimmte Reize eine **Prädisposition** als Angstauslöser erworben haben („preparedness"), wie etwa der Anblick von Schlangen, Spinnen, Ungeziefer, Tiefe und Weite, im Gegensatz zum Anblick von Steckdosen oder Autos, die nicht minder gefährlich sind.

Merksatz

Beim klassischen Konditionieren bzw. Signallernen werden Reize als Ankündigungen für solche Reize erlernt, die Reflexe, vegetative Reaktionen, Emotionen oder Triebverhaltensweisen auslösen.

6.8 | Instrumentelles Konditionieren – Erfolgslernen

E. L. Thorndike formulierte 1898 mit seinem „Gesetz des Effektes“ (engl. „law of effekt“) als Erster das Grundprinzip der instrumentellen Konditionierung: Verhalten ändert sich durch den Effekt, den es auslöst, belohnende Konsequenzen stärken die Verhaltenstendenz, bestrafende Konsequenzen schwächen sie (Box 6.1). Während bei der klassischen Konditionierung die Koppelung von zwei (oder mehreren) Reizen erlernt wird (S – S), sind es bei der instrumentellen Konditionierung Reize und Reaktionen (S – R), die miteinander verknüpft werden. Erlernt wird, unter welchen Bedingungen welche Reaktion zu welcher Konsequenz führt bzw. welches Verhalten erfolgreich ist.

Als „operant“ wurde diese Lernform von B. F. Skinner (1904–1990) deshalb bezeichnet, weil dabei Verhaltensweisen als Opera-

Box 6.1 | Skinners Vier-Felder-Schema

Seit Burrhus F. Skinner (1904–1990) werden in einem Vier-Felder-Schema die vier grundsätzlich möglichen Verhaltenskonsequenzen dargestellt, die unterschiedliche Wirkungen auf das zukünftige Verhalten ausüben:

	Positive Konsequenz	Negative Konsequenz
Eintreten der Konsequenz	(1) Positive Verstärkung Effekt: Verstärkung	(3) Bestrafung (Typ I) Effekt: Verhaltens-blockierung
Ausbleiben der Konse-quenz	(4) Bestrafung (Typ II) Effekt: Löschung	(2) Negative Verstärkung Effekt: Verstärkung

Beispiele:

(1) Kinder entwickeln jene Fähigkeiten, für die sie regelmäßig gelobt werden.

(2) Schlechte Gewohnheiten, wie Rauchen, Naschen, Alkoholtrinken, dienen meist der Reduktion von Anspannung, Stress oder Frustration.

(3) Verkehrsstrafen können Fehlverhalten im Verkehr reduzieren.

(4) Gegenüber Personen, die sich für Hilfeleistungen nicht bedanken, verliert man seine Hilfsbereitschaft.

tionen zur Veränderung der inneren und äußeren Realität aufgefasst werden. Das englische „operant" bezeichnet ein weitgehend spontanes Verhalten, das nicht (wie bei der klassischen Konditionierung) reflexartig durch einen Reiz ausgelöst wird. Die von Skinner als Forschungsprogramm eingeführte „experimentelle Verhaltensanalyse" bezweckte eine möglichst exakte Beschreibung des Einflusses von Umweltbedingungen auf die Auftrittswahrscheinlichkeit von Verhaltensformen.

Allgemein findet instrumentelles Lernen dann statt, wenn wiederholt eine bestimmte Situation wahrgenommen wird („diskriminativer Hinweisreiz"; SD), in der bestimmte Verhaltensweisen (R) zu bestimmten Konsequenzen führen (K). Wenn die Konsequenz die Auftrittswahrscheinlichkeit des Verhaltens erhöht, spricht man von **Verstärkung**, wenn sie die Auftrittswahrscheinlichkeit senkt, spricht man von **Bestrafung** (Box 6.1 und Box 6.2). Wie bei der klassischen Konditionierung unterscheidet man eine Phase des *Erwerbs*, der *Löschung* und der *Spontanerholung*. Folgende Aspekte kennzeichnen zusätzlich das instrumentelle Konditionieren als Lernform:

Merksatz

Instrumentelles Lernen findet dann statt, wenn ein Individuum wiederholt eine bestimmte Situation wahrnimmt, in der bestimmte Verhaltensweisen zu positiven oder negativen Konsequenzen führen.

- **Situations-** und **Reizkontrolle**: Situationen, Personen oder Objekte erlangen als diskriminative Hinweisreize einen „Aufforderungscharakter" zur Ausübung der belohnten Handlungen, wenn sie regelmäßig im Kontext der Verhaltensweisen mit Belohnungswirkung auftraten. Ein Sportplatz animiert Sportlerinnen und Sportler zum Training, eine bestimmte Lehrperson löst im Vergleich zu anderen Lehrpersonen bei Kindern undisziplinierteres Verhalten aus, eine bestimmte Musik reizt zum Tanzen. Auch ein „Hemmungscharakter" bestimmter Reize und Situationen wird auf diese Weise gelernt: So verhält man sich in Anwesenheit mancher Personen von vornherein zurückhaltender als gegenüber anderen Personen.
- **Kontingenz**: Je regelmäßiger und eindeutiger eine positive oder negative Konsequenz auf ein Verhalten wahrgenommen wird, desto schneller wird gelernt („Kontingenzregel").
- **Generalisation** und **Diskrimination**: Die Prozesse der Verallgemeinerung und der Spezifizierung können sich auf die Situationen, auf die Reaktionen sowie die Konsequenzen beziehen. In man-

Box 6.2 | „Sinnvolles Bestrafen“

Anhand seiner Forschungsergebnisse betonte Skinner immer wieder, dass Mensch und Tier hauptsächlich durch Belohnung lernen und dass Bestrafung nicht einfach die umgekehrte Wirkung von Belohnung hat. Auch spätere Lernforscher bestätigten, dass Belohnung eher dem Verhaltensaufbau und Bestrafung eher der Verhaltensblockierung dient, zumindest solange die Bestrafung real droht. Hinzu kommen die negativen emotionalen und sozialen Nebenwirkungen von Bestrafung, die ja oft eine Form sozialer Aggression darstellt (Schläge, Verbote, Drohungen, Kritik, Vorwürfe etc.). Für einen sinnvollen und dosierten Einsatz von Bestrafung haben Lernforscher Regeln aufgestellt (s. Zimbardo & Gerrig, 2004; Bourne & Ekstrand, 2005). Demnach sollten Bestrafungen ...

- unangenehm, schnell und kurz sein (Wirksamkeit)
- unmittelbar nach der unerwünschten Reaktion erfolgen (Kontingenz)
- dem Fehlverhalten angepasst sein (Intensitätsbegrenzung)
- als natürliche Konsequenzen des schadenverursachenden Verhaltens verstehbar sein
- sich auf das (veränderbare) Verhalten und nicht auf die Persönlichkeit beziehen (Verhaltensorientierung)
- auf die gegenwärtige Situation beschränkt bleiben (Situationseingrenzung, Aktualisierung statt Pauschalierung)
- keine körperlichen Schmerzen verursachen
- für die Zukunft ein Alternativverhalten nahelegen (z.B. als Vorschlag einer Wiedergutmachung oder einer zukünftigen Verhaltensänderung)

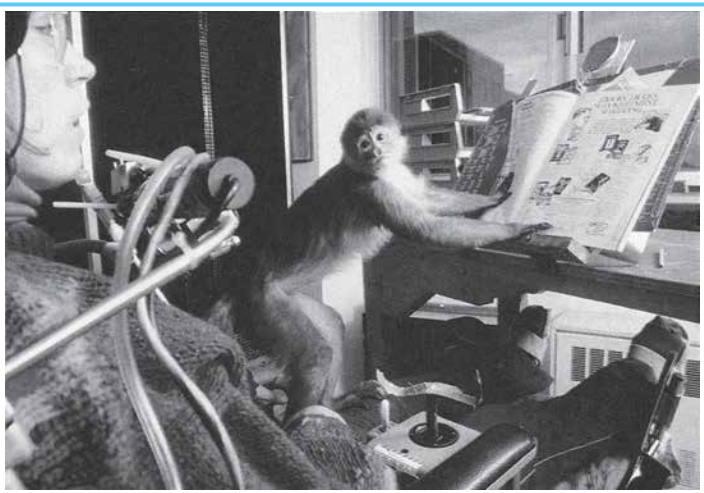
chen sozialen Situationen müssen zum Beispiel für die Auswahl eines erfolgreichen Verhaltens spezielle Bedingungen beachtet werden, für die eine gute *soziale Wahrnehmung* Voraussetzung ist. Ebenso können dem Erfolg enge Verhaltensgrenzen etwa im Sinne *sozialer Kompetenz* gesteckt sein. Aber auch angenehme Konsequenzen wirken auf manche Menschen nur in differenzierter Form (z.B. Lob nur, wenn nicht schulmeisterlich), während etwa Geld als generalisierter (erlernter) Verstärker gelten kann.

- **Shaping** (Verhaltensformung): In allen Lebensbereichen (Schule, Beruf, Privatleben) müssen Verhaltensweisen, die Erfolg bringen

oder Misserfolg vermeiden helfen, erst erworben und „trainiert“ werden. Die „richtigen“ Reaktionsweisen sind oft aus schwachen Verhaltensansätzen heraus zu entwickeln, was üblicherweise durch selektive Verstärkung der tendenziellen Reaktionen geschieht. Dieses Prinzip des Shapings (Abb. 6.18) wird von begabten Eltern in der Erziehung intuitiv genützt, wenn es etwa um den Erwerb von Pünktlichkeit, Ordnungssinn, Ehrlichkeit usw. durch das Kind geht. Diese Eltern ignorieren möglichst das unerwünschte Verhalten und nehmen alle Ansätze des erwünschten Verhaltens verstärkend wahr. Der Prozess des Shapings kann durch Vorbildwirkung (*Beobachtungslernen*) und verbale Aufforderungen oder Bitten unterstützt werden.

- **Primäre** und **sekundäre Verstärker**: Konsequenzen können entweder von vornherein angenehm (oder unangenehm) sein, oder sie erwerben diese Eigenschaft erst durch Koppelung mit anderen Verstärkern (*Signallernen*). Primäre, also unmittelbare Verstärker sind etwa Triebbefriedigungen, wie Essen, Trinken oder sexuelle Aktivitäten, sekundäre (erlernte) Verstärker dagegen sind Geld, Schulnoten, Lob oder Statussymbole.

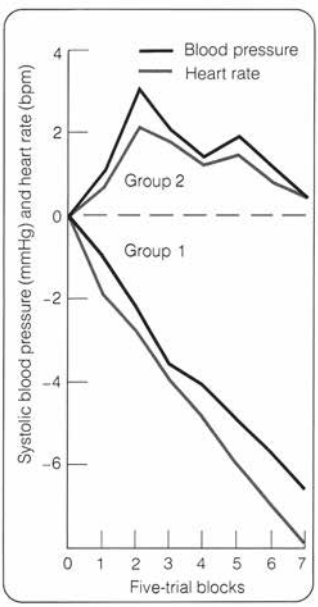
Abb 6.18



Shaping: Lange vor der wissenschaftlichen Erforschung des Konditionierens wussten Tierhalter, wie das Verhalten von Tieren durch den selektiven Einsatz von Belohnungen geformt werden kann. Für eine gelähmte Frau konnte zum Beispiel ein Äffchen dazu trainiert werden, ihre Haare zu kämmen, sie zu füttern, Buchseiten umzublättern etc.

Das instrumentelle Konditionieren bzw. Verstärkungs- oder Erfolgslernen zeigt sich im Alltag auf Schritt und Tritt und erklärt eine Vielzahl unserer beobachtbaren automatisierten oder bewussten Verhaltensweisen, insbesondere wenn man dabei auch die Wirksamkeit „verdeckter" Verstärker in Betracht zieht. Verdeckte Verstärker sind symbolische oder gedankliche Operationen mit stellvertretender Belohnungswirkung, wie sie von den *kognitiven Lerntheorien* (z.B. Bandura, 1977b) weiterentwickelnd beschrieben wurden (s. 6.11). Weniger bekannt dürfte sein, dass auch unwillkürliche physiologische Reaktionen wie Muskelentspannung oder Blutdruck durch Belohnungsrückmeldungen trainierbar sind, was therapeutisch etwa beim Einsatz von **Biofeedback** genutzt wird (Abb. 6.19). Diese Methode dient der Sichtbarmachung von physiologischen Prozessen wie Muskelaktivitäten, Atmung, Herzschlag, Durchblutung und anderen vegetativen Reaktionen, die oft schwer oder gar nicht wahrgenommen werden können und sich somit weitgehend der psychischen Beeinflussung entziehen. Damit ist Biofeed-

Abb 6.19



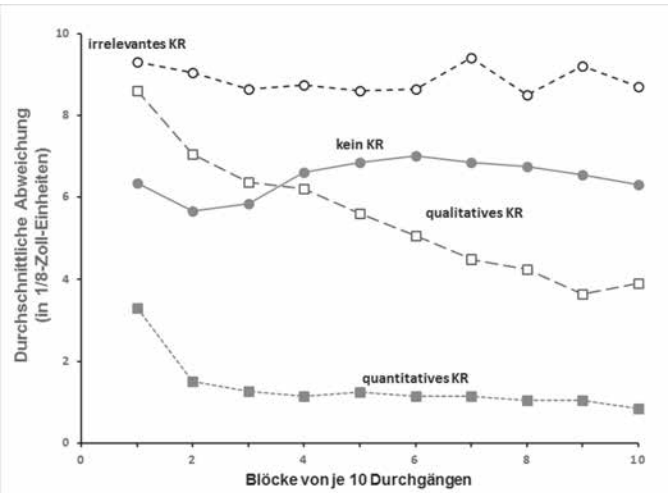
Als Biofeedback bezeichnet man die zumeist elektronische Registrierung und optische oder akustische Rückmeldung von physiologischen Reaktionen (z.B. Herzschlag, Blutdruck, Muskelspannung). Damit werden im physiologischen System Funktionsveränderungen trainierbar, die sonst nicht willkürlich steuerbar sind. Beispiel: Männliche Versuchspersonen bekamen in einem Experiment (Schwartz, 1975) immer dann ein Licht und einen Ton präsentiert, wenn durch natürliche Schwankungen die Herzrate und der Blutdruck minimal sanken (Gruppe 1) oder anstiegen (Gruppe 2). Nach jeweils zwölf sukzessiven „Erfolgen" bekamen sie als Belohnung Dias mit schönen Landschaften und hübschen Frauen zu sehen und erhielten zusätzlich einen Geldbonus. Nach 20–30 Durchgängen zeigte sich bei Gruppe 1 eine deutliche und konsistente „Entspannungsreaktion", während bei Gruppe 2 das Antrainieren höherer Blutdruck- und Herzratenwerte nicht so gut gelang.

back ein ideales Hilfsmittel, um die Eigenkompetenz von Klientinnen und Patienten bei der Behandlung von Verspannungen, vegetativen Stressreaktionen, muskulären Fehlfunktionen (z.B. bei Inkontinenz) und bei Schmerzerkrankungen zu fördern (s. Uher, 2008).

Fertigkeiten – Motorisches Lernen 6.9

Sprechen, Laufen, Radfahren, Autofahren, Klavierspielen, Turnen, Balletttanzen und Operngesang sind Beispiele für komplexe Fertigkeiten, die durch Übung erworben werden. Für alle Höchstleistungen an Geschicklichkeit gilt die Regel: „Ohne Fleiß kein Preis“ (engl. „No pain, no gain“). Untersuchungen an hochbegabten Personen („Genies“) im Bereich der Musik, der Wissenschaft und des Schachspiels ergaben, dass niemand seinen überragenden Leistungsstand ohne mindestens zehn Jahre intensiven **Trainings** erreicht hatte (s. Anderson, 1996).

Abb 6.20



In einem Experiment sollten Versuchspersonen mit verbundenen Augen eine drei Zoll lange Linie zeichnen (Towbridge & Cason, 1932). Die 1. Gruppe bekam keine Rückmeldung über die Länge der tatsächlich gezeichneten Linie (kein KR: „Knowledge of Result“), die 2. Gruppe erhielt nutzloses Feedback (irrelevantes KR), die 3. Gruppe hörte nur „richtig“ oder „falsch“ (qualitatives KR) und die 4. Gruppe bekam die quantitative Abweichung zu einer Drei-Zoll-Linie rückgemeldet (quantitatives KR). Am schnellsten lernten die Personen mit quantitativem Feedback.

Gedächtnis und Wissen

7

Inhalt

7.1 Einprägen und Vergessen
Einprägen als psychische Abbildung der Wirklichkeit
Lern- und Vergessenskurven
Vergessenstheorien

7.2 Kurzzeitspeicherung
Ultrakurzzeitgedächtnis (UKZG) – Sensorisches Gedächtnis (SG)
Kurzzeitgedächtnis (KZG)
Arbeitsgedächtnis

7.3 Langzeitspeicherung

7.4 Komponenten des Langzeitgedächtnisses

7.5 Stadien der Gedächtnisbildung
Aufnahme von Wissen (Enkodierung, Akquisition)
Festigung von Wissen
Abruf von Wissen

7.6 Gedächtnisregeln

7.7 Die PQ4R-Methode

Welche Informationen in welcher Form gespeichert sind, wie lange die Einprägungen gespeichert bleiben und wie die Speicherinhalte wieder reproduziert werden können, sind Grundfragen der Gedächtnisforschung. Sehr allgemein kann **Gedächtnis** als jene Instanz eines biologischen Systems verstanden werden, die die Aufgabe hat, verfügbare Information zu speichern und wiederzugeben. Da Gedächtnisprozesse schon beim Verhaltenslernen die Grundlage

Merksatz

Gedächtnis ist die mentale Fähigkeit, Informationen aufzunehmen, zu speichern und wiederzugeben.

bildeten, stellt dieses Kapitel sozusagen die Fortsetzung des vorigen Kapitels über das Lernen dar und richtet den Blick auf spezifische Aspekte der Informationsspeicherung, die im Alltag besonders im Zusammenhang mit Wissen interessant sind. Die gleichen Gesetzmäßigkeiten gelten jedoch grundsätzlich auch – sofern nicht ausdrücklich differenziert wird – für das implizite Wissen, welches nicht verbal abrufbar ist, sondern sich nur in Verhaltensänderungen manifestiert (Buchner & Brandt, 2002).

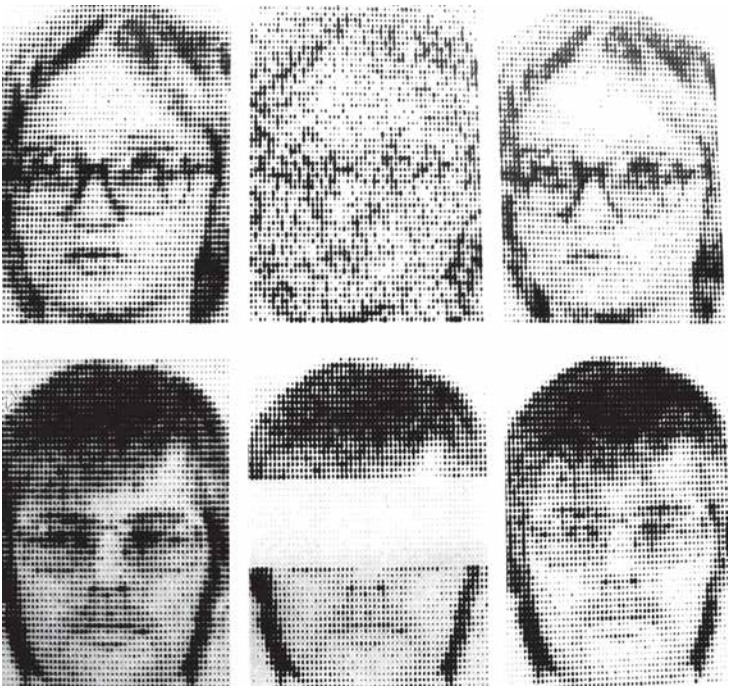
7.1 | Einprägen und Vergessen

7.1.1 | Einprägen als psychische Abbildung der Wirklichkeit

Als biologischer Zweck des Lernens im Allgemeinen und des **Einprägens** von Wissen im Besonderen kann die Entwicklung organismusinterner Modelle von Lebensumwelten gelten, die damit besser vorhergesagt und kontrolliert werden können. Die Speicherkapazität des Gehirns wäre jedoch überfordert, wenn die gesamte über die Sinne einströmende Informationsmenge (ca. 10^9 bit/s; Keidel, 1963) dauerhaft eingeprägt werden müsste (das wären ca. 57,6 Terabyte an einem 16-Stunden-Tag).

Es haben sich daher im Laufe der evolutionären Entwicklung im Nervensystem viele Filtermechanismen und Speicherstrategien herausgebildet, die auf eine sparsame und dennoch leistungsfähige Informationsverarbeitung abzielen. Diese besteht im Wesentlichen darin, dass aus den wechselnden Erscheinungsweisen unserer Wirklichkeit weitgehend konstante räumliche oder zeitliche Muster extrahiert werden, die in gespeicherter Form für die Selektion und Klassifikation von Ereignissen eingesetzt werden. Solche Muster betreffen Konfigurationen von Merkmalen (*Gestalten*), regelmäßige Aufeinanderfolgen von Reizen und Reaktionen bzw. Reiz-Reiz- und Reiz-Reaktions-Kombinationen (*Konditionierungen*), Ketten von Verhaltenselementen (*Fertigkeiten)* und abstrakte geistige Ordnungen (*Begriffe*, *Schemata* oder *Schlussfolgerungen*). „Gegeben ist jeweils die Erfahrung (also eine Teilmenge aller möglichen Erfahrungen der Realität), und gesucht (d.h. für den Organismus zu lernen) ist

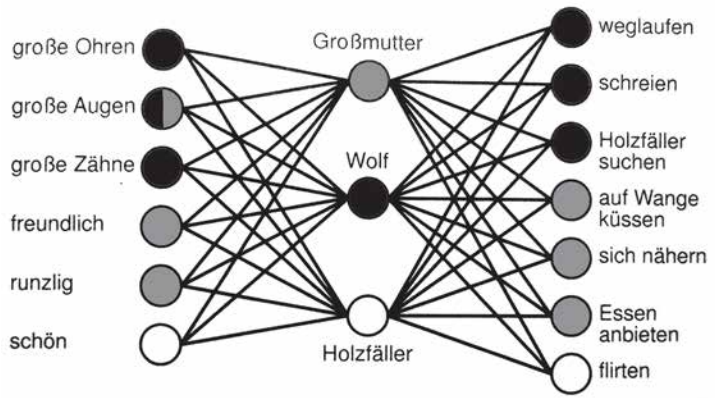
Abb 7.1



In sogenannten assoziativen Netzwerken („Hopfield-Netzwerke"), einer speziellen Form künstlicher neuronaler Netzwerke, kann jedes Element mit jedem anderen Verbindungen aufbauen. Dies geschieht durch Einsatz der „Hebb'schen Lernregel", nach welcher die wechselseitige Aktivierung zweier Neuronen in dem Maße gefördert wird, in dem sie zuvor gleichzeitig aktiviert waren. Werden nun verschiedene Bilder in Form von Neuronenaktivitätsmustern wiederholt eingegeben (z.B. Gesichter), dann speichert das Netz die inputcharakteristischen Hell-Dunkel-Muster (linke Bilder). Nach dem Lernprozess kann das Netzwerk auch bei reduzierter Informationseingabe (z.B. bei einem Fragment oder einer groben Bildstörung wie hier in der Mitte) das wahrscheinlichste Bild erzeugen (Bilder rechts). Die in einem künstlichen neuronalen Netzwerk maximale Zahl fehlerfrei speicherbarer unterschiedlicher Konfigurationen liegt etwa bei 13–15 % der Neuronenanzahl. In ähnlicher Weise könnten auch Vorstellungsleistungen des Gehirns zustande kommen (Palm, 1988).

die für den Organismus beste Abbildung der Realität" (Spitzer, 1996, 58). Wie von McClelland, McNaughton und O'Reilly (1995) hervorgehoben wird, hat diese Zielsetzung starke Ähnlichkeit mit der Schätzung von Parametern (z.B. von Mittelwerten) in der Statistik (s. Kap. 3).

Abb 7.2



Jones und Hoskins (1987) trainierten ein neuronales Netzwerk mit einer Input- und einer Outputschicht auf das Verhalten von Rotkäppchen im Wald: Hat jemand große Ohren, große Augen und große Zähne („Wolf"), dann sollte es weglaufen, schreien und den Holzfäller suchen; ist hingegen jemand freundlich, runzelig und hat ebenfalls große Augen („Großmutter"), dann sollte es sich nähern, auf die Wange küssen und Essen anbieten. Wenn nun bei einem solchen künstlichen neuronalen Netzwerk zwischen Input- und Outputschicht eine weitere Schicht Neuronen („hidden layer") eingezogen und das Netzwerk in obiger Weise (ebenfalls wieder durch Fehlerrückmeldungen) trainiert wird, dann bilden sich in den Zwischenneuronen die zu lernenden Eigenschaftskonfigurationen auch begrifflich ab.

Viele dieser Leistungen biologischer Speicher können mittels *künstlicher neuronaler Netzwerke* (s. 6.4) erforscht werden, so etwa durch Simulation von **Vorstellungen** (Abb. 7.1), von **Eigenschaftsräumen** im Kortex, von Reiz-Reaktions-Prozessen (Abb. 7.2) oder von **Klassifikationsprozessen** (Abb. 7.2).

7.1.2 | Lern- und Vergessenskurven

Nach welchen Kriterien arbeitet nun ein **biologischer Speicher**, um aus dem enormen Informationsinput jene stabilen Gesetzmäßigkeiten einzuprägen, die biologische Relevanz besitzen (s. 6.2), sowie hinreichend genau und trotzdem modifizierbar sind?

Da zu speichernde Erfahrungen und Wissensinhalte umso größere Lebensrelevanz besitzen, je häufiger sie in Erlebnisfolgen vorkommen, lösen dichte Wiederholungen von Erfahrungen einen ra-

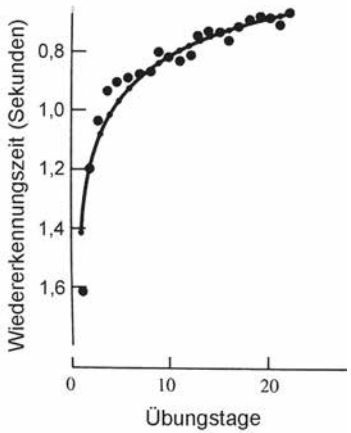


Abb 7.3

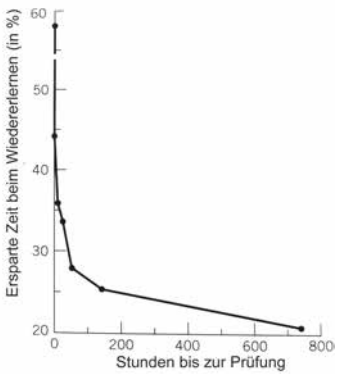
In einem Experiment von Pirolli & Anderson (1985) lernten Versuchspersonen über 25 Tage hinweg je zwei Stunden pro Tag 15 einfache Sätze, wie etwa „Der Doktor hasste den Rechtsanwalt", „Der Matrose erschoss den Friseur" usw. Überprüft wurde das Wissen täglich durch Vorgabe von korrekten Sätzen und solchen, die aus den Elementen der Lernsätze neu gebildet worden waren (z.B. „Der Doktor erschoss den Friseur"). Gemessen wurde die Zeit, wie schnell die Probandinnen und Probanden mit Knopfdruck entscheiden konnten, ob der Satz zum Lernmaterial gehörte oder nicht. Die Reaktionszeit (Z) – als inverses Lernkriterium – nahm mit der negativen Potenz der Übungstage (T) ab: $Z = 1{,}40 \cdot T^{-0{,}24}$ *(die Y-Achse ist zwecks Illustration des Lernfortschrittes umgedreht).*

schen Lernprozess aus. Tatsächlich lässt sich sowohl bei einfachen als auch bei komplexen Lernvorgängen (z.B. Silbenlernen, Addieren, mathematisches Beweisen, schriftstellerische Fertigkeit) der erzielte Lernfortschritt mittels einer positiven Potenzfunktion beschreiben (0,0 < Exponent < 1,0; Anderson, 2000). Dieses **Potenzgesetz des Lernens** („power law of learning"; Newell & Rosenbloom, 1981) besagt, dass erste Wiederholungen von gleichartigen Erfahrungen relativ schnell zur Einprägung führen und die nachfolgenden immer langsamer (Abb. 7.3).

In analoger Weise sollte nach den erwähnten Speicherprinzipien die Löschung von „statistisch unnützen" Einprägungen erfolgen, das sind solche, die nicht durch besondere Speicherfaktoren, wie zum Beispiel durch Aktivierung oder Emotionalität, gefestigt werden. Tatsächlich lässt sich bei vielen Lerninhalten auch der Vergessensprozess annähernd durch eine Potenzfunktion charakterisieren – nun aber mit negativem Exponenten (Abb. 7.4).

Dieses **Potenzgesetz des Vergessens** („power law of forgetting") bedeutet, dass eingeprägte Gedächtnisinhalte anfangs sehr rasch und dann immer langsamer vergessen werden. Der biologische Speicher hat also die Tendenz, alles wieder zu löschen, was nicht per-

Abb 7.4



Der Begründer der Gedächtnisforschung, Hermann Ebbinghaus (1850–1909), lernte dreizehn „sinnlose Silben" (z.B. DAX, TUF, NUP) auswendig, bis er sie zweimal fehlerfrei aufsagen konnte. Dann testete er, wie viel der ursprünglichen Lernzeit (in Prozenten) er sich ersparte (E), wenn er zu verschiedenen Zeitpunkten danach (S) die Liste wieder fehlerfrei lernen wollte. Diese „Vergessenskurve" lässt sich, wie viele andere auch, als Potenzfunktion darstellen: $E = 47{,}56 \cdot S^{-0{,}126}$.

Merksatz

Die Geschwindigkeit des Einprägens und Vergessens von Lerninhalten kann in vielen Fällen durch positive bzw. negative Potenzfunktionen beschrieben werden.

manent in seiner Lebensrelevanz bestätigt wird. Anderson (2000, 233) sieht die Ursache für diese schnelle Vergessensbereitschaft in der evolutionären Anpassung des Gedächtnissystems an die jeweilige „statistische Struktur der Realität". Als Indiz für diese Annahme führt er die Themen in den Headlines von Zeitungen an, deren Wahrscheinlichkeit, an einem bestimmten Tag in der Zeitung vorzukommen, sich relativ exakt über die (negative) Potenzfunktion ihres Erscheinens in vorangegangenen Zeitungsausgaben errechnen lässt.

Bei verschiedenen Beeinträchtigungen des Gehirns (z.B. durch Traumen, Infarkte, Drogenmissbrauch, Mangelernährung etc.) zeigen sich massive Gedächtnisausfälle (Pritzel et al., 2003). So etwa kann chronischer Alkoholmissbrauch zu einer Mangelernährung (Vitamin-B1-Mangel) und dieser zu einer Gehirnschädigung führen. Bei diesem sogenannten **Wernicke-Korsakoff-Syndrom** treten neben anderen Symptomen (Bewegungs-, Denk- und Sprechstörungen) relativ massive Gedächtnisausfälle bzw. **Amnesien** auf. Die Erinnerungsstörungen reichen dabei sowohl in die Vergangenheit, z.B. erkennt die Patientin oder der Patient berühmte Persönlichkeiten nicht wieder („retrograde Amnesie"), als auch in die Zukunft, sodass keine neuen Erlebnisse mehr eingeprägt werden können („anterograde Amnesie"). Allerdings betrifft diese Merkstörung nur

Amnesie: Gedächtnisverlust, der durch Verletzung, Krankheit, Drogenmissbrauch oder andere Gehirnbeeinträchtigungen hervorgerufen werden kann.

Amnesie | Box 7.1

Eine operative Entfernung der *Hippocampi* (basale Anteile des *Limbischen Systems*) hat eine (anterograde) Amnesie zur Folge. Ein bekanntes Beispiel ist der zu zweifelhafter Berühmtheit erlangte Patient H. M., dem wegen schwerer epileptischer Anfälle Teile beider Temporallappen entfernt worden waren, mit der Folge einer vollständigen und dauerhaften Einprägungsstörung – bei ansonsten erhaltender Intelligenz (Milner, 1954, 1970). Ab dem Zeitpunkt der Operation konnte sich der Patient an keine neuen Eindrücke mehr erinnern, wie etwa an den späteren Tod seines Vaters, an die neue Adresse seiner Familie oder an neue Bekanntschaften. Nach Aussage des Patienten konnte jeder gegenwärtige Moment klar erlebt werden, er sagte aber, er wisse überhaupt nichts darüber, was zuvor passiert sei, sodass er den Eindruck hätte, gerade aus einem Traum erwacht zu sein. Interessanterweise konnten aber weiterhin andere kognitive Leistungen (wie z.B. Gesellschaftsspiele) und Fertigkeiten erlernt werden, wenn auch die Situation des Erlernens selbst jeweils nicht mehr erinnert wurde (s. dazu Birbaumer & Schmidt, 1991).

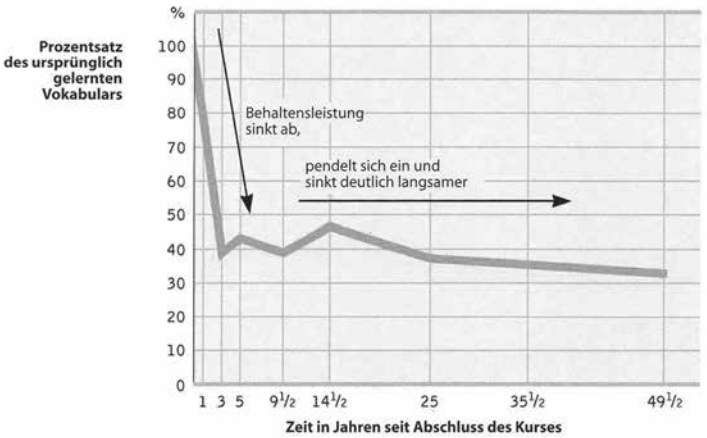
das *explizite*, nicht aber das *implizite Gedächtnis*, sodass gelernte Vokabeln zwar nicht reproduziert oder wiedererkannt werden, aber bei teilweiser Darbietung der Worte zumindest richtig ergänzt werden können (im „Wortergänzungstest“).

Vergessenstheorien | 7.1.3

In der psychologischen Fachliteratur werden für Vergessensprozesse (Abb. 7.5) im Wesentlichen drei Arten von Ursachen angeführt:

1. **Verfall**: Ebenso wie bei physikalischen Speichern mit der Zeit die Magnetisierung schwächer wird, nimmt auch in biologischen Systemen mit der Zeit die Ausprägungsstärke von Nervenver-

Abb 7.5



Bahrick (1984) überprüfte die Vokabelkenntnisse von Highschool- und Collegeabsolventinnen und -absolventen nach einem Spanischkurs über viele Jahre hinweg. Drei Jahre nach dem Kurs beherrschen Testpersonen nur mehr etwa 40 % der Vokabeln, dann aber blieb das Wortschatzniveau über viele Jahrzehnte weitgehend gleich, um 50 Jahre danach immer noch bei 30 % zu liegen.

bindungen ab. Dieser Zerfall von nervösen Gedächtnisspuren im Gehirn („Engrammen") kann als Folge des neuronalen Stoffwechsels, des Absterbens von Nervenzellen oder von unsystematischen Stör- und Wechselwirkungen im Nervensystem erklärt werden.

2. **Interferenzen**: Im Zentralnervensystem werden überaus große Mengen von Information sowohl zeitgleich (aus verschiedenen Sinnessystemen) als auch nacheinander gespeichert, sodass sich Informationsanteile überlappen und wechselseitig beeinflussen. Solche Interferenzen sind dann geringer, wenn zwischen den verschiedenen Speicherprozessen Pausen gemacht werden oder die einzelnen Lernmaterialien nicht verwechselbar sind, damit deren „postmentale" Erregungen (Rohracher, 1968) sich in Ruhe, nämlich ungestört durch weitere Speicherprozesse, *konsolidieren* können (s. 7.5.2).

interferieren: sich überlagern

3. **Abrufstörung**: Auch wenn Wissensinhalte optimal aufgenommen und gut gefestigt sind, können sie oft nicht wiedergegeben werden. Dies ist dann der Fall, wenn zwischen Abrufinhalten (z.B. der Art der Fragen) und gespeicherten Wissensinhalten keine passenden assoziativen Verknüpfungen vorhanden sind. Ein gutes Beispiel dafür ist das „Tip-of-the-tongue"-Phänomen, bei dem einem ein Wort „auf der Zunge liegt". Am leichtesten gelingt der Abruf aus dem Gedächtnis dann, wenn der **Kontexteffekt** stark ist, das heißt, wenn zwischen der Situation des Einprägens und jener der Prüfung eine möglichst große Übereinstimmung der psychischen und physischen Merkmale gegeben ist (Abb. 7.6).

Merksatz

Vergessen wird im Wesentlichen auf drei Ursachen zurückgeführt: 1. Zerfall des nervalen Engramms, 2. Interferenzen zwischen Speicherungen, 3. Abrufstörung.

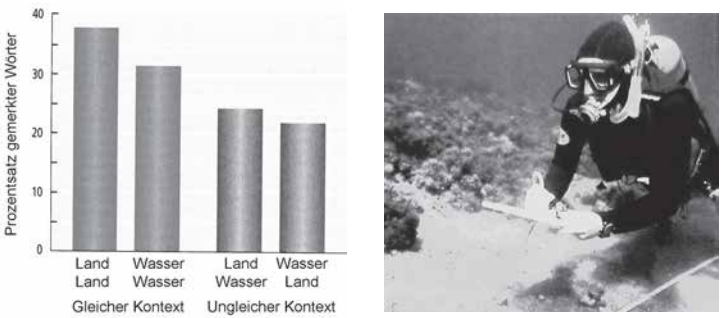


Abb 7.6

Im Experiment von Godden und Baddeley (1975) hatten Taucher Listen von Worten sowohl an Land als auch im Wasser zu lernen. Geprüft wurden sie im jeweils gleichen oder in verschiedenem Kontext. Die Reproduktionsleistungen waren bei Übereinstimmung zwischen Lern- und Prüfsituation um 10–12 % der Gesamtlernleistung besser als bei Nichtübereinstimmung.

7.2 | Kurzzeitspeicherung

Da in der psychologischen Gedächtnisforschung der Behaltensdauer von Merkinhalten besondere Aufmerksamkeit gewidmet wurde, hat sich in psychologischen Einführungswerken die Unterscheidung in *Ultrakurzzeitgedächnis*, *Kurzzeitgedächnis* und *Langzeitgedächtnis* eingebürgert, obwohl dadurch der falsche Eindruck entstehen kann, dass es sich um drei voneinander abgegrenzte Strukturen im Gehirn handle.

Abb 7.7 |

Bei etwa zehn Prozent vorwiegend jüngerer Personen (meist Kinder) hat das ikonische Gedächtnis eine Dauer von vielen Sekunden. In diesem Fall spricht man von einem eidetischen oder „fotografischen" Gedächtnis. Eine Person mit dieser Begabung betrachtet das obige Beispielbild eine halbe Minute lang und kann danach aus der Vorstellung die meisten der dargestellten Figuren wahrnehmungsgleich beschreiben.

Ultrakurzzeitgedächtnis (UKZG) – Sensorisches Gedächtnis (SG)

7.2.1

Wenn wir intensiven visuellen oder akustischen Wahrnehmungen ausgesetzt sind, kommt es danach häufig noch zu Nachbildern bzw. Nachklängen. Dies lässt erkennen, dass die aufgenommene Information in den primären sensorischen Gehirnarealen kurzfristig gespeichert bleibt, sie wird aber zumeist durch die nachfolgenden Wahrnehmungen überdeckt („maskiert") und somit schnell wieder gelöscht.

Forschungen zum ultrakurzen, sensorischen Gedächtnis beschränken sich weitgehend auf den visuellen und den akustischen Sinn. Experimente von Sperling (1960) zum (visuellen) **ikonischen Gedächtnis** zeigen, dass dessen Speicherkapazität zwar relativ groß ist, seine Dauer aber bei den meisten Menschen nicht mehr als 0,5 Sekunden beträgt. Vom (akustischen) **echoischen Gedächtnis** nimmt man an, dass es bis in den Sekundenbereich reicht, sodass es sich nur unscharf vom stabileren *Kurzzeitgedächtnis* abgrenzen lässt. Die Sonderbegabung weniger Personen, optische Wahrnehmungseindrücke auch noch nach vielen Sekunden detailreich aus der Vorstellung wiedergeben zu können, nennt man **Eidetik** oder **fotografisches Gedächtnis** (Abb. 7.7). Untersuchungen aus der Frühzeit der Psychologie lassen vermuten, dass diese Fähigkeit bei Kleinkindern stärker ausgeprägt ist als bei Erwachsenen und mit der Zunahme des begrifflichen Denkens verloren geht.

Merksatz

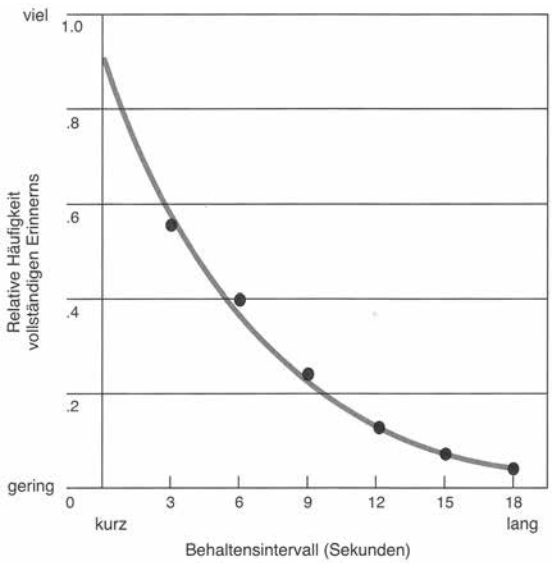
Das sensorische Gedächtnis oder Ultrakurzzeitgedächtnis hat zwar eine hohe Speicherkapazität, vermag aber im Allgemeinen die Speicherinhalte nur maximal eine Sekunde zu behalten.

Kurzzeitgedächtnis (KZG)

7.2.2

Die spontane Aufrechterhaltung von Bewusstseinsinhalten über eine Zeitspanne von bis zu zehn Sekunden (Abb. 7.8) wird der Leistung des Kurzzeitgedächtnisses zugerechnet. Dies ist zumeist auch jene Zeitspanne, innerhalb derer Bewusstseinsinhalte noch als „gegenwärtig" erlebt werden („Präsenzzeit"). Was die genaue zeitliche Erstreckung des Kurzzeitgedächtnisses betrifft, so variieren die

Abb 7.8



Im Experiment von Peterson und Peterson (1959) wurden Probandinnen und Probanden einzelne Dreierkombinationen von Konsonanten zum Merken dargeboten (z.B. FCV, RNL) und danach zu verschiedenen Zeitpunkten geprüft. Damit die Versuchspersonen die Tripletts in der Zwischenzeit nicht wiederholen konnten, mussten sie währenddessen dreistellige Zahlen nach rückwärts zählen („Distraktoraufgabe"). Bis zu einer Dauer von vier Sekunden konnten 50 % der bedeutungslosen Tripletts noch vollständig erinnert werden.

Angaben darüber in der psychologischen Literatur beträchtlich je nach Art der Experimente und Art der Merkinhalte (Silben, Wörter, Bilder, ...), nämlich von wenigen Sekunden bis zu einer halben Minute (vgl. Herkner, 1986).

Der Speicherumfang des Kurzzeitgedächtnisses kann mit etwa fünf bis neun Informationseinheiten (Miller, 1956: „magical number 7 ± 2") angegeben werden, wobei diese Einheiten sowohl elementare Inhalte (z.B. Ziffern, Objekte) als auch sogenannte **Chunks** sein können, d.h. assoziativ verbundene Konfigurationen von Einzelheiten. So etwa merkt man sich eine Auto- oder Telefonnummer etwas leichter, wenn man sie nicht als lange Ziffernfolge, sondern als Folge von zwei- oder dreistelligen Zahlen genannt bekommt (z.B. 478 – 23). Durch Bildung von Chunks (z.B. bei Zahlen: 222 – 567 – 369), für deren Zusammensetzung beliebige Gesetzmäßigkeiten gefunden werden können (Ähnlichkeit, Nähe, Zahlenrelatio-

engl. chunk: Brocken

nen, ...), ergeben sich wesentliche Einprägungserleichterungen, und zwar auch dann, wenn die einzuprägende Kombination früher bereits in anderem Zusammenhang verknüpft war (z.B. „Nine Eleven“ als Datum des Terroranschlags auf das World Trade Center).

Das KZG ist im akustischen Wahrnehmungsmodus länger ausgedehnt als im visuellen, geruchlichen oder geschmacklichen Bereich, aber es gibt von Person zu Person, je nach Begabung, Vortraining, Konzentration etc., erhebliche individuelle Unterschiede, zum Beispiel bezüglich der Merkfähigkeit für Melodien.

Wie Baddeley und Mitarbeiter (1975) nachweisen konnten, hängt die Behaltensleistung des Kurzzeitgedächtnisses nicht nur von der Menge an Informationseinheiten (*chunks*), sondern auch von deren Benennungslänge ab. Je länger die Symbolketten bzw. Wörter sind, mit denen die Informationseinheiten charakterisiert sind, desto weniger können davon gemerkt werden. Vier Informationseinheiten, deren Benennung insgesamt nicht länger als etwa 1,5 Sekunden dauert, können meist zu 80 % wiedergegeben werden (Box. 7.2).

Merksatz

Die Speicherkapazität im Kurzzeitgedächtnis wird auf etwa fünf bis neun Informationseinheiten geschätzt. Wenn kein Memorieren (Wiederholen) der Merkinhalte stattfindet, sind die Inhalte des Kurzzeitgedächtnisses nach etwa fünf Sekunden zu 50 % und nach etwa 20 Sekunden zu 100 % vergessen.

Wortlängeneffekt | Box 7.2

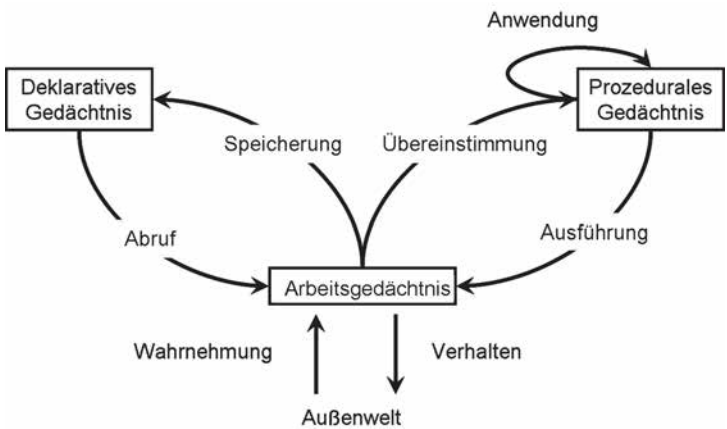
Da die Kapazität des Kurzzeitgedächtnisses begrenzt ist, wirkt sich auch die Länge der Wörter oder der Benennungen für Lerneinheiten auf die (kurzfristige) Speicherleistung aus: Den Versuchspersonen im Experiment von Baddeley und Mitarbeitern (1975) wurden in verschiedenen Durchgängen jeweils fünf kurze Wörter einmalig dargeboten (z.B. Tschad, Burma, Laos, Kuba, Malta), wovon sie sich sofort danach noch durchschnittlich 4,17 Worte merkten. Waren es hingegen längere Worte (z.B. Griechenland, Nicaragua, Afghanistan, Niederlande, Großbritannien), konnten nur mehr 2,8 Wörter im Mittel behalten werden.

7.2.3 | Arbeitsgedächtnis

In neuerer Zeit wird anstelle des Begriffs Kurzzeitgedächtnis eher das weitere Konzept **Arbeitsgedächtnis** (Computermetapher: Arbeitsspeicher) bevorzugt, in dem Informationen aus den Bereichen Wahrnehmung, Erinnerung, Emotion und Motivation zusammenfließen und entsprechend den aktuellen Handlungsanforderungen integrativ verarbeitet werden (Abb. 7.9). Baddeley (1986) formulierte drei Komponenten des Arbeitsgedächtnisses, nämlich eine sogenannte **zentrale Exekutive**, die für die Kontrolle der Aufmerksamkeit (d.h. der selektiven Aktivierung bestimmter Bewusstseinsinhalte) verantwortlich ist, eine **phonologische Schleife** und einen **visuell-räumlichen Notizblock**, die Funktionen des akustischen bzw. visuellen Memorierens erfüllen (d.h. der absichtlichen Aufrechterhaltung von Bewusstseinsinhalten im Arbeitsgedächtnis).

Die Leistungen des Arbeitsgedächtnisses werden im Wesentlichen im Frontallappen des Gehirns lokalisiert. Zugrunde liegt die

Abb 7.9 |



Das Arbeitsgedächtnis (AG) hat nach Anderson (1983a) folgende Funktionen: Es übernimmt Daten aus der Außenwelt, interpretiert diese durch Abruf von Begriffen, Fakten und Episoden aus dem deklarativen Gedächtnis und übergibt diesem erneut Daten zur Speicherung. Wenn Informationsmuster im Arbeitsgedächtnis mit Auslösern der im prozeduralen Gedächtnis gespeicherten Fertigkeiten übereinstimmen, dann werden Ausführungsimpulse an das AG rückgemeldet und dieses entscheidet, ob Verhaltensweisen ausgelöst werden. Zusätzlich können innerhalb des prozeduralen Gedächtnisses durch Verkettungen von Abläufen (Anwendungen) komplexere Wenn-dann-Produktionen erzeugt werden.

Beobachtung, dass bei Verletzungen des Frontallappens oder auch bei dessen mangelhafter Ausreifung (bei Kindern bis zum Alter eines halben Jahres) vorgegebene Objekte, die nach kurzer Betrachtung mehrere Sekunden verdeckt werden (z.B. wenn eine Spielzeugeisenbahn in einem Tunnel verschwindet), nicht am richtigen Ort erwartet werden können (mangelnde „Objektpermanenz“).

Wie schon erwähnt (Kap. 6.3), geht man inzwischen davon aus, dass das Zentralnervensystem eher einem stark verbundenen Gesamtnetzwerk gleicht („Modell des globalen Arbeitsspeichers“) als einer Ansammlung von autonomen Modulen mit streng getrennten Funktionszuweisungen. Das Ausmaß der neuronalen Aktivierung in einem Unterbereich des Gesamtnetzwerkes gibt die momentane mentale **Konzentration** auf die entsprechenden Wahrnehmungen, Gedanken, Erinnerungen oder Handlungen wieder. Aus neuropsychologischer Sicht kann daher der Arbeitsspeicher als jener fluktuierende Teilbereich des gesamten zentralen Speichernetzwerkes aufgefasst werden, dessen Neuronenaggregate zum jeweiligen Zeitpunkt gerade am stärksten aktiviert sind (s. auch D’Esposito & Postle, 2015).

Merksatz

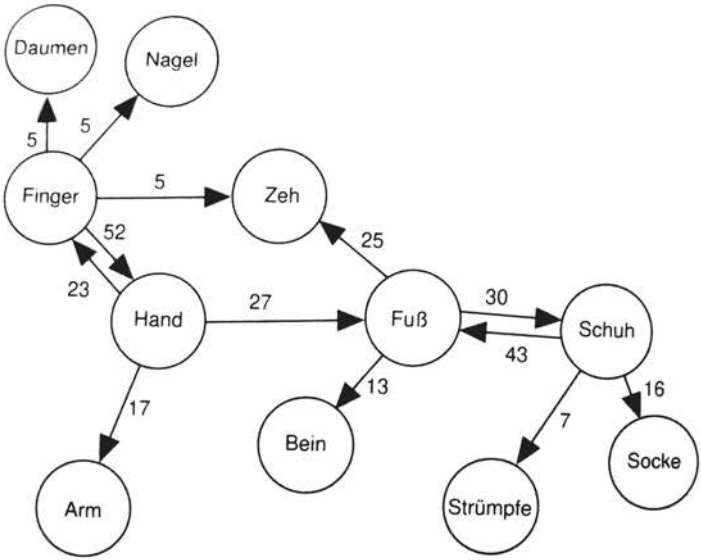
Das im Vergleich zum Kurzzeitgedächtnis theoretisch weitere Konzept des Arbeitsgedächtnisses schließt neben einer zentralen Exekutive (Aufmerksamkeitssteuerung) noch zwei Untersysteme zur Zwischenspeicherung akustischer und visueller Inhalte mit ein.

Langzeitspeicherung 7.3

Das **Langzeitgedächtnis** (LZG) enthält dauerhaft gespeicherte Erfahrungen und Erlebnisse, die nicht mehr im Kurzzeitgedächtnis (KZG) präsent sind bzw. aus diesem nicht mehr abrufbar sind. Darunter fallen im Allgemeinen jene Bewusstseinsinhalte, die länger als zehn Sekunden zurückliegen und nicht bewusst durch Wiederholen bzw. „Memorieren“ aufrechterhalten werden.

Bei den Speicherinhalten des LZG handelt es sich um eine nahezu unendliche Vielfalt von Informationen aus den Bereichen der Wahrnehmung, des Denkens und des Verhaltens: Gesichter, Tiere, Landschaftsformen, physikalische und chemische Gesetzmäßigkeiten, Lautgestalten und Schriftbilder von Worten und Sprachen, Melodien und Musikstücke, Gedichte, Sachverhalte und Erlebnisse, Begriffsinhalte, mathematische Algorithmen, berufliche Fertigkei-

Abb 7.10



In einer Untersuchung an etwa tausend Studentinnen und Studenten ermittelten Palermo und Jenkins (1964) die durchschnittliche Assoziationshäufigkeit zwischen verschiedenen Begriffen. Das resultierende semantische Netzwerk illustriert die (mittleren) Wahrscheinlichkeiten, mit denen in freier Assoziation (ohne Handlungsdruck) ein bestimmter Begriff (z.B. Finger) einen anderen (z.B. Hand) ins Bewusstsein ruft (z.B. p = 0,52).

ten und viele andere perzeptive, geistige oder motorische Zusammenhänge. Aufgrund neuropsychologischer Forschungen darf man annehmen, dass ein Großteil der psychischen Abbilder der inneren und äußeren Realität des Menschen in den Neuronennetzwerken des *Großhirns* (Neokortex), des *Kleinhirns* und des *Zwischenhirns* niedergelegt ist (s. 6.3).

Merksatz

Unter Langzeitgedächtnis versteht man die Gesamtheit an dauerhaft gespeicherten Erfahrungen im Zentralnervensystem.

Die Speicherinhalte sind miteinander vernetzt und beeinflussen sich gegenseitig in vielfacher Weise (Abb. 7.10). So etwa kann an bereits gespeicherten Begriffsstrukturen durch Umlernprozesse eine Veränderung in den Ähnlichkeitsrelationen nachgewiesen werden (Maderthaner & Kirchler, 1982) oder durch Anreicherung von Begriffen anhand von immer mehr Assoziationen („Fächer-Effekt“,

Anderson, 1996) ein Verarbeitungskonflikt erzeugt werden (*Interferenz*), der die Wiedergabe von Einprägungen beeinträchtigt. Bei Stimulation des neuronalen Gedächtnisnetzwerks durch konkrete Wahrnehmungen, Vorstellungen oder andere Kognitionen entstehen spezifische neuronale Aktivierungsmuster („Erregungskonstellationen", Rohracher, 1965, 57), die sich in Richtung der stärksten assoziativen Verbindungen ausbreiten („Theorie der Aktivierungsausbreitung", Collins & Loftus, 1975; Abb. 7.10). Auf Basis solcher Netzwerke entsteht ein dynamisches „inneres Modell" der Realität, aufgrund dessen für jeden Zeitpunkt antizipiert wird, welche Veränderungen in der Wahrnehmungswelt zu erwarten sind. Empirische Hinweise auf solche Voraktivierungen im psychischen System lieferten Experimente zum **Priming**, die zeigen, dass auf beliebige

engl. prime: vorbereiten, fertig machen

Priming | Box 7.3

Als „assoziatives Priming" bezeichnet man das Phänomen, dass Begriffe, die miteinander assoziiert sind, einander wechselseitig in der Reproduktion fördern, sofern einer davon zeitlich vor dem anderen wahrgenommen oder gedacht wurde. So etwa beschleunigt sich das Klassifikationsurteil, ob ein Wort der deutschen Sprache angehört oder nicht, wenn vorher (oder gleichzeitig) mit diesem Wort ein mit ihm häufig assoziiertes Wort dargeboten wird. Von den nachfolgend dargebotenen Wortpaaren aus dem Experiment von Meyer und Schvaneveldt (1971) steht das eine in starker und das zweite in schwacher assoziativer Verbindung. Die unter den Wortpaaren angeführte (durchschnittliche) Reaktionszeit lässt erkennen, wie rasch Versuchspersonen entscheiden können, ob es sich bei den beiden Wörtern um deutsche Vokabeln handelt: Ein Wort kann dann schneller beurteilt werden, wenn es bereits vorher durch ein assoziiertes Wort semantisch voraktiviert wurde.

	starke Assoziation	geringe Assoziation
1. Wort	Brot	Krankenschwester
2. Wort	Butter	Butter
Reaktionszeit	0,86 s	0,94 s

Reize (Bilder, Worte, ...) schneller richtig reagiert werden kann, wenn vorher an ähnliche Reize oder an mit ihnen assoziierte Inhalte gedacht wurde (Box 7.3; Kiefer, 2005).

7.4 | Komponenten des Langzeitgedächtnisses

Derzeit existieren einige Konzepte zum Langzeitgedächtnis, die unterschiedliche Einteilungen treffen, je nachdem, ob sie sich mehr an der Art der Inhalte oder an deren Entstehung orientieren (Tab. 7.1). So unterscheidet man häufig **episodische** Inhalte, welche sich auf konkrete Erlebnisse und deren Aufeinanderfolge beziehen, von **semantischen** Inhalten, die Abstraktionen der Erlebniswelt repräsentieren, nämlich Merkmalsdimensionen, Begriffe und Klassifikationsschemata, was auch als kontextfreies Faktenwissen bezeichnet wird (O'Reilly et al., 2011). Episodische Erinnerungen sind zumeist analog bzw. **imaginativ**, also in Form von Vorstellungen gespeichert, während semantische Inhalte eher **propositionalen** Charakter haben, d.h. in Aussageform gespeichert sind. Als **deklarative** Gedächtnisinhalte gelten (episodische) Erinnerungen und (semantisches) Wissen über Sachverhalte bzw. Fakten („Gewusst was"), während **prozedurale** (nicht-deklarative) Einprägungen jene geistigen oder motorischen Operationen zusammenfassen, die zu Veränderungen der inneren oder äußeren Realität führen („Gewusst wie"; s. auch Abb. 7.9). Schließlich wird oft **explizites** von **implizitem** Gedächtnis unterschieden, wobei im einen Fall ein bewusster Zugang zu den Wissensinhalten besteht, sie also formulierbar und erinnerbar sind, während sie sich im anderen Fall nur im Denken oder Verhalten manifestieren (Walla et al., 2001; Ray, Bates & Bly, 2004). Da es allerdings möglich ist, dass Wissen sowohl explizit wiedergegeben werden kann, als auch sich in Verhaltensänderungen niederschlägt (z.B. bei *Priming*, *Konditionierungen*), sollte eher von „expliziter" oder „impliziter Testung" als von unterschiedlichen Gedächtnisformen gesprochen werden.

Eine spezielle Variante des Langzeitgedächtnisses ist das **Ultralangzeitgedächtnis** (ULZG, „very long-term memory"), welches vorwiegend anhand von Tagebuchuntersuchungen erforscht wurde und sich oft auf Sprachkenntnisse (s. Abb. 7.5) oder auf autobiografische Inhalte bezieht (Abb. 7.11). Solche Untersuchungen zeigen, dass langjährige Gedächtnisleistungen im Wesentlichen von geeig-

Auswahl an verbreiteten Typisierungen des Langzeitgedächtnisses, bei denen entweder von der Art gespeicherter Inhalte oder ihres Zustandekommens ausgegangen wird. **Tab 7.1**

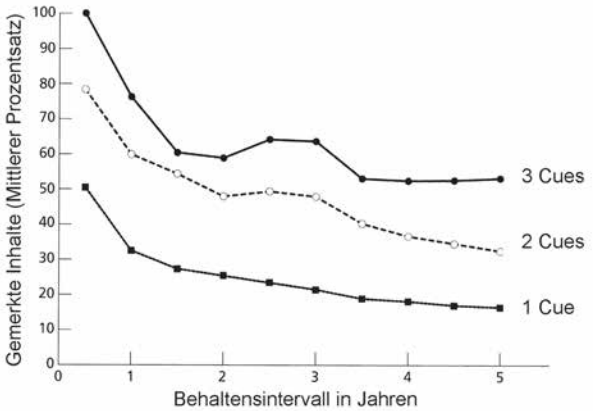
episodisch (räumlich und zeitlich definierte Erlebnisse)	semantisch (Begriffe, Klassifikationen, Begriffsnetzwerke)	Tulving (1972, 2002)
deklarativ (Sachverhalte, Wissen)	prozedural (Reiz-Reaktions-Folgen, Fertigkeiten)	Anderson (1976)
epistemisch (bildhaft)	heuristisch (aufgabenbezogen)	Dörner (1976)
propositional (aussagenlogisch, prädikatenlogisch)	analog/imaginativ (Vorstellungen, kognitive Landkarten)	Kosslyn, Ball & Reiser (1978), Paivio (1971)
implizit (automatisch Gelerntes)	explizit (absichtsvoll Gelerntes)	Graf & Schacter (1985)
deklarativ (Fakten, Sachverhalte)	nicht-deklarativ (Abläufe, Aktionen)	Squire (1987)

neten Erinnerungshilfen abhängen (s. auch Abb. 7.22) und dass 20–30 % konkreter Inhalte (z.B. Erlebnisse) oder lebensnah trainierter Sprachinhalte (z.B. Vokabeln) auch über Jahrzehnte erhalten bleiben können.

In der modernen Gedächtnisforschung strebt man einerseits nach einer Integration aller experimentell bestätigten Gedächtnisformen (Abb. 7.12) und versucht andererseits die Gehirnareale ausfindig zu machen, die für die einzelnen Gedächtniskomponenten zuständig sind (s. etwa Squire, 1987; Pritzel et al., 2003). Die Repräsentation von Gestalten und Schemata wird im Wesentlichen den unimodalen und polymodalen Projektionsfeldern des Kortex zugeschrieben. Episodische Inhalte werden überwiegend in den Assoziationszentren der rechten, semantische Inhalte überwiegend in jenen der linken *Gehirnhemisphäre* mit wesentlicher Beteiligung des Hippocampus vermutet, während das Limbische System eher für den Aufbau des Gedächtnisses und der frontale und temporale Kortex eher für den Abruf wichtig zu sein scheint. Die Basalganglien (Striatum, Pallidum), die prämotorische Region im *Vorderhirn* und das *Kleinhirn* dürften sowohl den Aufbau, die Konsolidierung und den Abruf von geistigen und motorischen Fertigkeiten steuern (s. auch 6.2).

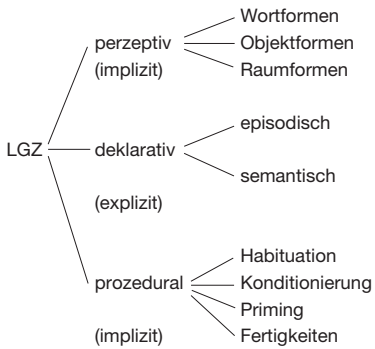
Unimodale Projektionsfelder sind nur für eine Sinnesart (visuell, akustisch, ...) zuständig, polymodale für mehrere.

Abb 7.11



Über einen Zeitraum von fünf Jahren notierte Wagenaar (1986) über 2000 Ereignisse in seinem Leben und vermerkte dazu, wer dabei war, wo es stattfand, wann es war, wie selten es war, den Emotionalitätsgrad usw. In halbjährlichen Zeitabständen zog er nach dem Zufall eine Ereignisbeschreibung und versuchte sich aufgrund einzelner Merkmale („cues") an die Gesamtsituation (andere „cues") zu erinnern. Nach fünf Jahren lag die Erinnerungsleistung mit einer Gedächtnishilfe bei 20 %, unter Zuhilfenahme von drei „cues" immerhin bei 60 %.

Abb 7.12



Gliederung der verschiedenen, in der psychologischen Fachliteratur vorgeschlagenen Typisierungen von Strukturkomponenten des Langzeitgedächtnisses. Im perzeptiven Gedächtnis vermutet man Gestalt- und Strukturinformationen über die Wahrnehmungswelt, im deklarativen Gedächtnis die Speicherung von Erlebnissen, Fakten und Begriffen, und im prozeduralen Gedächtnis alle Arten kognitiver oder verhaltensbezogener Programme, wie zum Beispiel erworbene Reiz-Reaktions-Muster und angelernte Denk- und Verhaltensketten.

Stadien der Gedächtnisbildung 7.5

Die Leistung des Langzeitgedächtnisses hängt davon ab, wie die zu speichernde Information aufgenommen wurde, ob und wie sie sich in gespeicherter Form verändert und wie sie wieder aus dem Speichermedium zurückgewonnen werden kann. Hinsichtlich dieser drei Phasen der Gedächtnisbildung konnten in der psychologischen Forschung viele verschiedene systematische Einflüsse bzw. Effekte nachgewiesen werden.

Aufnahme von Wissen (Enkodierung, Akquisition) 7.5.1

Entscheidend für die Gedächtnisleistung sind die Form, in welcher Informationen dem Gedächtnis zugeführt werden (z.B. als Melodien, sprachliche Inhalte, Vorstellungen, Gedanken usw.), und die Bedingungen, unter denen dies geschieht (z.B. Aufmerksamkeit, Bedürfnislage, Kontext, Gliederung der Inhalte, Reihenfolge der Einprägung usw.). Wie erwähnt, geht jeder Enkodierung von Speicherinformation bereits eine Filterung voraus, die durch Gestaltbildung, Aufmerksamkeitsausrichtung und Begriffskategorisierung zustande kommt, wobei auch irrelevant wirkende Details Einfluss ausüben, wie etwa die jeweilige Stimmung, der räumliche Kontext, verbale Kommentare oder Ähnliches (Abb. 7.13).

Merksatz

Der Erwerb und die Nutzung von Wissen werden durch die Form und die Bedingungen der Informationsaufnahme, mögliche Veränderungen im Speicher und durch die Art ihres Abrufs bestimmt.

Für die Aufnahme von Gedächtnismaterial in das Langzeitgedächtnis können zumindest fünf relevante Effekte unterschieden werden:

Arousal-Effekt: Wie schon im Abschnitt Aktivierung und Lernen (6.2) dargestellt, werden Lern- und Einprägungsprozesse sowohl durch ein mittleres (tonisches) Aktivierungsniveau als auch durch Aktivierungsschwankungen bei der Informationsaufnahme gefördert (phasische Aktivierung). Insbesondere steigt die Wahrscheinlichkeit für die Speicherung von Bewusstseinsinhalten, wenn sie eine psychische Stimulation oder Aktivierung auslösen (Interesse, Aufmerksamkeit, Emotion, Motivation; Storbeck & Clore, 2008; McGaugh, 2018) und danach positive Konsequenzen (Aha-Erlebnis,

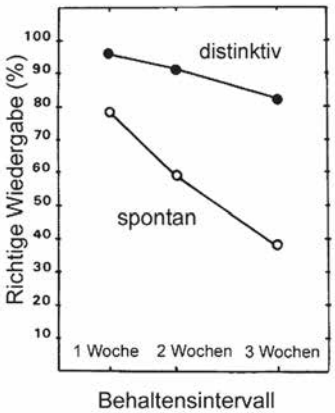
Abb 7.13

In einem Experiment von Carmichael, Hogan und Walter (1932) wurden 86 Studenten 12 Figuren vorgegeben, mit der Aufgabe, sie nachher aus dem Gedächtnis nachzuzeichnen, wobei jedoch die gleichen Figuren mit unterschiedlichen Wörtern kommentiert wurden, wie z.B. „Diese Figur ähnelt einer Brille" (Wortliste 1) oder „... ähnelt Hanteln" (Wortliste 2). Die danach produzierten Zeichnungen hatten zu etwa 75 % eine Ähnlichkeit mit den Wortbedeutungen, im Vergleich zu 45 % bei einer Kontrollgruppe ohne Begleitwörter.

Entspannung, Verstärkung, Lernpause) im Sinne einer zentralnervösen Desaktivierung nachfolgen (Kleinsmith & Kaplan, 1964). Dass manchmal bei aktivierenden Merkinhalten unmittelbar nach dem Lernen schlechtere Merkleistungen zustande kommen, wird mit dem Andauern eines intensiven Konsolidierungsprozesses erklärt (s. nächster Abschnitt), der offenbar einen sofortigen Zugriff auf die Speicherinhalte erschwert („perseverative consolidation").

Distinctiveness-Effekt: Die Originalität (Einmaligkeit, Besonderheit, Eigentümlichkeit, Exklusivität) von Speicherinhalten bzw. deren Unähnlichkeit zu anderen Speicherinhalten ist eine weitere Einprägungshilfe. Je markanter das Eigenschaftsprofil von Informationseinheiten subjektiv hervortritt, desto klarer (weniger verwechselbar) prägt sich dessen Inhalt im Gedächtnis ein. Die akzentuierte Verarbeitung bzw. „distinktive Enkodierung" eines Wortes, eines Bildes oder einer Aussage kann etwa dadurch gesteigert werden, dass unverwechselbare Charakteristika für die Merkinhalte herausgefunden werden müssen (Abb. 7.14). Wenn Versuchspersonen selbst jene Charakteristika erarbeiten, die sie für die jeweiligen Lerneinheiten als besonders typisch ansehen (und diese auch als Wiedergabehilfen einsetzen), sind die Merkleistungen wesentlich besser, als wenn Merkhilfen anderer Personen verwendet werden.

Abb 7.14



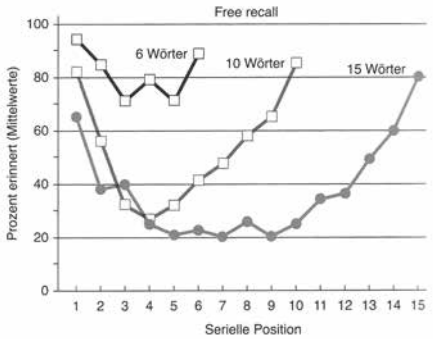
Die Merkleistung bei Wissensinhalten lässt sich erheblich steigern, wenn zu ihrer Kennzeichnung und später für ihren Abruf aus dem Gedächtnis distinkte, d.h. unverwechselbare Charakteristika verfügbar sind. Mäntylä und Nilsson (1988) ließen 24 Probanden für 30 Substantive (z.B. Admiral, Ballett, Zirkel) spontan drei charakteristische Beschreibungsmerkmale ausdenken, die ihnen (was sie vorher nicht wussten) nach einer, drei oder sechs Wochen als Erinnerungshilfen für die Abprüfung der Substantive vorgegeben wurden. Weitere 24 Probanden hatten die gleiche Aufgabe, wurden aber darauf fokussiert, drei distinktive, d.h. zu anderen Begriffen möglichst unterschiedliche Charakteristika zu finden. Die Gruppe mit distinktiven „recall cues" schnitt um 15–45 % besser ab als jene mit spontan entwickelten Merkhilfen. Sogar nach sechs Wochen konnten noch mehr als 80 % des distinktiv enkodierten Materials erinnert werden.

So etwa können unter Nutzung von drei selbst generierten distinkten Beschreibungsmerkmalen mehr als 500 Wörter unmittelbar nach dem Lernen zu mehr als 90 % und sieben Tage danach immer noch zu 65 % korrekt wiedergegeben werden. Bei Verwendung von Beschreibungsmerkmalen anderer Personen sind es hingegen nur 55 bzw. 43 % (Mäntylä, 1986).

Positionseffekt: Eine frühe experimentelle Entdeckung war auch die Bedeutung der Reihenfolge aufgenommener Informationen. Der sogenannte **Primacy-Effekt** besteht darin, dass Wörter, Bilder oder Gedanken, die nach einer Pause in der Informationsaufnahme als Erste ins Bewusstsein kommen, eine größere Chance haben, langzeitlich gespeichert zu werden, als solche inmitten anderer Inhalte. Der **Recency-Effekt** hingegen beschreibt die bessere Wiedergabe von Lerninhalten, die am Ende einer Lernserie stehen und sofort danach geprüft werden, gewissermaßen lediglich als Auswirkung des noch andauernden Kurzzeitgedächtnisses (Abb. 7.15).

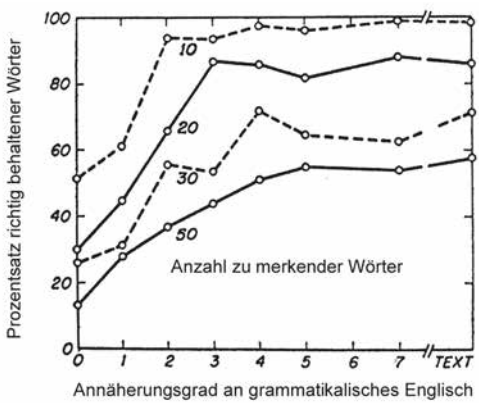
Gliederungseffekt: Empirisch gut bestätigt ist auch die plausible Regel, dass Lernmaterial dann besser behalten wird, wenn es sprachlich, begrifflich oder assoziativ klar geordnet ist. Je besser Merkinhalte sprachlich formuliert sind (Abb. 7.16), oberbegrifflich kategorisiert werden können oder assoziativ miteinander zusammenhängen, desto leichter lassen sie sich speichern, wieder auffinden oder logisch rekonstruieren. Besonders effektiv sind des-

Abb 7.15



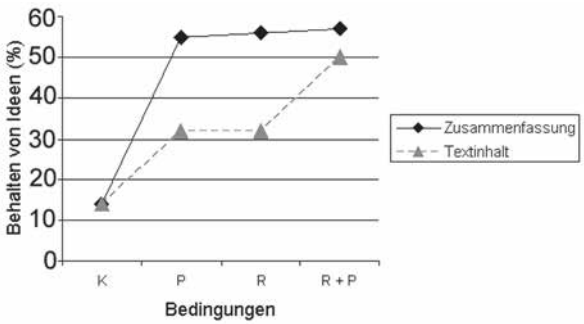
Im Experiment von Jahnke (1965) wurden 48 Studierenden unterschiedlich lange Wortlisten einmalig präsentiert, mit der Aufgabe, danach alle Worte, die erinnert werden konnten, wiederzugeben. Es wurde immer festgehalten, an welcher Position die reproduzierten Worte in der Liste standen: Sowohl die ersten als auch die letzten Elemente in den Serien konnten bis zu viermal besser gemerkt werden als die mittleren Elemente.

Abb 7.16



Miller und Selfridge (1950) gaben ihren Versuchspersonen verschieden lange Wortketten (10–50 Wörter) zum Lernen vor, deren grammatikalischer Ordnungsgrad kontinuierlich abgestuft war. Beginnend mit zufälliger Auswahl von Wörtern aus dem Englischen (Grad 0), über zufällige Wortpaare (Grad 1), Worttripel (Grad 2), Wortquartupel (Grad 3) usw. bis zu korrekt formulierten Sätzen (Text). Die Reproduktionsleistung nahm erwartungsgemäß von 10-Wort-Ketten bis zu 50-Wort-Ketten ab, aber innerhalb der gleich langen Wortketten nahm die Merkfähigkeit mit dem Grad der Annäherung an die grammatikalische Struktur der Sprache um 40–50 % zu.

Abb 7.17



Lernstoff beim vorliegenden Experiment (Glover, Bruning & Plake, 1982) war eine Abhandlung (800 Worte) über das Sonnensystem, die sich in 16 Absätze gliederte, von denen jeder durch einen Schlusssatz zusammengefasst wurde. Die im Gesamttext und in der Zusammenfassung enthaltenen Wissenseinheiten („Ideen") wurden auf eine bis zwei begrenzt und dienten bei der Wissensabfrage als Kriterium für die Einschätzung der Behaltensleistung (Zusammenfassung/Text). Verglichen mit einer Kontrollgruppe (K, Bloßes Lesen und Lernen) schnitt die Gruppe, die den zusammenfassenden Satz in eigenen Worten zu wiederholen hatte (P, Paraphrasieren), oder jene Gruppe, die den absichtlich orthografisch falsch geschriebenen Schlusssatz zu korrigieren hatte (R, Redigieren), sowohl beim Merken der Zusammenfassung (ca. 55 %) als auch beim Merken des Textes (ca. 30 %) besser ab. Der Textinhalt wurde aber dann am besten behalten, wenn beide elaborativen Vorgangsweisen zur Anwendung kamen.

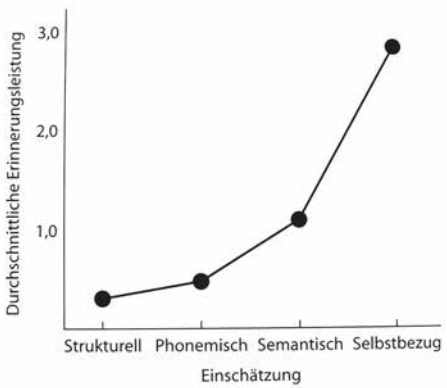
halb auch Lernstrategien bzw. **Mnemotechniken**, bei denen ursprünglich isolierte Lerneinheiten in Reimform gebracht, in eine Geschichte eingebaut (z.B. Bower, Clark, Lesgold & Winzenz, 1969) oder in der bildlichen Vorstellung assoziativ miteinander verbunden werden. Auf die Vorteile des *Imagery-Effekts* als besondere Form der assoziativen Gliederung wird nachfolgend noch genauer eingegangen.

Elaborationseffekt: Das bloße Wiederholen von Wissensinhalten, ohne darüber intensiver nachzudenken, ist als Lerntechnik heute veraltet. Craik und Lockhart (1972) heben die Bedeutung der „Tiefe der Verarbeitung" („levels of processing theory") für die Speicherung hervor und verstehen darunter das Reflektieren über die Bedeutung der Inhalte und über deren Anbindung an bereits aufgenommenes Wissen. Je mehr der gesamte Bedeutungsumfang neuer Information bewusst wird, je lebendiger die davon ausgelösten Vorstellungen sind und je mehr Assoziationen zu anderen Wissensinhalten geweckt werden, desto größer ist die Chance, dass diese Informatio-

nen langzeitlich abgelegt werden. Künstlich kann man die Elaboration des Stoffes in Experimenten etwa dadurch steigern, dass absichtlich falsch geschriebener Text korrigiert werden muss oder dass ersucht wird, eigene Formulierungen für die aufzunehmenden Wissensinhalte zu bilden (*Paraphrasieren*; Abb. 7.17). Eine spezielle Bedeutung hat dabei auch der Bezug zur eigenen Person. Je mehr Assoziationen die lernende Person von den Wissenseinheiten zu ihren eigenen Erlebnissen und Erfahrungen herstellen kann, desto größer ist der **Selbstbezugseffekt** und damit die Einprägungswahrscheinlichkeit für das jeweilige Lernmaterial (Abb. 7.18).

Imagery-Effekt: Der Spruch „Bilder sagen mehr als tausend Worte" kann durch Ergebnisse der Gedächtnisforschung bestätigt werden: Konkrete Objekte werden besser behalten als Abbildungen von diesen, und diese wieder besser als abstrakte Worte (Bevan & Steger, 1971). Paivio (1971) wies darauf hin, dass verbale Inhalte und Bildinhalte zwei komplementäre Quellen der Informationsaufnahme darstellen und dass viele Gedächtnisinhalte sowohl verbal-semantisch in Form von *Propositionen* als auch visuell-anschaulich in Form von

Abb 7.18



Im Experiment von Rogers, Kuiper und Kirker (1977) stuften Versuchspersonen 40 Wörter von Eigenschaften entweder nach deren geschriebener Wortgröße, nach möglicher Reimbildung, nach Bedeutungsübereinstimmung mit einem anderen Wort oder nach dem Selbstbezug der jeweiligen Eigenschaft ein. Danach wurden sie (überraschend) gebeten, so viele Eigenschaftswörter wie möglich wiederzugeben. Verglichen mit den Bedingungen einer oberflächlichen Beschäftigung mit dem Lernmaterial (strukturell, phonemisch) war die Leistung besser, wenn zuvor über die Bedeutung der Wörter nachgedacht werden musste (semantische Bedingung), am besten war sie aber, wenn die Eigenschaften vorher selbstbezogen zu reflektieren waren.

Vorstellungen gespeichert sind („Theorie der dualen Kodierung"). Ebenso wie die Verknüpfung von verbalen Elementen (Wörter) zu größeren Einheiten (Sätze, Reime, Geschichten) die Merkleistung verbessert (*Chunking*), trifft dies auch für visuelle Elemente zu (Box 7.4).

Aufgrund der besonderen Assoziierbarkeit visueller Vorstellungen wurde in der Antike bereits die sogenannte **Loci-Technik** („Technik der Orte") als Gedächtnisstütze von Rednern (z.B. Simonidis, Cicero) verwendet, indem sich diese eine gut bekannte Lokalität, wie etwa eine Straße, vorstellten und das erste Thema mit dem ersten Haus bzw. dessen Bewohner assoziierten, das zweite Thema mit dem zweiten Haus usw. Beim Vortrag riefen sie sich dann die zu merkenden Themen durch geistiges Abgehen der Orte ins Gedächtnis.

Eine ganz ähnliche Mnemotechnik ist die **Hakenmethode** („peg-word technique"), bei der vorausgehend eine fixe Koppelung von Ziffern mit Hakenwörtern, die man sich gut vorstellen kann, erlernt werden muss (z.B. 0 = Ei, 1 = Kerze, 2 = Schwan, 3 = Dreizack, 4 = Kleeblatt, 5 = Hand usw.). Die zu merkenden Inhalte (z.B. Objekte einer Einkaufsliste) werden dann mit den Hakenwortbildern interaktiv verknüpft (wenn z.B. der erste Einkaufsgegenstand ein Brot ist, wird dieses mit einer Kerze darauf vorgestellt; Abb. 7.19). Diese mnemotechnische Nutzung des Imagery-Effektes kann aber auch zur besseren Speicherung von Zahlen verwendet werden: Die Zahl

Verknüpfen in der Vorstellung | Box 7.4

Bower (1972) konnte zeigen, dass bei Merkaufgaben, in denen 5 x 20 Paare von konkreten Wörtern miteinander assoziiert werden mussten (je Paar 5 Sek. Lernzeit), bei einer Standardinstruktion (z.B. „Wenn das linke Wort gezeigt wird, versuchen Sie bitte das rechte Wort wiederzugeben") nur 50 % der Wortpaare erinnert werden konnten, im Gegensatz dazu aber 80 % der Wortpaare, wenn die Personen beim Lernen aufgefordert wurden, die Worte visuell miteinander zu verbinden (Bugelski, Kidd & Segmen, 1968). Die gut vorstellbaren Bedeutungen der Wörter, wie zum Beispiel Hund und Fahrrad, mussten dabei „interaktiv" verbunden werden, wie etwa durch die Vorstellung, wie der Hund auf dem Rad herumfährt. Welche interaktive Vorstellung gewählt wurde, blieb den Probandinnen und Probanden vorbehalten.

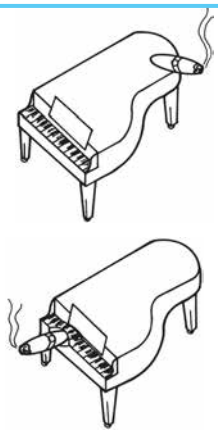
5024 zum Beispiel könnte vorgestellt werden als eigene Hand (= 5), in der ein Ei (= 0) liegt, auf dem ein Schwan (= 2) sitzt, der ein Kleeblatt (= 4) im Schnabel hat. Die Produktion solcher Vorstellungen kostet zwar zunächst Zeit, „amortisiert" sich aber später durch eine beträchtliche Einsparung an Wiederholungen bzw. Lernaufwand. Durch mnemotechnische Methoden der genannten Art lässt sich ein Lerngewinn von mindestens 20 bis 30 % erzielen, bei intensiverem Training auch mehr (z.B. Furst-Methode).

7.5.2 | Festigung von Wissen

Nach der Aufnahme der Informationen werden diese in den Speichersystemen meist noch nachverarbeitet und umstrukturiert (Meeter & Murre, 2004; Nadel et al., 2012). Erlebnisse werden durch Erzählen sprachlich nachformuliert und geistig nachgeformt, neues Wissen in vorhandenes Wissen eingepasst, logische Zusammenhänge werden hergestellt und Widersprüche geglättet, verbale Nachrichten in Vorstellungen transformiert usw. Hinsichtlich der Stabilisierung von kognitiven Einprägungen sind zumindest vier Einflussquellen hervorzuheben:

Konsolidierungseffekt: Die lange bekannte positive Wirkung des Schlafs auf Gedächtnisprozesse wird auch durch aktuelle Untersuchungen bestätigt. In der ersten Nachthälfte scheinen vor allem

Abb 7.19 |



In einem Paarlernexperiment zur mnemotechnischen Nutzung der Hakenwort-Technik fanden Wollen, Weber und Lowry (1972) bei visueller Vorstellung von Wörtern allein eine Merkleistung von 36 %, bei Verknüpfung der Vorstellungen aber eine Reproduktion von 74 % (z.B. für die Wortkombination Klavier und Zigarre). Spätere Untersuchungen (McDaniel et al., 1995) ergaben, dass humorige, absonderliche oder bizarre (d.h. distinkte) Vorstellungskombinationen (wie links unten) besser eingeprägt werden als realistische (wie links oben).

deklarative Gedächtnisinhalte gefestigt zu werden (Episoden, Faktenwissen), in der zweiten eher die prozeduralen Speicherinhalte (z.B. Konditionierungen, Fertigkeiten). Eine wesentliche Rolle spielen dabei offenbar die Stresshormone (z.B. Cortisol), deren Ausschüttung im Tiefschlaf gehemmt ist und im Traumschlaf zunimmt (Born & Plihal, 2000, Abb. 7.20). Auch aus anderen Bereichen der Forschung (z.B. Stressforschung) weiß man, dass in Stresssituationen deklarative, nicht aber prozedurale Gedächtnisleistungen beeinträchtigt sind.

Wiederholungseffekt: Der überwiegende Anteil unseres Wissens wird nicht durch einmalige Präsentation behalten, sondern durch Wiederholungen. Hinsichtlich der zeitlichen Verteilung der Wiederholungen bis zur Prüfung ist der **Spacing-Effekt** zu berücksichtigen, welcher besagt, dass das Wissen dauerhafter eingeprägt wird, wenn Wiederholungen nicht sofort hintereinander erfolgen, sondern über einen längeren Zeitraum verteilt sind (Abb. 7.21).

Wenn beispielsweise fünfzig Spanischvokabeln entweder an einem Tag fünfmal oder an fünf aufeinanderfolgenden Tagen jeweils einmal wiederholt werden, dann sind es im ersten Fall 68 % und im zweiten Fall 86 %, die beim Test nach 30 Tagen noch gewusst werden (Bahrick, 1984). In einer zusammenfassenden Auswertung von 254 Studien („Metaanalyse“) zu diesem Thema schätzen Cepeda und Mitarbeiter (2006), dass bei solcherart **verteiltem Lernen** eine um durchschnittlich 10 % bessere langfristige

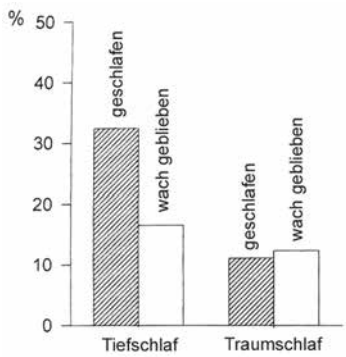
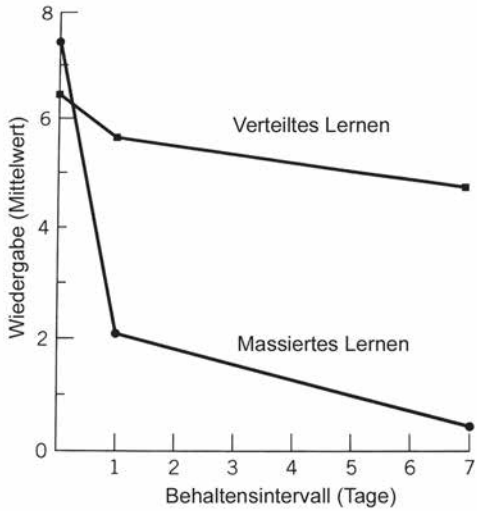


Abb 7.20

Um die Konsolidierungswirkung von Schlaf auf das Gedächtnis zu untersuchen, ließen Born und Plihal (2000) Gruppen von Probandinnen und Probanden Wortpaar-Assoziationen entweder um 23 Uhr – vor dem frühen Schlaf (Tiefschlaf) – oder um 2 Uhr – vor dem späten Schlaf (Traumschlaf) – lernen und prüften das Behalten jeweils etwa drei Stunden später. Eine Kontrollgruppe lernte das gleiche Material, blieb aber zwischen Lernen und Testung wach. Nach dem Tiefschlaf war eine Leistungssteigerung von 15 % zu verzeichnen, der Effekt des Traumschlafes unterschied sich bei dieser Aufgabe nicht von der Wachbedingung.

Abb 7.21



Wie sich die Verteilung von Lerninhalten über bestimmte Zeiträume auf die Behaltensleistung auswirkt, überprüfte Keppel (1964) mittels Assoziationslernen. Bei einer Gruppe von Probandinnen und Probanden gab es acht Lerndurchgänge an einem Tag (verdichtetes Lernen), bei der zweiten Gruppe je zwei Lerndurchgänge an vier aufeinanderfolgenden Tagen (verteiltes Lernen). Unmittelbar nach den Lernphasen war zwar die Gruppe mit massiertem Lernen etwas besser, fiel aber bereits nach einem Tag auf etwa ein Drittel der Leistung der Gruppe mit verteiltem Lernen ab.

Reproduktionsleistung als bei massiertem Lernen gegeben ist, und empfehlen die Aufteilung von Wissensstoff auf zumindest zwei Lernphasen.

Interferenz-Effekt: Auf den möglichen störenden Einfluss der Interferenz auf Wissensbestände wurde bereits bei den Vergessenstheorien hingewiesen (7.1.3). Wenn, wie im Zentralnervensystem, viele Speicherinhalte knapp hintereinander aufgenommen und nebeneinander abgelegt werden, kommt es notwendigerweise zu Wechselwirkungen und Störungen zwischen den Wissensinhalten. Bei knapp aufeinanderfolgenden Lernprozessen ist sowohl mit einer **proaktiven Interferenz** zu rechnen, bei der die Konsolidierungsprozesse der früheren Lernprozesse die Einprägung der späteren Lernvorgänge stören, als auch mit einer **retroaktiven Interferenz**, bei der die Konsolidierung gegenwärtig ablaufender Lernprozesse durch nachfolgende Lernaktivitäten beeinträchtigt wird. Um somit

gemäß dem (Hubert Rohracher zugeschriebenen) Spruch „Das Gehirn lernt länger als das Bewusstsein" der Nachverarbeitung von Speicherprozessen Rechnung zu tragen, sollten zwischen Lerninhalten möglichst viele kurze Pausen eingelegt und verwechselbare Lernstoffe in deutlichem Abstand voneinander gelernt werden.

Rekonstruktionseffekt: Neben diesen Formen der negativen Interferenz gibt es natürlich auch positive Wechselwirkungen zwischen Lerninhalten (Bradshaw & Anderson, 1982), etwa wenn sie die gleiche Grundstruktur haben (da z.B. die Funktionsweise eines Fernsehers jener eines Radios ähnlich ist, lässt sie sich danach gut lernen), wenn sie in einer Ursache-Wirkungs-Beziehung stehen oder wenn eine Oberbegriffsbeziehung vorhanden ist. Solche wechselseitigen Beziehungen zwischen Wissensmaterialien ermöglichen bei Vergessen von Einzelelementen oder Teileinheiten deren logische Rekonstruktion („reconstructive and inferential memory").

Abruf von Wissen

7.5.3

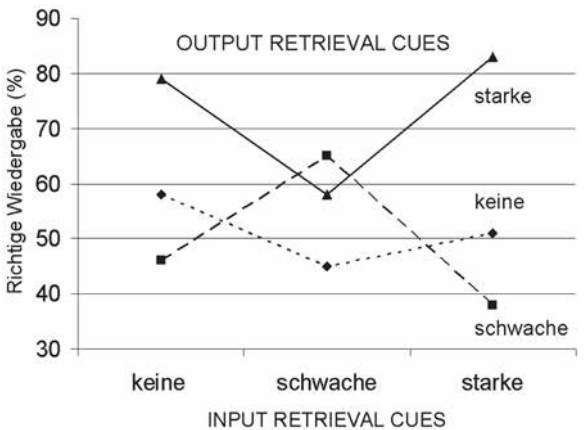
Auch in der dritten Phase der Gedächtnisprozesse, bei der Wiedergewinnung des gespeicherten Wissens, sind mehrere Aspekte von Bedeutung:

Abrufmodus-Effekt: Die Art, wie die gespeicherten Inhalte abgefragt bzw. angewendet werden, entscheidet über die zu erwartenden Gedächtnisleistungen. Auf die Differenzierung zwischen *impliziter und expliziter Gedächtnisleistung* wurde schon hingewiesen. Eine implizite Testung von Wissen bedeutet dessen Nutzung in der Wahrnehmung (z.B. durch Musterergänzen), im Denken (z.B. durch assoziatives *Priming*) oder im Verhalten (z.B. beim Signallernen, Fertigkeiten), womit also im weitesten Sinne seine **Verhaltensumsetzung** gemeint ist (z.B. Schachspielen, Gutachten erstellen, Sprechen einer Fremdsprache). Die explizite Testung von Wissen kann einerseits durch **Wiedererkennen** („recognition") erfolgen, andererseits durch **Wiedergabe** („recall"; Mandler, 1980). Wie schon mehrfach erwähnt, fällt es leichter zu entscheiden, ob der Speicherinhalt einer Erfahrung auf eine neue Erfahrung zutrifft, als die Speicherung selbst zu reproduzieren (Bruce & Cofer, 1967): Bei bloßem Wiedererkennen von Speicherobjekten sind also im Allgemeinen höhere Leistungen zu erwarten als bei aktiver Wiedergabe, bei Letzterer kann ohne („free recall") oder mit („cued recall")

leistungsverbessernden Erinnerungshilfen gearbeitet werden (s. etwa Abb. 7.11).

Retrieval-Cue-Effekt: Entsprechend dem **Encoding-Specifity Principle** von Tulving und Thomson (1973) werden Gedächtnisinhalte dann besser reproduziert, wenn in der Prüfsituation die gleichen Merkmale als Abrufhilfen herangezogen werden, die auch bei der Aufnahme der Information als deren Charakteristika wahrgenommen wurden (Meyer & Schwanefeldt, 1971; s. auch Abb. 7.6). Dies kann man auch im Alltag beobachten, wenn man zum Beispiel beim Gehen in ein anderes Zimmer eine beabsichtigte Tätigkeit vergessen hat und sich bei der Rückkehr in die ursprüngliche Situation wieder daran erinnert. Besonders wirksame Retrieval Cues sind solche, die in enger logischer oder kausaler Beziehung zu den Merkinhalten stehen und bereits in der Einprägungssituation mit diesen assoziiert waren (Abb. 7.22; Kenealy, 1997).

Abb 7.22



Um die Wirksamkeit von Erinnerungshilfen (retrieval cues) zu untersuchen, ließen Thomson und Tulving (1970) 24 Wörter unter drei Bedingungen lernen: Wörter allein, gemeinsam mit einem schwachen oder gemeinsam mit einem starken Hilfswort (welches entweder einen geringen oder einen starken Assoziationsbezug zu den Merkwörtern hatte, z.B. Merkwort: MANN, schwacher Cue: HAND, starker Cue: FRAU). Anschließend wurde die Liste wieder unter diesen drei Bedingungen geprüft (keine, schwache oder starke Erinnerungshilfe). Wenn die Abrufhilfe stark und auch die Einprägungshilfe stark war, konnten 83 % der Wörter wiedergegeben werden, aber fast ebenso viele (79 %) bei starker Abrufhilfe und überhaupt keiner Einprägungshilfe. Bei schwacher oder fehlender Abrufhilfe dagegen sank die Leistung auf etwa 50 % bis maximal 65 % (im Falle schwacher, aber immerhin gleicher Erinnerungshilfe).

Schwerwiegende Probleme können in Zusammenhang mit Zeugenaussagen vor Gericht entstehen, weil die Art zu fragen („Suggestivfragen“) die Antwort in eine bestimmte Richtung lenkt (Abb. 7.23). Fisher, Geiselman und Amador (1989) haben für Zeugenbefragungen ein „kognitives Interview“ entwickelt, dessen Anwendung in Befragungssituationen die Anzahl an brauchbaren Zeugenauskünften im Vergleich zu üblichen Kreuzverhörtechniken um fast 50 % erhöht und die Verfälschung von Aussagen nachweislich reduziert. Die Kriterien dabei sind: Vermeidung von Angst und Stress, Unterlassen von beurteilenden und persönlichen Kommentaren, sprachliches Nachvollziehen (*Paraphrasieren*) und offenes Interpretieren, Pausen zwischen den Befragungen, Anpassung der Sprache etc.

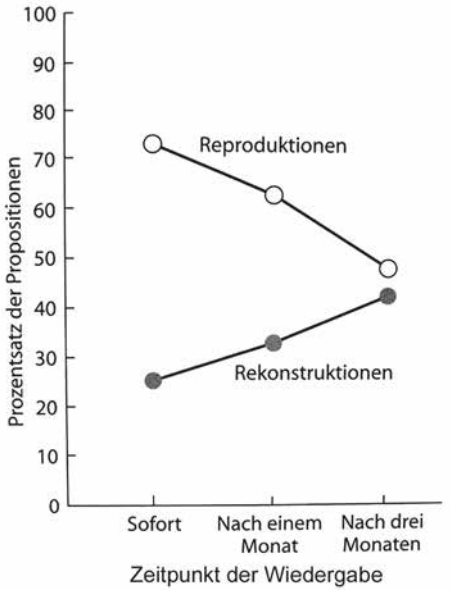
Schematisierungseffekt: Die Reproduktion von Speicherinhalten im Gedächtnis ist nicht mit der Abrufung von Daten im Computer zu verwechseln, sondern gleicht eher einem kreativen Rekonstruktionsprozess. Alle durch die Frage angesprochenen Speicherinhalte müssen reaktiviert und unter Zuhilfenahme von Altwissen miteinander verknüpft werden. Erlebnisse, die nur teilweise gemerkt und eventuell nur verbal formuliert abgespeichert sind, müssen bei Bedarf aus den Erinnerungsfragmenten, zumeist nach einem logischen **Schema**, abgeleitet werden. Bereits in den Dreißigerjahren hatte Frederic Bartlett (1886–1969) in Experimenten festgestellt, dass Kurzgeschichten beim Wiedererzählen systematisch verzerrt werden, indem man sie vereinfacht („Nivellierung“), Details hervorhebt und überbetont („Akzentuierung“) oder sie dem eigenen Verständnis und Hintergrundwissen anpasst („Assimilation“). Je komplexer die zu erinnernden Episoden sind und je länger sie zu-

Abb 7.23

Bei der Befragung von Zeuginnen und Zeugen kann die Formulierung der Frage beträchtliche Verfälschungen hervorrufen. Loftus und Palmer (1974) ließen Studierende einen kurzen Film über einen Zusammenstoß zweier Autos sehen und befragten sie anschließend über die vermutete Geschwindigkeit. Die Schätzungen der Personen sind zu den Frageformulierungen in Klammern hinzugefügt: „How fast were the two cars going when they smashed into (66 km/h)/collided with (62 km/h)/bumped into (61 km/h)/hit (54 km/h)/contacted (50 km/h) each other?“

rückliegen, desto stärker sind die rekonstruktiven Anteile (Abb. 7.24). Ein Beispiel für dynamische Schemata sind die schon erwähnten **Skripts**, in denen häufig erlebte Sequenzen von Erfahrungen gespeichert sind. So dient das schon beschriebene „Restaurant-Skript" (Eintreten – Platzsuchen – Hinsetzen – Bestellen – Konsumieren – Bezahlen) bei der Schilderung eines weit zurückliegenden Restaurantbesuches zur Rekonstruktion vergessener Teilerlebnisse.

Abb 7.24



In einer sehr umfangreichen Studie (Kintsch & van Dijk, 1978) zum Textverstehen hatten Studierende eine Abhandlung von 1300 Wörtern zu lesen, die danach in einem Bericht wiederzugeben war. Um den semantischen Gehalt quantifizieren zu können, wurden der Text und die Berichte in ihre semantischen Einheiten (Propositionen) zerlegt. Dadurch konnte erfasst werden, welche Inhalte bei den Erinnerungsberichten richtig reproduziert und welche nur rekonstruiert wurden. Im Vergleich zur sofortigen Wiedererzählung gehen bei der Testung nach drei Monaten die Reproduktionen von 72 auf 48 % zurück, während die Rekonstruktionen (logische und plausible Schlussfolgerungen aus dem Text) von 25 auf 44 % ansteigen.

erschwert werden (Abb. 8.11). Durch die Aufschaukelung bedingte Übersteuerungsreaktionen im täglichen Leben finden sich etwa beim Duschen durch hektischen Heiß-kalt-Wechsel, beim Übersteuern eines ins Schleudern geratenen Autos oder bei sich negativ aufschaukelnden E-Mails („Flaming").

Denken und Schlussfolgern 8.4

Wie bereits mehrfach erwähnt, besteht geistiges Problemlösen im zielgerichteten Einsatz von mentalen Operationen zur Transformation von psychischen Zuständen. Bei den dabei involvierten Denkprozessen handelt es sich im Wesentlichen um zwei Formen des Schlussfolgerns, nämlich um das **induktive** und das **deduktive Denken**. Bei Ersterem geht es darum, aus Einzelheiten (z.B. Sachverhalten) allgemeine Regeln (*Gesetze*) abzuleiten, bei der zweiten Form um den Schluss vom Allgemeinen (z.B. Menge von Annahmen) auf das Besondere (z.B. Konsequenzen daraus).

Induktives Denken 8.4.1

Beim **Begriffslernen** (s. 6.10) sollen für Gruppen von Objekten, Situationen oder Prozessen jene Merkmale und Merkmalsrelationen herausgefunden werden, die ihnen gemeinsam sind. Der Prozess ist vergleichbar mit jenem in der *Inferenzstatistik* (s. 3.6.2), wo von Fällen mit variierenden Variablenausprägungen auf einen gesetzmäßigen Zusammenhang der Variablen geschlossen wird.

Die Ausbildung von Begriffen ist einerseits als kontinuierlicher Lernprozess vorstellbar – wie bei der Mustererkennung (z.B. McClelland & Rogers, 2003; Jones & Hoskins, 1987; s. Abb. 7.1) – und andererseits als rückmeldungsgesteuerter Prozess des **Hypothesentestens**, bei dem vorerst Annahmen über Merkmale und deren Zusammenhänge gebildet und diese Hypothesen danach auf Basis von Erfahrungen überprüft werden. Bei Ersterem bilden sich Begriffe nur aufgrund der statistischen Beziehungen zwischen ihren Merkmalen (z.B. Korrelationen) – quasiautomatisch – aus, beim zweiten Prozess hingegen wird die hypothetische Begriffsstruktur anhand von Fallbeispielen verifiziert oder falsifiziert. Die Hypothese zum Beispiel, ob ein handgeschriebener Buchstabe ein A ist oder

nicht, wird durch entsprechendes Feedback vom Schreiber bestätigt oder verworfen.

Eine deutlich ausgeprägte, wahrscheinlich für viele Fehlleistungen des Alltags verantwortliche Denkneigung betrifft die Bestätigungstendenz beim Prüfen von Hypothesen. Wenn wir allgemeine Aussagen (z.B. Meinungen, Vorurteile) im Kopf haben, testen wir üblicherweise positiv, das heißt, wir suchen Beispiele, die der Annahme entsprechen, und nicht Gegenbeispiele, die Chancen für eine Widerlegung böten. Wason (1960) konnte diesen **Bestätigungsfehler** („confirmation bias") auch bei einfachen Hypothesen über Gesetzmäßigkeiten in Zahlenreihen überzeugend nachweisen.

8.4.2 | Deduktives Denken

„Logisches oder schlussfolgerndes Denken bezieht sich auf den Prozess, durch den der Mensch von schon Bekanntem zu weiterem Wissen gelangt" (Anderson, 1996, 303). Fast immer können aus einem explizit formulierten komplexen Wissensbestand auch implizite (indirekt erschließbare) Wissensinhalte abgeleitet werden. Wenn man zum Beispiel weiß, dass Gehirnerkrankungen das geistige Leistungsvermögen beeinträchtigen können, und wenn Enzephalitis eine Gehirnerkrankung ist, dann weiß man, dass diese Erkrankung eine intellektuelle Schädigung hervorrufen kann. Während in der Logik korrekt durchgeführte deduktive Schlüsse („Syllogismen") immer wahr sind, trifft dies für „psycho-logische" Schlüsse nicht unbedingt zu. Vielmehr gibt es mehrere Möglichkeiten, wie psychologische Ableitungen zustande kommen können (Oberauer, 2006; Johnson-Laird, 1999):

1. Ableitungen können auf Tatsachenwissen basieren, das Vorhersagen über die Wahrscheinlichkeit des Überganges von einem Realitätszustand in einen anderen gestattet (kognitive Modelle: „Produktionssysteme"). Beispiel: Wenn brennbare Substanzen entzündet werden, dann entwickelt sich Hitze.
2. Ableitungen können – ähnlich wie in der Aussagenlogik – als formaler Prozess ablaufen („abstract-rule theory"). Beispiel: Die Aussage „A impliziert B" ist wahr, „A" ist wahr, daher ist „B" ebenfalls wahr.
3. Ableitungen können als Folge der situationsspezifischen Konstruktion eines mentalen Modells im *Arbeitsgedächtnis* verstan-

den werden („model theory“, s. auch Abb. 8.6). Beispiel: Das Burgtheater in Wien ist gegenüber dem Rathaus, die Universität befindet sich rechts vom Rathaus (wenn man frontal davor steht), daher ist die Universität links vom Burgtheater zu finden, wenn man frontal vor diesem steht.

Kognitionspsychologische Untersuchungen zum deduktiven Denken beziehen sich vor allem auf konditionale und kategoriale Schlüsse. **Konditionale Schlüsse** sind jene spezielle Klasse von *aussagenlogischen* Schlussfolgerungen, bei denen Wenn-dann-Formulierungen (*Implikationen*) verwendet werden. Ein aussagenlogischer Schluss besteht im Allgemeinen aus **Prämissen** (Voraussetzungen), aus denen unter Verwendung von **Schlussformen** (Ableitungsoperationen) **Schlüsse** (Konklusionen) gezogen werden (Tab. 8.3).

Tab 8.3

Konditionalschluss				Beispiele	
Konklusion zulässig	Modus ponens	P1:	Wenn P, dann Q	P1:	Wenn Karin vor der Tür steht, dann ist Hans glücklich.
		P2:	P	P2:	Karin steht vor der Tür.
		K:	Q	K:	Hans ist glücklich.
	Modus tollens	P1:	Wenn P, dann Q	P1:	Wenn Karin vor der Tür steht, dann ist Hans glücklich.
		P2:	nicht Q	P2:	Hans ist nicht glücklich.
		K:	nicht P	K:	Karin steht nicht vor der Tür.
Konklusion nicht zulässig	Bestätigung des Hinterglieds	P1:	Wenn P, dann Q	P1:	Wenn Karin vor der Tür steht, dann ist Hans glücklich.
		P2:	Q	P2:	Hans ist glücklich.
		K:	P	K:	Karin steht vor der Tür.
	Verneinung des Vorderglieds	P1:	Wenn P, dann Q	P1:	Wenn Karin vor der Tür steht, dann ist Hans glücklich.
		P2:	nicht P	P2:	Karin steht nicht vor der Tür.
		K:	nicht P	K:	Hans ist nicht glücklich.

Darstellung von zwei erlaubten und zwei unerlaubten Schlussformen, worunter man gebräuchliche geistige Operationen versteht, um aus Prämissen Schlüsse zu ziehen. (Modifiziert aus Knoblich, 2002, 673)

Eine sehr gebräuchliche Schlussform ist der **Modus ponens**, bei dem von einer Wenn-dann-Gesetzmäßigkeit und einer gegebenen Wenn-Bedingung auf die Gültigkeit einer Dann-Bedingung geschlossen wird. Diese Schlussform wird im Alltag häufig eingesetzt und wird allgemein sehr gut beherrscht.

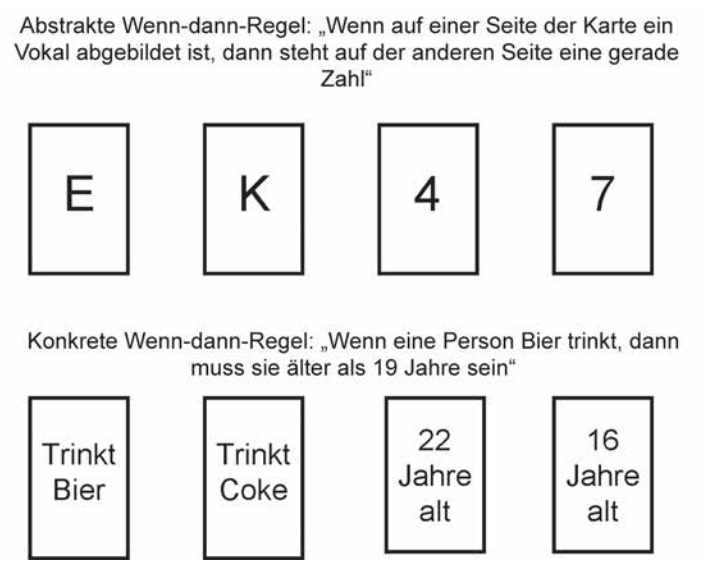
Wesentlich weniger im Denken genützt und nicht so gut in seiner Logik durchschaut ist der **Modus tollens**. In ihm wird ebenfalls eine gültige Wenn-dann-Gesetzmäßigkeit vorausgesetzt, und man schließt (korrekterweise) von einer nicht gegebenen Dann-Bedingung auf das Nichtvorliegen der Wenn-Bedingung (Tab. 8.3). Wenn Schlüsse des logisch korrekten Modus tollens zur Beurteilung vorgegeben werden, stimmen nicht 100 %, sondern nur zirka 50 bis 60 % der Versuchspersonen seiner Gültigkeit zu. Aber auch logische Schlüsse, die nur manchmal, aber nicht immer richtig sind („Bestätigung des Konsequens“ oder „Ablehnung des Antecedens“) erhalten in Experimenten immerhin zwischen 20 und 30 % unbedingte Zustimmung (Rips & Marcus, 1977). Die geringere Akzeptanz des Modus tollens als korrekte Schlussfolgerungsfigur lässt vermuten, dass logische Schlüsse immer dann schwerfallen, wenn in ihnen Verneinungen vorkommen. Offenbar um die Kanalkapazität des Arbeitsgedächtnisses nicht zu überfordern, scheinen in mentale Modelle (als Basis von Schlüssen) hauptsächlich solche Propositionen einbezogen zu werden, die als wahr gelten, während verneinende eher unbeachtet bleiben (Johnson-Laird, 1999). Allgemein kann man außerdem sagen, dass logische Ableitungen umso weniger fehleranfällig sind, je öfter sie in alltäglichen Schlussfolgerungen vorkommen, wie Abb. 8.12 zeigt („domain-specific-rule theories“).

Kategoriale Schlüsse sind bereits aus der Antike bekannt. Sie enthalten in den Prämissen sogenannte **Quantoren** (Mengenangaben: Allquantor, Existenzquantor usw.) und basieren auf Operationen der Prädikatenlogik. Ein Beispiel für einen richtigen Schluss mit den Quantoren „kein“ und „alle“ ist demnach:

Kein Rechteck ist ein Kreis.
Alle Quadrate sind Rechtecke.
Es folgt: Alle Quadrate sind keine Kreise.

Mengenrelationen und die darauf basierenden Schlussfolgerungen werden häufig in Form von „Venn-Diagrammen“ illustriert,

Abb 8.12



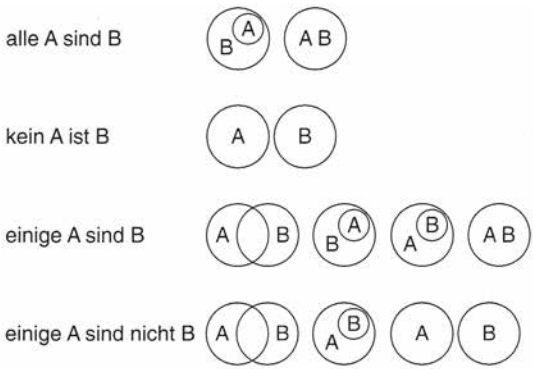
In der Auswahlaufgabe („selection task“) von Wason (1966) mussten die Probandinnen und Probanden eine abstrakte Wenn-dann-Regel (oben) auf ihre Gültigkeit prüfen, indem sie bei Vorlage von vier Karten mit zwei möglichen Buchstaben auf der einen und zwei möglichen Ziffern auf der anderen Seite so wenige Karten wie möglich umdrehen sollten. Etwa 90 % drehen richtigerweise die E-Karte um, weil eine 7 auf der anderen Seite die Regel falsifizieren würde. 62 % schauen unnötigerweise hinter die 4, wo sowohl ein E als auch ein K stehen könnte. Nur 25 % entscheiden sich richtigerweise für die 7, weil auf ihrer Rückseite kein E sein dürfte. Formuliert man die Aufgabe als konkretes „Erlaubnisschema“ (unten), dann wählen 74 % der Versuchspersonen die beiden richtigen Karten, nämlich „Trinkt Bier“ und „16 Jahre alt“ (Oaxford & Chater, 1994).

wodurch die den Prämissen innewohnenden Mengenaussagen leichter überschaubar und besser nachvollziehbar sind (Abb. 8.13).

Systematisch auftretende Fehler beim deduktiven Schlussfolgern können auf unterschiedliche Ursachen zurückgeführt werden (s. etwa Mayer, 1979; Anderson, 1996):

1. Vergessen von einzelnen Annahmen (bzw. von „Modellen“ gemäß der *Theorie mentaler Modelle*, s. Abb. 8.6), vor allem, wenn mehrere oder komplexe Prämissen gegeben sind.
2. Missinterpretation von logischen Relationen, wie etwa, wenn einseitige Wenn-dann-Beziehungen auch umgekehrt interpretiert werden („Konversionsfehler“), was in realen Situationen

Abb 8.13



Mittels „Venn-Diagrammen" lassen sich die Quantoren (alle, einige, einige nicht, kein) von kategorialen Schlüssen grafisch durch Mengenrelationen darstellen. Beispielsweise ist die Aussage „einige A sind B" mehrdeutig interpretierbar, weil ihr vier Mengenrelationen zugeordnet werden können, während die Behauptung „kein A ist B" nur durch die einander ausschließenden Mengen A und B darstellbar ist.

sogar zu richtigen Schlüssen führen kann (z.B. „Person A findet Person B sympathisch").

3. Logische Gesetzmäßigkeiten werden zum Teil zu Recht nicht mit 100 %, sondern nur mit einer bestimmten Wahrscheinlichkeit als gültig angesehen, weil viele Ereignisse im Alltag tatsächlich mehrere Ursachen und Voraussetzungen haben (z.B. „Wenn jemand Grippe hat, dann bekommt er wahrscheinlich Fieber"; s. auch 3.4).
4. Logische Bezeichnungen werden im Alltag anders interpretiert als in der Mathematik. So etwa versteht man in der Logik unter dem *Quantor* „einige" eine Menge von einem bis zu allen Elementen, während umgangssprachlich damit häufig zwischen zwei und fünf gemeint sind.
5. Wenn Schlussaufgaben nicht mehr gut durchschaut werden können und dennoch ein Urteil abzugeben ist, lassen sich Versuchspersonen oft zu gleichen Formulierungen verleiten (wie etwa „keine", „einige", „alle"), wie sie auch in den Prämissen vorkommen („Atmosphärenfehler").
6. Insbesondere bei Schlussfolgerungen im Alltag spielt auch die Erwünschtheit von Konsequenzen für ihre Richtigkeitsbeurtei-

Merksatz

Beim induktiven Denken wird aus einzelnen Erfahrungen auf ihnen zugrunde liegende allgemeine Regeln geschlossen. Beim deduktiven Denken werden aus postulierten allgemeinen Regeln implizit enthaltene Konsequenzen abgeleitet.

lung eine Rolle, insbesondere wieder dann, wenn die Struktur logischer Annahmen eher komplex und unüberschaubar ist („Hedonistischer Bias").
Sehr wahrscheinlich werden allerdings viele Schlussfolgerungen des Alltags häufig nicht mittels kategorialer oder konditionaler Schlussformen gezogen, sondern über Assoziationen („Expertise", implizite Eindrücke und das sogenannte „Bauchgefühl") (Arkes et al., 2016; Uleman, Saribay & Gonzalez, 2008; Gigerenzer, 2007), wie überhaupt der Einfluss von Emotionen auf Entscheidungen und Urteile („bounded rationality") bisher in der Forschung unterschätzt worden sein dürfte (Lerner et al., 2015).

Entscheidungsfindung und Urteilsbildung 8.5

Urteils- und Entscheidungsprozesse sind Phasen des Denkens und Problemlösens und lassen sich begrifflich kaum scharf voneinander abgrenzen. Da Urteile in allen Bereichen unseres Alltags weitreichende und mitunter sogar gefährliche Konsequenzen haben, fanden in der psychologischen Forschung insbesondere jene Effekte großes Interesse, die nicht formal-logisch erklärbar waren.

Ähnlichkeitsurteile 8.5.1

Das Zustandekommen eines Ähnlichkeitsurteils ist nur scheinbar auf eine einfache Übereinstimmung in den Merkmalsausprägungen der Vergleichsobjekte zurückzuführen. Tversky (1977) konnte nachweisen, dass bei unterschiedlicher Gruppierung von Objekten diese als unterschiedlich ähnlich erlebt werden (Abb. 8.14). Ebenso konnte er zeigen, dass man beim **Ähnlichkeitsurteil** stärker die Gemeinsamkeiten von Objekten beachtet, während beim **Unähnlichkeitsurteil** die Unterschiede stärker gewichtet werden. Allgemein nimmt das Ähnlichkeitsurteil für zwei Objekte in dem Ausmaß zu, in dem ihnen gleiche Merkmale zugeschrieben werden und es ihnen an verschiedenartigen Merkmalen mangelt, wobei allerdings die Auffälligkeit und subjektive Gewichtung der Merkmale eine Rolle spielt (z.B. Intensität, Bekanntheit, *Prototypikalität*, Informationsgehalt).

Abb 8.14

Schweden 49 %	Polen 15 %	Ungarn 36 %	Österreich
Schweden 14 %	Norwegen 26 %	Ungarn 60 %	Österreich

Zwei Objekte können als sehr unterschiedlich ähnlich beurteilt werden, je nach Art ihrer Gruppierung. Teversky (1977, 343) präsentierte 70 Versuchspersonen Kombinationen von vier Ländern mit der Aufforderung, jeweils für ein Land (z.B. Österreich) aus drei anderen Ländern jenes auszuwählen, welches am ähnlichsten ist. Das Beispiel zeigt, dass bei der ersten Vergleichsgruppe 49 % der Probandinnen und Probanden Österreich und Schweden als einander am ähnlichsten einstuften (im Kontrast zu damals kommunistischen Ländern), während sich bei der zweiten Gruppe (nun mit Norwegen statt Polen im Set) 60 % für Ungarn als ähnlichstes Land entschieden (offenbar in Abgrenzung zu den skandinavischen Ländern).

8.5.2 | Urteilsheuristiken und Urteilsrahmung

Heuristiken sind „Urteilsstrategien, die relativ schnell und mit vergleichsweise geringem Aufwand Erklärungen, Vorhersagen und Schlussfolgerungen ermöglichen" (Strack, 1985). Sie sind notwendige Instrumente unserer Denkökonomie, weil im Alltag kaum genug Informationen und Zeit für streng logische Urteile zur Verfügung stehen. Vor allem drei solcher Urteilsheuristiken („judgmental heuristics") waren häufig Gegenstand von Untersuchungen (s. Strack & Deutsch, 2002):

Merksatz

Heuristiken sind notwendige Instrumente unserer Denkökonomie, weil im Alltag kaum genug Information und Zeit für streng logische Urteile zur Verfügung stehen.

Die **Verfügbarkeitsheuristik** („availability heuristic") bewirkt, dass wir für Denk- und Urteilsprozesse jene Informationen heranziehen, die im Moment kognitiv verfügbar sind. Das können auffällige Wahrnehmungen, schnell aktivierbare Gedächtnisinhalte und logisch leicht reproduzierbare Denkergebnisse sein. Alles, was uns leicht in den Sinn kommt (s. *Priming*), wird bevorzugt für Urteilsprozesse verwendet. So etwa werden Personen mit auffallenden Eigenschaften (z.B. Kleidung, Hautfarbe, Bewegung) oder in auffallender Position (z.B. Sitzordnung, Beleuchtung) als einflussreicher wahrgenommen als andere. Todesrisiken, die in Presse oder Fernsehen oft genannt oder gezeigt werden, werden überschätzt (z.B. Mord, Sturmflut, Flugzeugabsturz) und kaum erwähnte Risiken unterschätzt (Herzkrankheiten, Krebs, Diabetes; s. 8.5.4).

Die **Repräsentativitätsheuristik** („representativeness heuristic“) wird dann wirksam, wenn ein Erfahrungsinhalt (z.B. ein Objekt, ein Vorgang) in ein kognitives Schema passt (z.B. Begriff, Kausalprozess) und somit von diesem repräsentiert wird. Dies ist etwa der Fall, wenn eine Person als repräsentativ (typisch) für ihre Nationalität, eine Handlung als repräsentativ für den Handelnden oder eine *Stichprobe* als repräsentativ für eine *Population* empfunden wird (Box 8.1). Unsere kognitive Repräsentation von Zufallsfolgen lässt uns etwa die Serie 1, 2, 3, 4, 5, 6 aus dem Glücksspiel „6 aus 45“ als weniger zufällig erscheinen als etwa die Folge 29, 4, 35, 17, 41, 18.

Missachtung der Grundrate | Box 8.1

In dem Ausmaß, in dem eine kognitive Repräsentation (z.B. Klassifikation, *Schema*, *Skript*) auf einen Erfahrungsinhalt zutrifft, werden auch die in der Repräsentation gespeicherten Gesetzmäßigkeiten als gegeben angenommen. Je mehr zum Beispiel vorhandene Symptome einer bestimmten Krankheit entsprechen, desto eher wird diese als gegeben angenommen, und zwar auch dann, wenn die Auftrittswahrscheinlichkeit für die Krankheit (*Prävalenz*) eigentlich sehr gering ist. Gegen diese sogenannte Missachtung der Grundrate („base rate fallacy“) bei der Einschätzung von Störungen sind nicht einmal Ärztinnen und Ärzte gefeit, wie folgendes Experiment zeigt: Mediziner in verschiedenen Ausbildungsstufen der Harvard Medical School sollten für einen Patienten die Wahrscheinlichkeit für das Vorliegen einer bestimmten Erkrankung einschätzen. Anzunehmen war, (1) dass ein krankheitsspezifischer Labortest positiv anspräche und dass dieser bei Vorliegen der Krankheit zu 100 %, aber auch bei Gesunden in 5 % der Fälle positiv ist und (2) dass die Krankheit in der Bevölkerung mit einer relativen Häufigkeit von 0,001 aufträte. Die richtige Antwort, nämlich etwa 2 % Erkrankungswahrscheinlichkeit (s. 8.5.3) wurde verblüffenderweise nur von 18 % der befragten Mediziner geschätzt, während die meisten 95 % als Erkrankungswahrscheinlichkeit angaben (Cascells, Schoenberger & Graboys, 1978). Das heißt, dass sich auch medizinische Fachleute von der Repräsentativität eines Symptoms (Tests) dazu verleiten lassen, in ihrem Urteil die Auftrittshäufigkeit der Krankheit zu ignorieren.

Eine dritte, viel verwendete und automatisierte Urteilsheuristik ist die **Anker-** oder **Anpassungsheuristik** („anchoring", „adjustment"). Wie aus der Bezeichnung hervorgeht, führt sie zu einer Anpassung von Urteilen an vorhandene Orientierungsrichtlinien. Dies mag zwar die Einschätzungsprozesse oft beschleunigen, stellt aber umgekehrt wieder eine Quelle für „kognitive Täuschungen" dar (z.B. Hell, Fiedler & Gigerenzer, 1993). Eine ankerbedingte Verfälschung von Urteilen kann mittels einer einfachen Multiplikationsrechnung demonstriert werden, sobald die Probandinnen und Probanden nur fünf Sekunden Rechenzeit haben und ihnen nur eine grobe Schätzung des Ergebnisses möglich ist. Im diesbezüglichen Experiment von Tversky und Kahneman (1974) kamen die Versuchspersonen bei der Vorgabe von

8 x 7 x 6 x 5 x 4 x 3 x 2 x 1

auf ein mittleres Produkt von 2250 („Median"), bei der Vorgabe der gleichen Zahlen, jedoch in ansteigender Reihenfolge

1 x 2 x 3 x 4 x 5 x 6 x 7 x 8

im Mittel nur mehr auf 512. In der ersten Bedingung bilden die Produkte der ersten Zahlen einen wesentlich größeren Anker (z.B. 8 x 7 = 56, 56 x 6 = 336 usw.) als in der zweiten Bedingung (z.B. 1 x 2 = 2, 2 x 3 = 6 usw.). Die erhebliche Unterschätzung des richtigen Ergebnisses, nämlich 40.320, zeigt außerdem, dass arithmetische Reihen ähnlich wie exponentielle Funktionen in ihrer Entwicklung schlecht vorhergesagt werden können.

Eine andere Art von Ankerheuristik ist die sogenannte **Rahmung** („framing"), bei der eine Urteilssituation so beschrieben wird, dass in ihr bereits richtungsweisende Bezugsgrößen für eine Schätzung oder Klassifikation enthalten sind (Box 8.2). Eine solche Rahmung kann durch die Formulierung von Konsequenzen einer Entscheidung erzeugt werden (Kahneman & Tversky, 1983): Wenn zum Beispiel Patienten mit Krebs sich für eine bestimmte Behandlungsform (Operation oder Strahlentherapie) entscheiden sollen, wird ihre Entscheidung davon beeinflusst, ob die jeweiligen Risiken anhand der Überlebensrate (z.B. 95 %) oder anhand der Mortalitätsrate (z.B. 5 %) angegeben werden.

Merksatz

Unter Rahmung eines Urteils versteht man die Beschreibung einer Urteilssituation, in der die Richtung des Urteils bereits eingegrenzt ist.

Die **Prospekttheorie** (Tversky & Kahneman, 1981) erklärt viele Rahmungseffekte durch Annahme einer nichtlinearen *psychophysischen*

Rahmung von Problemsituationen | Box 8.2

Die *Rahmung* von Problemsituationen hat wesentlichen Einfluss auf die Art von Entscheidungen, die getroffen werden: Kahneman und Tversky (1983) ersuchten Versuchspersonen (n = 152), sich in der Rolle als Gesundheitsbeauftragte im Kampf gegen eine gefährliche Infektionskrankheit mit 600 möglichen Todesopfern für eine von zwei Strategien zu entscheiden:

Konsequenzen des Gesundheitsprogramms	Zustimmung
Strategie A: 200 Personen werden gerettet	72 %
Strategie B: Mit 1/3 Wahrscheinlichkeit werden alle 600 gerettet und mit 2/3 Wahrscheinlichkeit wird niemand gerettet	28 %

Weiteren Versuchspersonen (n = 155) wurde eine andere Beschreibung des gleichen Problems vorgelegt:

Konsequenzen des Gesundheitsprogramms	Zustimmung
Strategie C: 400 Personen werden sterben	22 %
Strategie D: Mit 1/3 Wahrscheinlichkeit wird niemand sterben und mit 2/3 Wahrscheinlichkeit werden alle 600 sterben	78 %

Wenn also, wie im ersten Beispiel, die Aufmerksamkeit auf Gewinn gelegt wird (Strategie A), dann wird eher die sichere Alternative gewählt, wenn das gleiche Ergebnis als Verlust beschrieben wird (Strategie C), entscheidet man sich lieber für die riskante Alternative. Gewinnrahmung macht risikoscheu und Verlustrahmung risikofreudig. Wie spätere Experimente zeigten, schwächt sich dieser Rahmungseffekt ab, wenn statt 600 nur 60 oder nur 6 Personen gerettet werden können und wenn es sich bei den Betroffenen um nahe Verwandte handelt.

Wertfunktion: Der subjektive Nutzen nimmt nicht linear mit der Größe eines Gewinnes zu, sondern schwächt sich immer mehr ab, was ebenso auf den subjektiven Schaden bei Verlusten zutrifft. Die Praxisrelevanz der Prospekttheorie zeigten Wong und Kwong

(2005) am Beispiel der Evaluation des Zuspätkommens von Beschäftigten auf, wo bei Verwendung kleiner Zahlen (z.B. Abwesenheitsrate von 10 %) individuelle Unterschiede größer erschienen, als wenn große Zahlen herangezogen wurden (z.B. Anwesenheitsrate von 90 %). Für seine Beiträge zur Wirtschaftstheorie erhielt der Psychologe Daniel Kahneman übrigens im Jahr 2002 den Nobelpreis für Wirtschaftswissenschaften.

8.5.3 | Wahrscheinlichkeitsurteile

Eine wesentliche Aufgabe unseres kognitiven Systems ist die Bildung und Änderung unserer Meinungen über die Realität durch Auswertung einschlägiger empirischer Erfahrungen. Aufgrund der Komplexität unserer Wirklichkeit können wir oft keine eindeutigen Urteile fällen, sondern müssen uns mit Wahrscheinlichkeitsurteilen begnügen. Wie wahrscheinlich ist es zum Beispiel, dass ein Mensch lügt oder dass eine schwere Krankheit vorliegt?

Solche Schätzungen sollten einerseits die sogenannte **Apriori-Wahrscheinlichkeit** (*Grundrate*, p(H)) der jeweiligen Hypothese (H) berücksichtigen, d.h. die Auftrittswahrscheinlichkeit der postulierten Gesetzmäßigkeit allgemein (z.B., dass eine Person überhaupt ein Lügner ist), und andererseits – bei Hinzukommen neuer Daten (D) – auch die **bedingte Wahrscheinlichkeit** p(D/H), mit der eine hypothesenkonforme Erfahrung bei Gültigkeit der Hypothese auftritt (z.B., wie wahrscheinlich es ist, dass ein Lügner lügt). Die Wahrscheinlichkeit für eine Hypothese nach deren Revision aufgrund neuer Erfahrungen nennt man **Aposteriori-Wahrscheinlichkeit** p(H/D). Das **Bayes-Theorem** verbindet diese Wahrscheinlichkeiten und erlaubt wahrscheinlichkeitstheoretisch korrekte Schätzungen über die Gültigkeit von Hypothesen:

Merksatz

Annahmen über die Realität manifestieren sich in Wahrscheinlichkeitsurteilen, die durch Einbeziehung neuer Realitätserfahrungen modifiziert werden.

Thomas Bayes (1702–1761): britischer Mathematiker und Geistlicher

$$p(H/D) = \frac{p(H) \cdot P(D/H)}{p(D)} = \frac{p(H) \cdot p(D/H)}{p(H) \cdot p(D/H) + p(\overline{H}) \cdot p(D/\overline{H})}$$

Wenn zwei Hypothesen in ihrem Wahrscheinlichkeitsverhältnis überprüft werden sollen (z.B. wie viel wahrscheinlicher ist H_1 im Vergleich zu H_2), kann p(D) weggekürzt werden, und man erhält:

$$\frac{p(H_1/D)}{p(H_2/D)} = \frac{p(H_1)}{p(H_2)} \cdot \frac{p(D/H_1)}{p(D/H_2)}$$

(Der Index kennzeichnet die Wahrscheinlichkeiten für Hypothesen 1 und 2.)

Wenn sich in Denk- und Schätzprozessen die in den obigen Formeln aufgezeigten Verhältnisse zumindest annähernd abbilden, spricht man auch von **statistischem Denken** oder von **Bayes-Logik** („Bayesian reasoning"). Die Alltagserfahrung und einschlägige psychologische Experimente zeigen jedoch, dass es in der Praxis durch Missachtung der Wahrscheinlichkeitsgesetze mitunter zu drastischen Fehleinschätzungen der Wahrscheinlichkeit von Ereignissen kommt (s. auch 8.5.2). Die Fehleranfälligkeit des menschlichen Denkens lässt sich an verschiedenartigen Urteilseffekten demonstrieren. Die Bedingungen hierfür suchten Gigerenzer und Hoffrage (1995) – im Gegensatz zu anderen Autoren – jedoch gerade in der Leistungsfähigkeit und der Entwicklungsgeschichte unseres Denkens und Urteilens. Sie meinen, dass die kognitiven Denkmuster in der evolutionären Entwicklung des Menschen nicht auf den Umgang mit Wahrscheinlichkeiten zugeschnitten waren, sondern auf Häufigkeiten (Kleiter, 1994; Hoffrage et al., 2015; „natural sampling"). Sie überprüften 15 häufig in Urteilsuntersuchungen verwendete Beispiele und konnten zeigen, dass die Versuchspersonen bei Vorgabe der Probleme im **Häufigkeitsformat** zu etwa 48 % die logisch korrekte Bayes-Logik verwendeten, während im **Wahrscheinlichkeitsformat** nur etwa 22 % wahrscheinlichkeitstheoretisch richtige Schätzungen abgaben (s. Abb. 8.15).

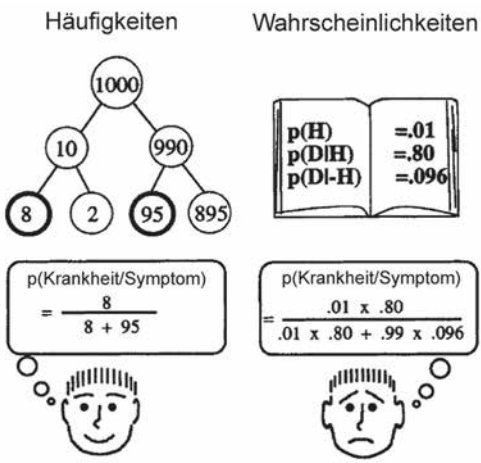
Die einfache, allgemeine Regel für eine praxisgerechte Revision der Wahrscheinlichkeit von Hypothesen (H) aufgrund von hypothesenkonformen oder hypothesendiskrepanten Häufigkeitsdaten (D) lautet nach Gigerenzer und Hoffrage (1995) wie folgt:

$$p(H/D) = \frac{f(H \cap D)}{f(H \cap D) + f(\overline{H} \cap D)}$$

$$Treffer - Frequenz : f(H \cap D)$$
$$Falscher - Alarm - Frequenz : f(\overline{H} \cap D)$$
$$Hypothesen - Negation : \overline{H}$$

Ein praktisches Beispiel könnte die Einschätzung des bedingten Risikos eines Unfalles (H) bei Trunkenheit (D) sein: Wie groß ist die Wahrscheinlichkeit, dass ein alkoholisierter Autofahrer einen Verkehrsunfall verursacht, wenn man weiß, dass früher 55 Betrunkene einen Unfall (H∩D) und 500 Betrunkene keinen Unfall ($\bar{H}$∩D) verschuldet haben? Bei Missachtung anderer Einflussfaktoren kann nach obiger Formel mit einer Wahrscheinlichkeit von p ≅ 0,10 (= 55/[55+500]), d.h. mit einer relativen Häufigkeit von 10 % das Auftreten eines Unfalles durch Alkoholisierung vermutet werden.

Abb 8.15



Das in vielen Urteilsexperimenten verwendete „Mammografie-Problem" kann im Häufigkeitsformat und im Wahrscheinlichkeitsformat dargeboten werden. Annahme: Bei etwa 10 von 1000 Frauen (p = 0,01) im Alter von 40 Jahren wird bei einer Routineuntersuchung Brustkrebs festgestellt. Bei 8 von den 10 Frauen, die Krebs haben, ist auch der Mammografie-Befund positiv, von jenen 990 Frauen, die nicht Krebs haben, ist aber bei 95 der Mammografie-Befund ebenfalls positiv. Wie wahrscheinlich ist es, dass eine Frau mit positivem Untersuchungsbefund Brustkrebs hat? Wie die Grafik anschaulich zeigt, ist die Wahrscheinlichkeitsberechnung komplizierter als die Häufigkeitsberechnung, führt aber zum selben Ergebnis, nämlich 7,8 % bzw. p = 0,078. Im Häufigkeitsformat muss nur der Quotient zwischen richtig diagnostizierten Kranken (8) zur Gesamtanzahl aller positiv (d.h. der 8 richtig und 95 falsch) diagnostizierten Personen (103) in Bezug gesetzt werden (Gigerenzer & Hoffrage, 1995).

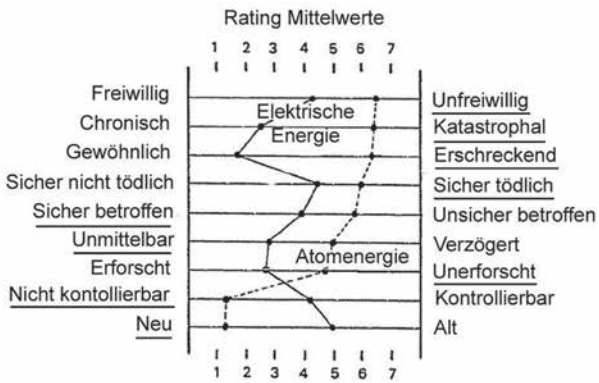
Ein anderes Beispiel betrifft die Vorhersage von Krankheit (H) durch Stress (D): Wie gut kann von einer Stressbelastung auf das spätere Auftreten einer Infektionskrankheit geschlossen werden, wenn sich bei einer Person bisher in acht Fällen nach psychischen Belastungen eine Erkältungskrankheit einstellte und in drei Fällen nicht? Nach der Bayes-Logik kann ein stressbedingter Krankheitsfall mit einer Wahrscheinlichkeit von $p \cong 0{,}727$ (= 8/[8+3]) bzw. 73 % vermutet werden (s. auch Prinz et al., 2015).

Risikowahrnehmung

8.5.4

Die Risikoeinschätzung von Gefahren unterscheidet sich von Person zu Person oft erheblich, sowohl bezüglich ihrer Auftrittswahrscheinlichkeit als auch bezüglich ihrer Folgen (Gigerenzer, 2013). Während Laien die Folgen von Risiken eventuell aufgrund konkreter Erfahrungen oft wesentlich differenzierter sehen als Expertinnen und Experten, sind umgekehrt Letztere aufgrund ihres statis-

Abb 8.16



Risiken werden sowohl hinsichtlich ihrer Auftrittswahrscheinlichkeit als auch bezüglich ihres Schadenspotenzials subjektiv unterschiedlich beurteilt (z.B. Atomkraft – elektrische Energie). Mittels Risikocharakteristika oder Risikodimensionen sollen daher individuelle und öffentliche Reaktionen auf Bedrohungen besser vorhersagbar gemacht werden. Slovic, Fischhoff und Lichtenstein (1985) fanden als subjektiv risikosteigernd zusätzlich zu den neun unterstrichenen Polaritäten auch die Aspekte „Bedrohung späterer Generationen", „globale Katastrophe", „zunehmende Gefahr", „ungerechte Risikostreuung", „schwierige Reduzierbarkeit" und „geringe präventive Kontrolle".

tischen Zuganges oft genauer in der Einschätzung der Wahrscheinlichkeit von Risiken. Untersuchungen zur Risikowahrnehmung von Slovic, Fischoff und Lichtenstein (1980, 1985; Slovic, 1987) ergaben insgesamt 18 differenzierende Charakteristika zur Klassifikation von Risiken mit interessanten Rückschlüssen auf die Risikoeinschätzung in der Bevölkerung. Mit erhöhter Risikoeinschätzung verbunden waren insbesondere folgende Merkmale (s. auch Abb. 8.16):

- **Unfreiwilligkeit** („involuntary"): Freiwillig eingegangene Risiken werden in der Regel als weniger gefährlich eingestuft.
- **Katastrophenartigkeit** („catastrophic"): Ereignisse, die mit übermäßig großen Verlusten einhergehen, wirken bedrohlicher.
- **Ungewöhnlichkeit** („dread"): Seltene, furchtbesetzte Vorfälle oder Szenarien werden – obzwar nicht immer rational nachvollziehbar – als hoch riskant eingeschätzt.
- **Tödlichkeit** („certainly fatal"): Je mehr Tote durch ein Risiko zu erwarten sind, desto höher die subjektive Risikoeinschätzung.
- **Betroffenheit** („known to exposed"): Ereignisse, die einen selbst betreffen können, erscheinen gefährlicher.
- **Unmittelbarkeit** („effect immediate"): Plötzlich zu erwartende Ereignisse wirken bedrohlicher.
- **Unerforschtheit** („not known to science"): Je weniger erforscht die Risiken scheinen, desto bedrohlicher wirken sie.
- **Unkontrollierbarkeit** („not controllable"): Eine Situation erscheint umso riskanter, je weniger man darauf Einfluss zu haben scheint (z.B. als Beifahrer im Auto oder als Fluggast).
- **Neuheit** („new risk"): Unbekannte Gefahren (z.B. COVID-19) werden mehr gefürchtet als bekannte (z.B. Lungenentzündung).

Hinsichtlich der Einschätzung des Auftretens von Risiken wirkt sich besonders die erwähnte *Verfügbarkeitsheuristik* verfälschend aus: Jene Gefahren, über die häufig gesprochen oder in den Massenmedien berichtet wird (wie Tornados, Flutkatastrophen, Geburtskomplikationen, Impfungen usw.), werden als wahrscheinlicher eingestuft, während triviale Bedrohungen (wie Herzkrankheiten, Krebs, Schlaganfall, Diabetes usw.) um das 10- bis 15-Fache unterschätzt werden (Slovic, Fischoff & Lichtenstein, 1980). Auch jene Aktivitäten oder Technologien, welche am häufigsten für Todesfälle verantwortlich sind, werden drastisch unterschätzt (z.B. Rauchen, Alkohol, Kraftfahrzeuge, Handfeuerwaffen, Elektrizität, Motorräder, Schwimmen etc.).

Hinsichtlich der Einschätzung der subjektiven Gefährdung scheint überdies ein sogenannter **Risiko-Optimismus** wirksam zu sein. In einer Befragung von Ruff (1990) schätzten die Probanden ihre eigene Gefährdung durch Krebs, Herzinfarkt, AIDS, Arbeitslosigkeit, Verkehrsunfall, Überfall, Umweltverschmutzung und Atomkrieg jeweils geringer ein als jene der übrigen Bevölkerung. Eine Untersuchung der Risikobeurteilung von Atomkraftwerken (Maderthaner et al., 1978) hat überdies ergeben, dass die dauernde Konfrontation mit einem Gefahrenobjekt (in diesem Fall durch die Nähe des Wohnortes) eine Abwertung des Risikos zur Folge hatte (vgl. dazu auch die *Theorie der kognitiven Dissonanz*, Kap. 10).

Einen anderen Einfluss auf das Urteilsverhalten zeigen Tversky und Köhler (1994) mit ihrer **Support-Theorie** auf, welche besagt, dass die subjektive Wahrscheinlichkeit einer Gefährdung auch wesentlich von der Beschreibung der Ereignisse abhängt: Je detaillierter die Gefahren geschildert oder subjektiv vorgestellt werden, desto höhere Wahrscheinlichkeiten werden ihnen in Summe zugeschrieben. Die Auswirkungen des expliziten Beschreibens von Ereignissen („unpacking") zeigte etwa folgende Studie: Natürliche und unnatürliche Todesursachen (Vorkommenshäufigkeit: 92 % bzw. 8 %) wurden von den Probanden zunächst auf 58 % bzw. auf 32 % geschätzt (ohne die Summierbarkeit auf 100 % zu beachten). Wenn nun die Todesursachen aufgeschlüsselt zu beurteilen waren, betrug plötzlich die Summe der subjektiven Wahrscheinlichkeiten für die natürlichen Todesursachen 73 % (Einzelschätzungen: Herzkrankheiten: 22 %, Krebserkrankungen: 18 %, andere natürliche Erkrankungen: 33 %) und für die unnatürlichen Todesursachen 53 % (Unfall: 32 %, Mord: 10 %, andere unnatürliche Ursachen: 11 %).

Intelligenz – Geistige Leistungsfähigkeit | 8.6

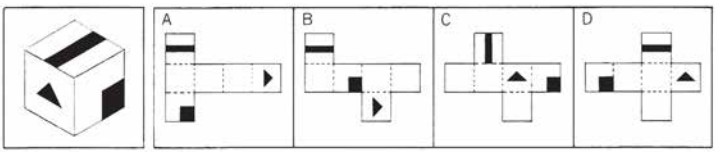
Intelligenz lässt sich definieren als „Fähigkeit, aus Erfahrung zu lernen, Probleme zu lösen und Wissen einzusetzen, um sich an neue Situationen anzupassen" (Myers, 2005, 460; Spektrum-Lexikon, 2020). Als im Jahre 1905 der französische Bildungs- und Erziehungsminister zur Entwicklung besserer Lehrmethoden für Kinder mit Entwicklungsstörungen aufrief, waren der ursprüngliche Jurist und spätere Psychologe Alfred Binet (1857–1911) und der Arzt Theodore Simon (1873–1961) überzeugt, dass zunächst die Mög-

Merksatz

Intelligenz ist ein theoretisches Konstrukt, das die geistige Leistungsfähigkeit einer Person charakterisiert: die erfolgreiche Verarbeitung von Information, der erfolgreiche Einsatz von Lernprozessen und die Nutzung von Wissen zur Lösung von Problemen.

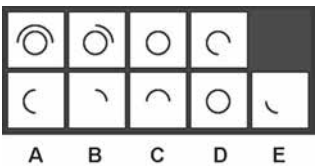
lichkeit einer Selektion der Kinder nach ihrer geistigen Leistungsfähigkeit geschaffen werden müsse. Dies war der Beginn der wissenschaftlich begründeten **Intelligenzdiagnostik** (s. Vanecek, 2003) und der *psychologischen Diagnostik*, welche im „Einsatz festgelegter Testverfahren zur Bewertung von Fähigkeiten, Verhaltensweisen und Persönlichkeitseigenschaften von Personen" besteht (Zimbardo & Gerrig, 2004, 399). Die eingesetzten Testverfahren müssen heute strengen *Gütekriterien* (s. 3.7.4) genügen, und ihre Weiterentwicklung erfordert daher einen erheblichen zeitlichen und finanziellen Aufwand (Aufgabenbeispiele: Abb. 8.17 und Abb. 8.18).

Abb 8.17



Beispiel einer Intelligenztestaufgabe zur Raumwahrnehmung: Welches der vier Muster (A, B, C oder D) entsteht, wenn der Würfel aufgefaltet wird? (Lösung s. S. 293)

Abb 8.18



Beispiel einer Intelligenzaufgabe zum logischen Schlussfolgern: Welches der Bilder in der unteren Bildreihe stellt eine logische Fortsetzung der obigen Bildfolge dar? (Lösung s. S. 293)

8.6.1 | Intelligenzdiagnostik

Binet (1911) ging mit seiner „Binet-Simon-Intelligenzskala" ursprünglich davon aus, altersgerechte Intelligenzaufgaben für normal leistungsfähige Kinder zwischen 3 und 15 Jahren zu finden. Die Lösung der Aufgaben einer bestimmten Altersgruppe erlaubte die Zuordnung des entsprechenden „mentalen Niveaus" bzw. **Intel-**

ligenzalters, welches mit dem Lebensalter in Bezug gesetzt wurde (Herle, 2003). Je größer die Differenz zwischen Intelligenzalter und Lebensalter war, desto mehr wich die Intelligenzleistung vom Durchschnitt der Altersgruppe ab. Da aber bei einem dreijährigen Kind eine Differenz von zwei Jahren mehr bedeutet als bei einem fünfzehnjährigen Jugendlichen, schlug William (alias Wilhelm) Stern (1911/1994) den **Intelligenzquotienten** vor:

$$IQ = \frac{IA}{LA} \cdot 100$$

(Die Multiplikation mit hundert soll ganze Zahlen ermöglichen.)

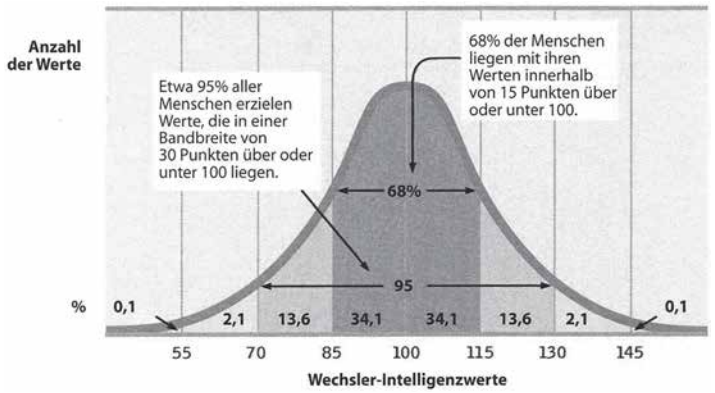
Nachdem weiters festgestellt worden war, dass ab dem frühen Erwachsenenalter bestimmte Intelligenzleistungen wieder abnehmen (analytisches Denken, reasoning, *fluide Intelligenz*, s.u.), musste schließlich auch dieses Intelligenzmaß ersetzt werden durch einen altersrelativierten **Abweichungsquotienten**, wie er bereits von Lewis Terman (1877–1956) in seinem 1916 publizierten „Stanford-Binet-Test" eingeführt worden war. Dieser bis heute gültige Quotient bezieht die Abweichung der individuellen Testleistung vom Mittelwert auf die Streuung aller Testleistungen der altersentsprechenden Bevölkerung:

$$IQ = 100 + \frac{(Punktewert - Punktemittelwert)}{Punktestreuung} \cdot 15$$

Damit wird also eine gruppenbezogene Normierung der Intelligenzwerte vorgenommen. Ein IQ von 100 bedeutet, dass 50 % der altersspezifischen Bevölkerung besser und 50 % schlechter abschneiden (s. Abb. 8.19), und ein IQ-Wert von 70, dass nur etwa 2,5 % der entsprechenden Altersgruppe noch geringere Leistungen aufweisen („Intelligenzminderung" bzw. „geistige Behinderung").

Ein interessantes Phänomen ist der **Flynn-Effekt**, der lineare Anstieg in den Intelligenzleistungen seit Beginn der Intelligenzmessung in mindestens 14 Ländern, der bis zu 25 IQ-Punkte von einer Generation zur anderen (30 Jahre) beträgt. Dieser Trend zeigt sich besonders bei kulturunabhängigen Intelligenzleistungen (z.B. *fluid intelligence*) und wird teilweise mit der Verbesserung der Ernährung, der Hebung der Schulbildung, mit zunehmender Umweltkomplexität und Urbanisierung sowie mit ansteigender Testvertrautheit erklärt (Flynn, 1987; Pietschnig, Tran & Voracek, 2013;

Abb 8.19



Da es bei den meisten Leistungs- und Eignungstests keine absoluten, d.h. allgemein gültigen Leistungsniveaus gibt, werden die individuellen Ergebnisse mit jenen anderer Personen verglichen. Dafür wird eine große Stichprobe aus der Gesamtbevölkerung getestet und deren (Normal-)Verteilung zur Bestimmung von Standard- oder Normwerten verwendet (Normierung; s. 3.7.4). Die Abweichung der individuell erreichten Testpunktezahl vom Mittelwert der Bezugsgruppe dividiert durch die Standardabweichung der Verteilung ergibt den Standardwert einer Person (s. 3.6.1) und damit ihre relative Position auf dem allgemeinen Leistungskontinuum. Der Mittelwert aller Intelligenztestwerte ist mit 100 festgelegt und die Streuung mit 15 (oder 10). Aufgrund der angenommenen Normalverteilung der Werte können Prozentsätze für die Intelligenzbereiche bestimmt werden.

Pietschnig et al., 2019). In neueren Veröffentlichungen wird auch ein inverser Flynn-Effekt diskutiert (Dutton et al., 2016).

Im deutschsprachigen Raum sind folgende beiden Intelligenztests sehr gebräuchlich, die jeweils mehrere *Intelligenzformen* bzw. *Intelligenzdimensionen* einbeziehen (Testkatalog-Hogrefe, 2018/19):

- Wechsler Adult Intelligence Scale (WAIS-IV)
- Intelligenz-Struktur-Test 2000 R (IST 2000 R)

Einige andere Tests berücksichtigen nur einzelne Intelligenzdimensionen, wie etwa:

- Raven-Matrizen-Test (Standard Progressive Matrices; SPM)
- Dreidimensionaler Würfeltest (3DW)
- Multifactor Emotional Intelligence Scale

Merksatz

Mittels Intelligenztests wird eine quantitative und auf die Altersgruppe normierte Schätzung der geistigen Leistungsfähigkeit angestrebt. Vor einer Überinterpretation des globalen Intelligenzquotienten in einer Weise, als handle es sich um eine genetisch fixierte, angeborene Eigenschaft, muss dringend gewarnt werden.

Wichtig ist es, im Auge zu behalten, dass geistige Leistungsfähigkeit immer nur für einen bestimmten Zeitpunkt und nur für bestimmte Intelligenzformen erfasst werden kann. Die psychische Verfassung spielt dabei eine große Rolle. Zum Beispiel drückt Angst, Stress und vor allem Depression die Messwerte unweigerlich nach unten (s. *Trainingsweltmeister*). Vor einer Überinterpretation des globalen Intelligenzquotienten in einer Weise, als handle es sich um eine genetisch fixierte, angeborene Eigenschaft kann daher nicht eindringlich genug gewarnt werden. Manche Forscher (z.B. Guthke, 2003) unterscheiden überhaupt zwischen „Intelligenzanlage" (den genetischen Voraussetzungen), „Intelligenzstatus" (dem momentan erfassbaren Leistungsniveau) und „Intelligenzpotenz" (potenziell entwickelbare intellektuelle Leistungsfähigkeit).

Um auch die vielfältigen Bedingungen und Voraussetzungen für intellektuelle Leistungen richtig einschätzen zu können („interaktionistische" Sicht individueller Begabungen; s. 2.3.2), ist für eine zuverlässige Intelligenzdiagnostik auch die zusätzliche Erhebung von leistungsbezogenen Persönlichkeits- und Einstellungsmerkmalen nötig, wie zum Beispiel emotionale Labilität, Belastbarkeit und Leistungsmotivation.

Das in neuerer Zeit bevorzugte computergestützte Testen („computer adaptive/assisted testing", CAT) machte es möglich, die Testaufgaben von Leistungstests nicht mehr strikt in der gleichen Sequenz vorzugeben, sondern flexibel und maßgeschneidert (und damit zeitsparender) auf das Fähigkeitsniveau der Probandinnen und Probanden anzupassen, wie dies etwa beim AID 3 („Adaptives Intelligenz Diagnostikum 3") von Kubinger und Holocher-Ertl (2012) der Fall ist. Voraussetzung dafür ist allerdings eine probabilistische Testentwicklung („Item Response Theory") mit einer wahrscheinlichkeitstheoretischen Berechnung von Test- und Personenkennwerten („Itemparameter" und „Personenparameter", s. Roskam, 1996; Fischer & Molenaar, 1995).

Formen der Intelligenz

8.6.2

Für diagnostische Zwecke wird fast immer ein **Intelligenzprofil** erstellt, das die für die jeweilige Fragestellung (z.B. Berufseignung) interessierenden Stärken und Schwächen einer Person enthält. Die vollständige Erfassung aller intellektuellen Qualitäten eines Men-

schen ist weder theoretisch noch praktisch möglich. Die bis heute psychologisch analysierten Intelligenzformen sind äußerst vielfältig und reichen von einem durch Charles Spearman (1863–1945) postulierten „Generalfaktor" bis zum „Intelligenzstrukturmodell" von Guilford (1897–1988) mit 120 Intelligenzaspekten. Für viele solcher Intelligenzbereiche existieren auch psychometrisch entwickelte Testverfahren mit statistisch begründeten **Intelligenzdimensionen**. Häufig eingesetzte Methoden zur Gewinnung statistisch fundierter „Mess- und Strukturmodelle" (Erdfelder et al., 1996) sind die linearkombinatorische *Faktorenanalyse* und in neuerer Zeit die erwähnten wahrscheinlichkeitstheoretischen Testmodelle („Item Response Theory"; Fischer & Molenaar, 1995).

Eine bereits in den Dreißigerjahren entwickelte, statistisch fundierte Intelligenzkonzeption beruht auf der Idee sogenannter **Primärfähigkeiten** oder **Primärfaktoren** der Intelligenz („primary mental abilities"), von denen angenommen wird, dass sie in unterschied-

Tab 8.4 **„Primary Mental Abilities" (Thurstone & Thurstone, 1963)**

Benennung der Fähigkeit	Aufgabenart
1. Verbales Verständnis – Verbal comprehension	Erfassen von Wortbedeutungen (Synonyme)
2. Wortflüssigkeit – Word fluency	Lösen von Anagrammen, Bilden von Reimen
3. Rechenfähigkeit – Number	Erkennen korrekter oder inkorrekter Additionen
4. Räumliches Vorstellungsvermögen – Space	Erkennen von zweidimensionalen gedrehten Figuren
5. Merkfähigkeit – Memory	Auswendiglernen von Paarbildungen (Buchstaben und Ziffern)
6. Wahrnehmungsgeschwindigkeit – Perceptual speed	Wiedererkennen von Figuren, Bildvergleiche
7. Schlussfolgerndes Denken – Reasoning	Fortsetzen von Buchstabenreihen (Regeln erkennen)

Dieses historische Modell der sogenannten „Primärfaktoren" der Intelligenz wurde erstmals im Jahre 1934 an der University of Chicago durch statistische Analyse (Korrelations- und Faktorenanalyse) der Daten von 240 Personen anhand von 56 Testaufgaben postuliert.

licher Beteiligung die intellektuelle Leistungsfähigkeit des Menschen charakterisieren (Tab. 8.4). Bis zu 20 und mehr solcher Primärfähigkeiten wurden postuliert (Visualisierungskapazität, Informationsverarbeitungsgeschwindigkeit, Einprägungsfähigkeit, Einfallsreichtum usw.) und – ausgehend von ihren internen *Korrelationen* – zu Fähigkeitsdimensionen höherer Ordnung zusammengefasst („Comprehensive Ability Battery"; Hakstian & Cattell, 1978).

Multiple Intelligenzen | Box 8.3

Ausgehend von der Beobachtung, dass bei Gehirnschäden oft sehr spezifische Ausfälle stattfinden, dass bei geistiger Minderbegabung manchmal außergewöhnliche Spezialfähigkeiten in eng begrenzten Bereichen auftreten (z.B. Musik, Rechnen, Raumvorstellung) und dass in Testungen zahlreiche unkorrelierte Intelligenzdimensionen nachgewiesen werden konnten, geht Howard Gardner (2000) von zumindest acht voneinander unabhängigen Intelligenzformen aus:

- Sprachliche Intelligenz (Sprachverstehen, Schreiben, Reden und Lesen)
- Logisch-mathematische Intelligenz (logisches Schlussfolgern, Gleichungen lösen)
- Visuell-räumliche Intelligenz (Stadtpläne interpretieren, Gegenstände ordnen)
- Musikalische Intelligenz (Musikverständnis, Rhythmusgefühl, Gehör, Musizieren, Komponieren)
- Motorisch-kinästhetische Intelligenz (Geschicklichkeit, Körperkontrolle, Tanzen, Turnen, Ballspiele)
- Interpersonale Intelligenz (Einfühlung, Kommunikation)
- Intrapersonale Intelligenz (Selbstverständnis, Kenntnis der eigenen Stärken und Schwächen)
- Naturalistische Intelligenz (Klassifikation und Verständnis von Strukturen und Vorgängen in der Natur)

Obwohl dieses Konzept plausibel erscheint und in viele pädagogische Programme eingeflossen ist, wird es noch nicht als ausreichend empirisch evaluiert angesehen.

Zwei solcher Hauptfaktoren fanden in der Intelligenzforschung besondere Beachtung: Die **kristalline Intelligenz** („crystalized intelligence"), welche kulturabhängige Fähigkeiten erfasst (verbale und motorische Fertigkeiten, Expertenwissen), und die **fluide Intelligenz** („fluid intelligence"), welche sich auf die weitgehend kulturunabhängige kognitive Grundausstattung der geistigen Leistungsfähigkeit bezieht (z.B. induktives, kombinatorisches und figurales Denken). Die wissenschaftliche Produktivität liegt bei Fächern mit hoher Anforderung in fluider Intelligenz (Mathematik, Physik, Chemie) im Alter zwischen 25 und 40 Jahren (Salthouse, 2012).

Eine Ausweitung dieses Konzepts unabhängiger *Intelligenzdimensionen* stellt die **Theorie der multiplen Intelligenzen** von Gardner, 2000; 2006) dar (Box 8.3), bei der auch Fähigkeitsbereiche einbezogen wurden, die bisher in Intelligenztests kaum als solche Berücksichtigung fanden (z.B. Musikverständnis, Selbstkenntnis, Geschicklichkeit).

Merksatz

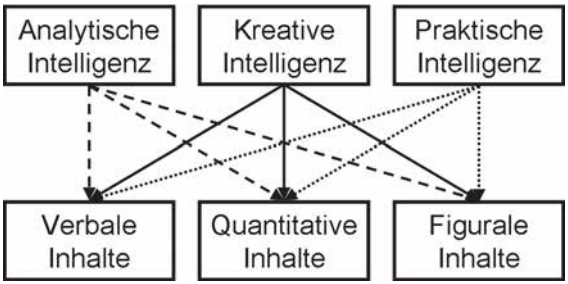
Geistige Leistungsfähigkeit manifestiert sich in sehr verschiedenen, voneinander weitgehend unabhängigen Bereichen, für die derzeit nur teilweise Testinstrumente zur Verfügung stehen.

Ein häufig geäußerter Kritikpunkt an klassischen Intelligenztests war ihre relativ geringe Prognoseleistung (ca. 25 %) für schulischen oder beruflichen Erfolg (Neisser et al., 1996) und ihre geringe Korrelation mit komplexen Problemlöseleistungen (z.B. Dörner, 1989; Salthouse, 2012). Eine besondere Annäherung der Intelligenzmessung an Alltagserfordernisse bezweckt das Konzept der **Erfolgsintelligenz** („Theory of Successful Intelligence") von Sternberg (1997). Sie wird definiert als

- Fähigkeit, im persönlichen Rahmen innerhalb des soziokulturellen Kontextes Erfolg zu haben,
- als Begabung zur Nutzung eigener Stärken und zur Kompensation eigener Schwächen,
- und erfordert eine Balance zwischen *analytischer* (schulischer, akademischer), *kreativer* und *praktischer* Intelligenz (Abb. 8.20).

Einen ebenfalls lebensnahen, in klassischen Intelligenztests kaum berücksichtigten Aspekt beschreibt der Begriff der **emotionalen Intelligenz** (Goleman, 1996), welcher die Geschicklichkeit charakterisieren soll, einerseits die eigenen Emotionen richtig einzuschätzen, mit ihnen vernünftig umzugehen und sie unmissverständlich mitzuteilen, und andererseits auch die Emotionslage anderer richtig zu interpretieren und erfolgreich zu beeinflussen.

Abb 8.20



In der triarchischen Theorie der Intelligenz („Triarchic Theory of Intelligence") von Sternberg und Kaufman (1998) wird von drei selbst wieder aus anderen Unterkomponenten bestehenden Intelligenzformen ausgegangen, die sich in je drei Inhalts- bzw. Präsentationsebenen manifestieren. Mittels einer entsprechenden Testbatterie („Sternberg Triarchic Abilities Test", STAT) konnten diese (neun) Faktoren an einer internationalen Stichprobe von mehr als 3000 Probandinnen und Probanden gut bestätigt werden (Sternberg et al., 2001). Die analytische Intelligenz einer Person (der Hauptaspekt konventioneller Intelligenztests) wird durch kreative und praktische Fähigkeiten („tacit knowledge") hier ergänzt gesehen, welche nach Sternberg in vielen Lebensbereichen mindestens ebenso wichtig sind (Sternberg, 2002).

Dass bisher kaum Formen emotionaler, sozialer oder praktischer Intelligenz in gängigen Intelligenztests vorkamen, lag wohl weniger an der Unterschätzung ihrer Alltagsrelevanz durch die Psychologinnen und Psychologen, sondern eher an den Schwierigkeiten ihrer Erhebung. Die in den entsprechenden Tests verwendeten Aufgaben (z.B. Erkennen von Stimmungen, Einfühlung in andere, Bewältigung sozialer Konfliktsituationen) lassen oft keine eindeutige Bewertung der Antworten zu und verursachen damit Probleme der *Objektivität*, *Reliabilität* und *Validität* (s. 3.7.4).

Genetische Veranlagung und Umweltfaktoren der Intelligenz

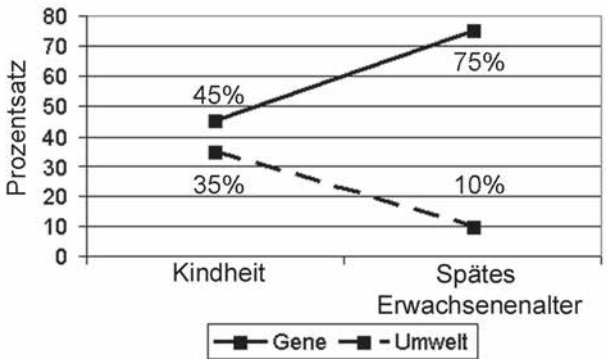
8.6.3

Die **Erblichkeitsschätzung** (Ausmaß des genetischen Einflusses auf ein Persönlichkeitsmerkmal) für Intelligenz ist von einigen verwerflichen und folgenschweren Entgleisungen in der Frühzeit der Intelligenzforschung geprägt. Ohne die Umwelt- und Kulturabhängigkeit von intellektuellen Leistungen zu beachten, wurden zum Beispiel in den USA Immigrantinnen und Immigranten jüdischer,

italienischer und russischer Herkunft, aber auch Latinos und Schwarze aufgrund angeblich niedrigerer Intelligenzwerte diskriminiert und als genetisch „minderwertig“ eingestuft, was im Jahre 1924 zu entsprechenden Einwanderungsbeschränkungen („Immigration Restriction Act“) führte.

Nachdem nicht nur unter nationalsozialistischer Herrschaft, sondern auch später immer wieder ideologisch gespeiste Vermu-

Abb 8.21



Die Einflüsse der Gene und der Umweltfaktoren (Familie, Bildung, Beruf) auf Intelligenzleistungen können mittels statistischer Verfahren geschätzt werden. Vergleiche zwischen Eltern und Adoptivkindern, Eltern und leiblichen Kindern, Geschwistern sowie eineiigen Zwillingen (unähnliches, ähnliches, gleiches Erbgut) in übereinstimmenden oder getrennten Lebensbedingungen erlauben eine Schätzung des Ausmaßes genetischer und umweltbedingter Effekte auf die Variation von Intelligenzmessungen. Nach Neisser und Mitarbeitern (1996, 85) nimmt der Einfluss der Umwelt mit dem Alter ab (35 % → 10 %) und jener der genetischen Disposition zu (45 % → 75 %).

tungen über angeblich konstante, ethnisch spezifizierbare Intelligenzausstattungen des Menschen auftauchten, hat eine Arbeitsgruppe von Intelligenzforschern („Board of Scientific Affairs“) der American Psychological Association (APA) in einem wissenschaftlichen Überblicksartikel zu wesentlichen Aspekten der geistigen Leistungsfähigkeit Stellung bezogen (Neisser et al., 1996). Demnach sind die Unterschiede im IQ zwischen US-Amerikanern asiatischer, hispanischer, indianischer und afrikanischer Herkunft minimal

(bis zu 10 Punkten) und gleichen sich im Laufe der kulturellen Eingliederung weitgehend an jene der „weißen“ Bevölkerung an.

Was die allgemeine Frage der genetischen Determiniertheit von Intelligenz betrifft, so wird aus Zwillings- und Geschwisterstudien geschlossen (s. auch Abb. 2.1 in 2.3.2), dass der Anteil an Erblichkeit für die Leistung in Intelligenztests etwa 50 % beträgt und jener für familiäre oder Umwelteinflüsse etwa 25 % (Rest ist Zufall). Allerdings nimmt mit dem Alter der Einfluss der genetischen Ausstattung zu und jener der Umwelt ab (Abb. 8.21). Dies wird dadurch erklärt, dass mit dem Alter die Selbstständigkeit wächst und somit die Chance, sich jene Umweltbedingungen auszusuchen oder zu schaffen, die der zugrunde liegenden genetischen Ausstattung bzw. den erblich bedingten Neigungen am besten entsprechen („genotyp-environment correlation“; Plomin & Spinath, 2004).

Merksatz

Der Anteil an Erblichkeit für die Leistung in Intelligenztests beträgt etwa 50 %, jener für familiäre oder Umwelteinflüsse etwa 25 %. Mit dem Alter nimmt der Einfluss der genetischen Ausstattung zu und jener der Umwelt ab.

Frauen und Männer unterscheiden sich in Gesamt-IQ-Werten im Allgemeinen nicht, sind aber in einigen Intelligenzaufgaben differenzierbar (Abb. 8.22). Insgesamt werden die gefundenen **ge-**

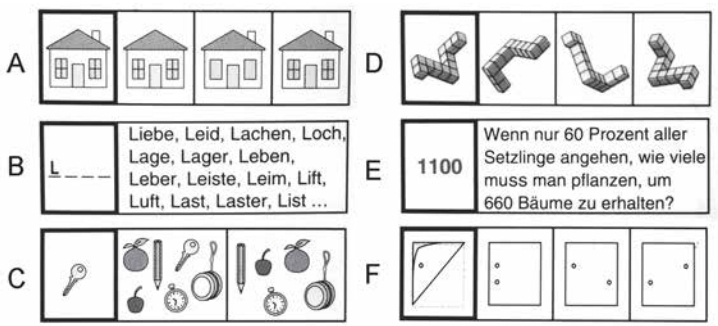


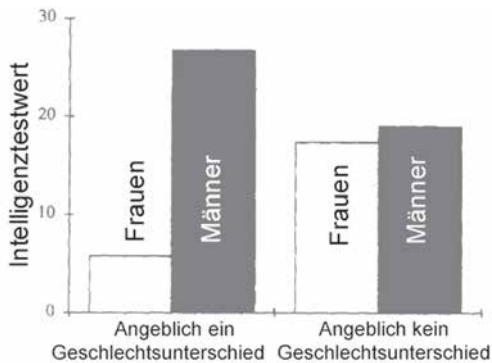
Abb 8.22

Frauen sind durchschnittlich im schnellen Identifizieren von Bildern (A), im Finden von Wörtern (B) und Vergleichen von Objektlisten (C) besser, während Männer im Allgemeinen besser rotierte Figuren identifizieren (D), mathematische Schlussfolgerungen lösen (E) und Strukturveränderungen erkennen (F).

schlechtsspezifischen Intelligenzdiskrepanzen oft übertrieben interpretiert und übermäßig verallgemeinert: Frauen schneiden tendenziell in Verbal- und Kommunikationsleistungen besser ab (Lesen, Rechtschreibung, Ideen- und Wortflüssigkeit, Interpretation der Körpersprache) und weisen weniger Lese- und Sprachstörungen auf, während Männer tendenziell bessere visuell-räumliche und mathematische Leistungen zeigen (Rotation von Strukturen in der Vorstellung, Kartenlesen, mechanische Probleme lösen; Kimura, 1993; Neisser et al., 1996).

Neben den bereits genannten Verfälschungsmöglichkeiten von Leistungstestkennwerten (z.B. durch Kulturabhängigkeit, Trainingseffekte, Umweltstimulation, nicht repräsentative Stichprobenauswähl) sind auch soziale Einflussfaktoren gefunden worden, wie etwa die Hemmung von Leistungen durch ein gesellschaftliches **Stereotyp**. Stereotype sind Kategorisierungen von Personengruppen, wobei bestimmte Konfigurationen von Eigenschaften als typisch angenommen werden. Mehr oder weniger eingestanden

Abb 8.23



In einem Experiment (Steele, 1997) nahmen Schülerinnen und Schüler einer Highschool teil, die alle als gleich gut und gleich interessiert in Mathematik angesehen werden konnten. Sie bekamen einen schwierigen Mathematiktest vorgelegt, wobei einer Gruppe gesagt wurde, dass in diesem Test Männer im Allgemeinen besser abschnitten als Frauen, während in einer zweiten Gruppe keine geschlechtsspezifischen Unterschiede suggeriert wurden. Die Ergebnisse zeigten, dass Frauen dann schlechter abschnitten, wenn sie erwarten konnten, dass der Test ihre stereotypiebedingten Schwächen erkennen ließe.

existieren solche Auswirkungen *impliziter Persönlichkeitstheorien* für alle Bevölkerungsgruppen, oft in Form gegensätzlicher Stereotype: Männer – Frauen, Schwarze – Weiße, Fremde – Einheimische etc. Stimmen nun bestimmte Eigenschaften (z.B. Mathematikfähigkeit) bei einer Person (z.B. einer Frau) nicht mit jenen überein, die hinsichtlich des einschlägigen Stereotyps erwartet werden (z.B. „Frauen sind schlecht in Mathematik"), dann resultiert für die Person eine subjektive Bedrohung der Zugehörigkeit zu ihrer Bezugsgruppe. Sie wird infolgedessen entweder ihre Leistungen den stereotypen Erwartungen anpassen oder eine Abschwächung der Identifikation mit dem Leistungsbereich vornehmen (Steele, 1997). Leistungshemmungen, die durch solche **Stereotypbedrohungen** („stereotype threats") erklärbar sind, konnten bei Frauen in Mathematik und bei Afroamerikanern in Verbaltests eindeutig nachgewiesen werden (Steele, 1997; Steele & Aronson, 1995; Abb. 8.23). Sie treten aber wahrscheinlich in allen Situationen auf, in denen aufgrund sozialer Stereotype negative Leistungserwartungen erzeugt werden.

Zusammenfassung

Der Zweck jeglichen menschlichen Handelns ist im Grunde die kurzfristige oder langfristige Adaptation an Umweltgegebenheiten. Probleme können daher psychologisch als Situationen definiert werden, in denen die gewünschte Überführung von einem Ist- in einen Soll-Zustand mit Aufwand verbunden ist. Die Lösung eines Problems wird durch Informationen über die Problemsituation und ihre Veränderungsmöglichkeiten (über den Problemraum) erleichtert. Theoretisch werden als Grundlage des Problemlösens mentale Repräsentationen (kognitive Abbildungen) des Problemraumes angenommen, um korrektes oder fehlerhaftes Problemlöseverhalten zu erklären.

Wichtige psychische Teilprozesse des Problemlösens sind Schlussfolgerungen, welche entweder induktiv ablaufen, indem aus Einzelerfahrungen auf Regeln geschlossen wird, oder deduktiv, indem aus Prämissen implikative Ableitungen durchgeführt werden. Letztere unterteilen sich wieder in Konditionalschlüsse (Wenn-dann-Formulierungen) und in Kategorialschlüsse (Mengenaussagen). Schlussfolgerungen sind fehleranfällig, besonders wenn