Kritisch denken mit wissenschaftlicher Psychologie

David G. Myers

2.1 Die Notwendigkeit der Psychologie als Wissenschaft – 18
2.1.1 Wussten wir das schon lange? Verzerrung durch nachträgliche Einsicht (Hindsightbias) – 19
2.1.2 Übertriebene Selbstsicherheit – 20
2.1.3 Wahrnehmung von Ordnung bei zufälligen Ereignissen – 21
2.1.4 Die wissenschaftliche Haltung: Neugierig, skeptisch und bescheiden – 22
2.1.5 Kritisches Denken – 24

2.2 Wie stellen und beantworten Psychologen Fragen? – 25
2.2.1 Die wissenschaftliche Methode – 25
2.2.2 Beschreibung – 26
2.2.3 Korrelation – 30
2.2.4 Experiment – 33

2.3 Statistische Argumentation im Alltagsleben – 36
2.3.1 Datenbeschreibung – 36
2.3.2 Signifikante Unterschiede – 39

2.4 Häufig gestellte Fragen zur Psychologie – 41

2.5 Kapitelrückblick – 46
2.5.1 Verständnisfragen – 46
2.5.2 Schlüsselbegriffe – 46
2.5.3 Weiterführende deutsche Literatur – 47

D. G. Myers, *Psychologie*,
DOI 10.1007/978-3-642-40782-6_2, © Springer-Verlag Berlin Heidelberg 2014

Millionen von Menschen interessieren sich für »Psychologie«, weil sie neugierig auf Menschen sind und weil sie hoffen, ihre eigenen kleinen Leiden heilen zu können. Sie hören sich psychologische Beratungsgespräche im Radio an, lesen Artikel über die Kräfte der Seele, nehmen an Hypnoseseminaren zur Raucherentwöhnung teil, tauchen ein in Selbsthilfe-Internetseiten und verschlingen Selbsthilfebücher über die Bedeutung der Träume, den Pfad zur ekstatischen Liebe und darüber, wie man persönliches Glück erlangt.

Andere wiederum fragen sich angesichts mancher psychologischer Wahrheiten: Stimmt es, dass die Bindung zwischen Mutter und Kind in den ersten Stunden nach der Geburt entsteht? Wie – und in welchem Ausmaß – prägt die Erziehung die Persönlichkeit und Fähigkeiten eines Kindes? Sind erstgeborene Kinder stärker leistungsorientiert? Kann Psychotherapie heilen?

Wie können wir bei solchen Fragen simple Meinungen von stichhaltigen Schlussfolgerungen unterscheiden? *Wie können wir die Psychologie so sinnvoll einsetzen, dass wir verstehen, warum die Menschen so und nicht anders denken, fühlen und handeln?*

2.1 Die Notwendigkeit der Psychologie als Wissenschaft

2.1 Wie illustrieren der Verzerrungseffekt durch nachträgliche Einsicht (Hindsightbias), die übertriebene Selbstsicherheit und die Tendenz, in zufälligen Ereignissen eine Ordnung wahrzunehmen, weshalb wissenschaftlich fundierte Antworten der Intuition und dem gesunden Menschenverstand überlegen sind?

Manche Menschen sagen, Psychologie sei letztlich nichts anderes als in einen Fachjargon gepresstes Allgemeinwissen. »Gibt es etwas Neues unter der Sonne? Ihr werdet dafür bezahlt, dass ihr mit ausgefallenen Methoden das beweist, was meine Großmutter schon immer wusste.« Andere Menschen glauben an die intuitiven Fähigkeiten des Menschen: »Tief drin schlummert in jedem von uns ein instinktives, inniges Bewusstsein, das uns – wenn wir es zulassen – die verlässlichste Führung gibt«, sagte Prinz Charles (2000). Der frühere US-Präsident George W. Bush beschrieb dem Journalisten Bob Woodward (2002) dieses Gefühl, um zu erklären, wieso er den Irak-Krieg eröffnet habe: »Ich bin ein Bauchmensch. Ich verlasse mich auf meine Instinkte.«

Prinz Charles und der ehemalige Präsident Bush befinden sich in guter Gesellschaft, wenn man die lange Liste an populären Psychologiebüchern zu »intuitivem Management«, »intuitivem Handel« und »intuitiver Heilung« betrachtet. Die heutige psychologische Wissenschaft zeigt, dass Intuition tatsächlich eine große Rolle spielt. Wie wir noch sehen werden, bewegen sich unser Denken, unser Gedächtnis und unsere Einstellungen auf zwei Ebenen – einer bewussten und einer unbewussten –, wobei der größere Teil dieser Prozesse automatisch abläuft, sozusagen »off-screen«. Wie Jumbojets fliegen wir meistens mit Autopilot.

Abb. 2.1 Die Grenzen der Intuition. Personalchefs neigen dazu, ihren »Bauchgefühlen« bei der Beurteilung von Stellenbewerbern zu sehr zu vertrauen, zum einen, weil sie sich an die Fälle erinnern, in denen sich ihr guter Eindruck als richtig erwies, zum anderen, weil sie nicht wissen, dass ein von ihnen abgelehnter Bewerber in einer anderen Firma erfolgreich war. (© contrastwerkstatt / Fotolia)

Handeln wir also klug, wenn wir auf das Flüstern unserer inneren Weisheit hören, uns einfach auf die »Kraft in uns« verlassen? Oder sollten wir unsere intuitiven Ahnungen öfters skeptisch überprüfen?

So viel steht fest: Oftmals unterschätzen wir, wie fehleranfällig die Intuition ist. Meine geografische Intuition sagt mir, dass Reno östlich von Los Angeles liegt, dass sich Rom südlich von New York und Atlanta östlich von Detroit befindet. Aber das ist falsch, falsch und nochmals falsch.

Die in den folgenden Kapiteln dargestellten Experimente zeigen, dass Menschen ihre Fähigkeit, Lügen zu entlarven, als besser wahrnehmen als sie eigentlich ist. Ebenso überschätzen Menschen die Genauigkeit ihrer Augenzeugenberichte, ihre Beurteilungen von Stellenbewerbern (Abb. 2.1), ihre Gefahrenvorhersagen oder ihr Talent bei der Auswahl von Aktien. »Das erste Prinzip«, sagte Richard Feynman (1997), »besteht darin, dass du dich selbst nicht hinters Licht führen darfst – und du bist die Person, welche am leichtesten zu täuschen ist.«

Tatsächlich beobachtete die Schriftstellerin Madeleine L'Engle (1973): »Der bloße Intellekt ist ein ausgesprochen ungenaues Werkzeug.« Drei Phänomene – der *Verzerrungseffekt durch nachträgliche Einsicht (Hindsightbias)*, die *Überschätzung unserer Urteilsfähigkeit* und unsere *Tendenz, in zufälligen Ereignissen eine Ordnung wahrzunehmen* – machen deutlich, warum wir uns nicht nur auf Intuition und gesunden Menschenverstand verlassen können.

» Wer sich auf seinen Verstand verlässt, ist ein Tor.
Sprüche 28, 26

2.1.1 Wussten wir das schon lange? Verzerrung durch nachträgliche Einsicht (Hindsightbias)

Es ist leicht, schlau zu sein und das Schwarze erst dann um den Pfeil zu malen, wenn er schon getroffen hat. Nach jedem Abwärtstrend der Börse sagen die Investmentgurus, die Börse sei doch ganz offensichtlich überreif für eine Korrektur gewesen. Nach dem Fußballspiel loben wir den Coach für das »kämpferische Spiel«, welches zum Sieg führte. Hat die Mannschaft jedoch keinen Erfolg, machen wir den Coach hinterher für das »lahme Spiel« verantwortlich. Nach einem Krieg oder einer Wahl scheint uns das jeweilige Ergebnis offensichtlich. Obwohl uns die Geschichte im Rückblick wie eine Folge unvermeidlicher Ereignisse vorkommen kann, wird die tatsächliche Zukunft selten vorhergesehen. Keine Person hat je in ihrem Tagebuch festgehalten: »Heute begann der 100-jährige Krieg.«

» Wir leben das Leben vorwärts, aber wir verstehen es rückwärts.
Der Philosoph Søren Kierkegaard (1813–1855)

Der **Hindsightbias** (auch bekannt als *Rückschaufehler*) ist leicht zu demonstrieren: Geben Sie der Hälfte der Mitglieder einer Gruppe einen angeblich wissenschaftlichen Befund, während Sie der anderen Hälfte das genaue Gegenteil präsentieren. Zu der ersten Gruppe sagen Sie: »Psychologen haben herausgefunden, dass Trennung die romantische Anziehung abschwächt. Das Sprichwort sagt es ja auch: ›Aus den Augen, aus dem Sinn‹«. Dann bitten Sie die Teilnehmer, darüber nachzudenken, warum das so ist. Die meisten Menschen können und werden an diesem wahren Befund nichts Erstaunliches finden.

Der zweiten Gruppe erzählen Sie das genaue Gegenteil, nämlich: »Psychologen haben herausgefunden, dass romantische Anziehung durch Trennung stärker wird. Wie das Sprichwort sagt: ›Liebe wächst mit der Entfernung‹«. Die Teilnehmer werden auch dieses unwahre Resultat mühelos erklären, und die überwiegende Mehrheit wird nicht überrascht sein. Wenn zwei gegensätzliche Befunde mit dem »gesunden Menschenverstand« erklärt werden können, dann haben wir ein Problem.

» Alles scheint ein Gemeinplatz zu sein, wenn es erst einmal erklärt ist.
Dr. Watson zu Sherlock Holmes

Hindsightbias (Verzerrung durch nachträgliche Einsicht, »hindsight bias«) – Tendenz, nach dem Eintreten eines Ereignisses zu glauben, man hätte es vorhersehen können (auch bekannt als *Rückschaufehler*).

Derartige Irrtümer bei unserem Erinnerungsvermögen und unseren Erklärungsversuchen machen deutlich, weshalb wir psychologische Forschung brauchen. Die Antwort auf die Frage, was ein Mensch gefühlt und warum er so und nicht anders gehandelt hat, kann irreführend sein, und zwar nicht deshalb, weil der gesunde Menschenverstand uns immer in die Irre schickt, sondern weil die Antwort *nach* dem Ereignis gegeben wird. Mit dem gesunden Menschenverstand lässt sich viel leichter beschreiben, was passiert *ist,* als was passieren *wird*. So sagte der Physiker Niels Bohr angeblich: »Vorhersagen sind recht schwierig, vor allem wenn es um die Zukunft geht.«

Der Hindsightbias wurde in etwa 100 Studien sowohl bei Kindern als auch bei Erwachsenen in vielen unterschiedlichen Ländern beobachtet (Blank et al. 2007; ◘ Abb. 2.2, ◘ Abb. 2.3). Dennoch hat der gesunde Menschenverstand häufig Recht. Yogi Berra sagte einmal: »Du kannst vieles beobachten, wenn du genau hinschaust.« (Wir müssen Berra für weitere Diamanten aus der Sprüchekiste danken, wie beispielsweise »Niemand kommt je hierher – es ist ein zu großes Gedränge« und »Wenn die Leute nicht raus auf den Baseballplatz kommen wollen, wird sie niemand daran hindern können.«) Wir alle sind Verhaltensbeobachter. Deshalb wäre es schon erstaunlich, wenn nicht zumindest ein paar Ergebnisse der psychologischen Forschung schon vorher bekannt gewesen wären. Viele Menschen glauben, dass Liebe Glück hervorbringt – und sie haben Recht (wir alle haben ein tiefes »Bedürfnis nach Zugehörigkeit«, wie wir in ► Kap. 12 erfahren werden). Daniel Gilbert, Brett Pelham und Douglas Krull (2003) merken sogar an: »Gute Ideen aus dem Bereich der Psychologie kommen uns oft seltsam vertraut vor. Und in dem Augenblick, in dem wir auf sie stoßen, meinen wir, sicher zu sein, dass wir schon einmal nahe daran waren, das Gleiche zu denken, und es einfach nur nicht geschafft haben, den Gedanken niederzuschreiben.« Gute Ideen sind wie gute Erfindungen; einmal vorhanden, sind sie nicht mehr wegzudenken (Warum hat es so lange gedauert, bis jemand Reisekoffer mit Rollen oder »Post-its« erfunden hat?).

Obwohl der gesunde Menschenverstand und unsere Intuition auf Informationen beruhen, die wir aus unzähligen beiläufigen Beobachtungen gewonnen haben, liegen sie manchmal völlig daneben. In späteren Kapiteln werden wir sehen, wie die Wissenschaft verbreitete Vorstellungen über den Haufen geworfen hat – dass Aufdringlichkeit zu Geringschätzung führt, dass Träume die Zukunft vorhersagen oder dass die meisten von uns nur 10% des Gehirns nutzen. Wir werden auch sehen, welche Überraschungen uns die Forschung mit ihrer Entdeckung der chemischen Botenstoffe des Gehirns bereitet, die unsere Stimmungen und unsere Erinnerungen steuern, oder mit den Forschungsergebnissen dazu, über welche Fähigkeiten Tiere verfügen und wie sich Stress auf unsere Fähigkeit zur Krankheitsabwehr auswirkt.

Abb. 2.2 Hindsightbias. Bei den Arbeiten an der Ölbohrplattform Deepwater Horizon im Jahr 2010 nahmen BP-Angestellte einige »Abkürzungen« und ignorierten ein paar Warnschilder. Dabei hatten sie nicht die Absicht, ihre Arbeitskollegen oder die Umwelt gravierenden Risiken auszusetzen. *Nach* der Ölkatastrophe wurde – mit dem Vorteil der späteren Einsicht – die Dummheit dieser Entscheidungen offensichtlich. (© EPA / US COAST GUARD / HANDOUT/ picture alliance / dpa)

Abb. 2.3 Hindsightbias (© Claudia Styrsky)

2.1.2 Übertriebene Selbstsicherheit

Beispiele für übertriebene Selbstsicherheit gibt es in der Geschichte viele:

» Ihr Sound gefällt uns nicht. Gitarrengruppen sind nicht mehr gefragt.
Erklärung von Decca Records, warum sie mit den Beatles 1962 keinen Plattenvertrag schließen wollten

» Die Computer der Zukunft werden wahrscheinlich nicht einmal eineinhalb Tonnen wiegen.
Popular Mechanics, 1949

» Sie könnten nicht einmal einen Elefanten treffen auf diese Entfernung.
Die letzten Worte des Generals John Sedgwick vor seinem Tod in einer Schlacht des Amerikanischen Bürgerkriegs, 1864

» Für unsere amerikanischen Vettern mag das Telefon ja eine nützliche Erfindung sein, aber nicht für uns. Wir haben genügend Botenjungen.
Urteil einer britischen Expertengruppe über die Erfindung des Telefons

Wir Menschen neigen dazu, zu glauben, wir wüssten mehr, als wir tatsächlich wissen. Auf die Frage, wie sicher wir sind, die richtigen Antworten auf Sachfragen zu wissen (*Liegt Boston nördlich oder südlich von Paris?*) antworten wir eher mit Selbstvertrauen als mit korrektem Wissen.[1] Schauen Sie sich einmal die drei folgenden Anagramme an:

Serwas → Wasser
Tessmy → System
Hartox → Thorax

Was glauben Sie, wie viele Sekunden Sie etwa gebraucht hätten, um jedes dieser Anagramme aufzulösen? Sobald man die Lösung kennt, sorgt die nachträgliche Einsicht dafür, dass sie uns absolut selbstverständlich erscheint. Das führt zu übertriebenem Selbstvertrauen. Wir glauben, wir hätten die Lösung in höchstens 10 Sekunden gefunden, während tatsächlich der Durchschnitt bei 3 Minuten liegt. Und diese 3 Minuten hätten Sie auch gebraucht, wenn Sie die Lösung nicht gekannt hätten. Probieren Sie es mit einem weiteren Anagramm aus: ACHENFI.[2]

Sind wir besser, wenn es darum geht, soziales Verhalten vorherzusagen? Der Psychologe Philip Tetlock (1998, 2005) von der University of Pennsylvania sammelte mehr als 27.000 Vorhersagen von Experten zu Weltereignissen, beispielsweise zur Zukunft von Südafrika oder zur Frage, ob sich Quebec von Kanada unabhängig machen wird. Sein wiederholter Befund: Experten, die angegeben hatten, ihrer Sache zu durchschnittlich 80% sicher zu sein, hatten in weniger als 40% der Fälle Recht behalten. Nichtsdestotrotz behielten sogar die Experten mit den falschen Prognosen ihr Selbstvertrauen, indem sie anmerkten, »beinahe Recht« gehabt zu haben. »Die Separatisten in Quebec haben das Sezessionsreferendum *beinahe* gewonnen.«

Prüfen Sie Ihr Wissen

- Wenn Freunde von uns anfangen, miteinander auszugehen, haben wir oft das Gefühl, wir hätten von Anfang an *gewusst*, dass die beiden füreinander bestimmt sind. Wieso ist das so?

1 Boston befindet sich südlich von Paris.
2 Die Lösung des Anagramms: EINFACH.

2.1.3 Wahrnehmung von Ordnung bei zufälligen Ereignissen

In unserer naturgegebenen Bereitschaft, unserer Welt einen Sinn zu verleihen – eine Eigenschaft, die der Dichter Wallace Stevens unsere »Ordnungswut« nennt – neigen wir dazu, Muster wahrzunehmen. Menschen sehen ein Gesicht im Mond, hören satanische Botschaften in Musikstücken oder erblicken das Abbild der Jungfrau Maria auf einem gegrillten Käsesandwich. Selbst in zufällig zusammengewürfelten Informationen finden wir oftmals Ordnung, denn – und damit müssen wir uns wohl oder übel abfinden – *eine zufällige Abfolge von Daten sieht häufig nicht zufällig aus* (Falk et al. 2009; Nickerson 2002, 2005). Tatsächlich treten in Zufallssequenzen Muster oder Reihen (wie wiederholte Ziffern) öfter auf, als die Menschen glauben (Oskarsson et al. 2009). Um mir das einmal selbst zu demonstrieren – und Sie können das natürlich selbst auch probieren –, habe ich eine Münze 51-mal geworfen. Hier die Ergebnisse:

1. K	11. Z	21. Z	31. Z	41. K	51. Z
2. Z	12. K	22. Z	32. Z	42. K	
3. Z	13. K	23. K	33. Z	43. K	
4. Z	14. Z	24. Z	34. Z	44. K	
5. K	15. Z	25. Z	35. Z	45. Z	
6. K	16. K	26. Z	36. K	46. K	
7. K	17. Z	27. K	37. Z	47. K	
8. Z	18. Z	28. Z	38. Z	48. Z	
9. Z	19. K	29. K	39. K	49. Z	
10. Z	20. K	30. Z	40. Z	50. Z	

Beim Anschauen dieser Ergebnisse springen einem Muster in die Augen: Die Würfe 10 bis 22 haben ein fast perfektes Muster von Paaren aus Kopf und Zahl. Von 30 bis 38 hatte ich eine »Pechsträhne« und warf bei 9 Würfen nur einmal Kopf. Aber dann war mir Fortuna gnädig, und ich warf bei den nächsten 9 Würfen 7-mal Kopf. Ähnliche Glücks- und Pechsträhnen kommen in etwa so häufig, wie man es bei zufälligen Folgen erwarten würde, beim Basketball, beim Fußball und beim Zusammenstellen eines Portfolios von Kapitalanlagen vor (Gilovich et al. 1985; Malkiel 2007; Myers 2002). Ob es nun um den Münzwurf, um Basketball oder darum geht, die Leistung eines Anlageberaters zu überprüfen, Zufallsfolgen sehen oft nicht zufällig aus und werden deshalb oft überinterpretiert. (»Er hat eine Glückssträhne, lassen wir ihn den Elfmeter ausführen.«)

Was beweist nun dieses Muster von Glücks- und Pechsträhnen? Habe ich meine Münze auf irgendeine paranormale Art gesteuert? Bin ich aus meiner Pechsträhne mit Zahlwürfen in eine Glückssträhne mit Kopftreffern gerutscht? Aber eine Erklärung ist gar nicht erforderlich, denn das sind genau die Zahlen, die man bei solchen vom Zufall gesteuerten Daten antrifft. Vergleicht man jeden Wurf mit dem nächsten, dann ergibt sich, dass 24 der 50 Würfe jeweils ein anderes Ergebnis erbrachten: Das ist genau das beinahe 50:50-Ergebnis, das beim Münzwurf zu erwarten ist. Trotz des scheinbaren Musters gibt das Ergebnis

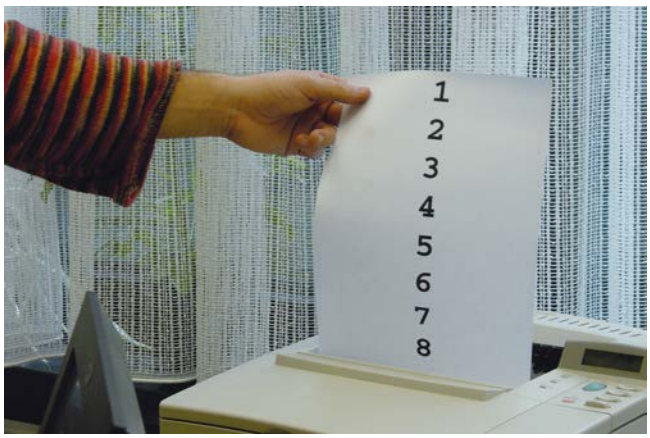

Abb. 2.4 Bizarre Sequenz computergenerierter Zufallszahlen. Sieht zwar merkwürdig aus, ist aber tatsächlich nicht unwahrscheinlicher als jede andere Zahlenfolge. (© Renate Schulz)

Abb. 2.5 In einer ausreichend großen Reihe zufälliger Geschehnisse treten Ereignisabfolgen auf, die uns merkwürdig erscheinen. Während der Fußballweltmeisterschaft 2010 stellte man einem deutschen Oktopus – Paul, »dem Orakel von Oberhausen« – zwei Boxen zur Verfügung. Beide enthielten Muscheln und auf der einen Seite eine Nationalflagge. Paul entschied sich bei 8 Spielen 8-mal für die richtige Box. Er sagte damit den Ausgang der 7 Spiele mit deutscher Beteiligung sowie Spaniens Triumph im Finale korrekt voraus. (© Roland Weihrauch dpa / picture alliance)

eines Wurfes keinerlei Hinweis auf das Ergebnis des folgenden Wurfes.

Manche Dinge sind jedoch so ungewöhnlich, dass wir nicht ohne Weiteres eine zufallsbezogene Erklärung akzeptieren können (wie es beim Münzenwerfen doch der Fall war). Statistiker finden das allerdings weniger geheimnisvoll. Als Evelyn Marie Adams zweimal die New-Jersey-Lotterie gewann, meldeten die Zeitungen, dass die Quote, dieses Kunststück zu vollbringen, bei 1 zu 17 Billionen liegt. Seltsam? Nun, die Chance von 1 zu 17 Billionen ist tatsächlich genau die Chance, mit der eine bestimmte Person, die zwei Lose der New-Jersey-Lotterie kauft, zweimal gewinnt. Die Statistiker Stephen Samuels und George McCabe (1989) berechneten, dass bei den Millionen Menschen, die ein Los der staatlichen Lotterie kaufen, ziemlich sicher irgendeiner irgendwann zweimal einen staatlichen Jackpot knacken wird. Tatsächlich können die unglaublichsten Dinge geschehen, sagen die Statistiker Persi Diaconis und Frederick Mosteller (1989), »wenn die Stichprobe nur groß genug ist«. Ein Ereignis, das täglich nur einen Menschen aus einer Milliarde betrifft, kommt immerhin 7-mal täglich vor, das heißt, 2500-mal pro Jahr (Abb. 2.4, Abb. 2.5).

» Ein wirklich außergewöhnlicher Tag wäre ein Tag ohne außergewöhnliche Ereignisse.
Statistiker Persi Diaconis (2002)

Merken Sie sich: Hindsightbias, übertriebene Selbstsicherheit und die Wahrnehmung von Ordnung bei zufälligen Ereignissen bringen uns dazu, unsere Intuition zu überschätzen. Aber die wissenschaftliche Forschung kann uns dazu verhelfen, Realität und Täuschung voneinander zu unterscheiden.

2.1.4 Die wissenschaftliche Haltung: Neugierig, skeptisch und bescheiden

? **2.2** Wie hängen die drei Hauptkomponenten der wissenschaftlichen Haltung mit dem kritischen Denken zusammen?

Was aller Wissenschaft in erster Linie zugrunde liegt, sind hartnäckige *Neugier* und die Leidenschaft, Dinge zu erforschen und zu verstehen. Dabei will man weder Irrtümer in die Welt setzen noch ihnen erliegen. Manche Fragen gehen über die Wissenschaft hinaus *(Gibt es ein Leben nach dem Tod?)*. Die Antwort auf solche Fragen erfordert auch ein Stück Glauben. Wie bei vielen anderen Ideen *(Gibt es Menschen mit übersinnlicher Wahrnehmung? Und lässt sich dies nachweisen?)* ist auch hier die Frage des Beweises ausschlaggebend. Lassen Sie die Fakten für sich selbst sprechen.

Der Zauberer James Randi bedient sich dieser *empirischen Methode*, wenn er die auf den Prüfstand stellt, die behaupten, die Aura des Menschen sehen zu können (Abb. 2.6):

Randi: Sehen Sie eine Aura um meinen Kopf?
Aura-Seher: Ja, ich sehe tatsächlich Ihre Aura.
Randi: Und wenn ich diese Zeitschrift vor mein Gesicht halte, können Sie die Aura dann immer noch sehen?
Aura-Seher: Natürlich.
Randi: Wenn ich mich also hinter diese Wand stelle, die kaum höher ist als ich, dann könnten Sie feststellen, wo ich stehe, weil die Aura über meinem Kopf zu sehen ist. Richtig?

Nach Randis Aussage war noch nie ein Aura-Seher bereit, sich dieser Prüfung zu unterziehen.

Ganz gleich, wie sinnvoll oder wie verrückt eine Idee auch sein mag: Die hartnäckige Frage lautet: *Klappt es?* Lässt sich das, was vorhergesagt wird, durch Überprüfung bestätigen? Manch-

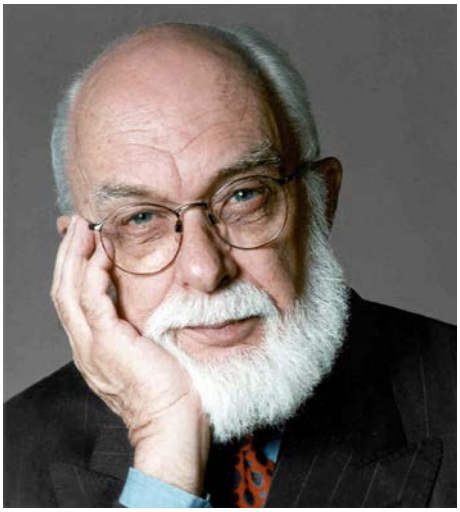

Abb. 2.6 Der erstaunliche Randi. Der Zauberer James Randi zeigt beispielhaft die skeptische Haltung. Er hat eine Vielfalt übersinnlicher Phänomene überprüft und als Tricks entlarvt. (© terabyte, this work with the title »Randi.jpg« [http://commons.wikimedia.org/wiki/File:RANDI.jpg?uselang=de] is licenced under CC-BY-SA-2.0-DE [http://creativecommons.org/licenses/by-sa/2.0/de/deed.de], no modifications were made)

mal finden verrückt klingende Ideen Unterstützung, wenn sie einer so genauen Prüfung unterzogen werden. Im 18. Jahrhundert spotteten Gelehrte über die Idee, dass Meteoriten von außerhalb der Erde kämen. Als zwei Forscher aus Yale die gängige Meinung in Frage stellten, höhnte Thomas Jefferson: »Meine Herren, eher glaube ich, dass diese zwei Yankee-Professoren lügen, als dass Steine vom Himmel gefallen sind.« Manchmal verwandelt wissenschaftliche Forschung Buhrufe in Beifall.

Doch meistens nutzt die Gesellschaft die Wissenschaft zur Abfallbeseitigung: Verrückt klingende Ideen werden auf den Müllhaufen der vergessenen Behauptungen geworfen, wo sich bereits angebliche Perpetuum-mobile-Maschinen, Wunderheilmittel gegen Krebs und körperlose Reisen in längst vergangene Jahrhunderte befinden. Wenn wir Phantasie von Wirklichkeit und Sinn von Unsinn unterscheiden wollen, brauchen wir die Einstellung eines Wissenschaftlers: Skepsis ohne Zynismus und Offenheit ohne Leichtgläubigkeit.

Ein polnisches Sprichwort sagt: »Wer Gewissheit im Glauben will, muss mit Zweifeln anfangen.« Die wissenschaftlich tätigen Psychologen betrachten Verhalten mit *neugieriger Skepsis*. Sie stellen ständig zwei Fragen: »*Was meinen Sie damit?*« und »*Woher wissen Sie das?*«.

In der Arena der miteinander konkurrierenden Ideen kann ein skeptischer Test aufzeigen, welche Idee am besten zu den Fakten passt (Abb. 2.7). Übt das Verhalten der Eltern einen entscheidenden Einfluss auf die sexuelle Orientierung ihrer Kinder aus? Kann ein Astrologe aufgrund der Planetenposition im Augenblick Ihrer Geburt Ihre Zukunft vorhersagen? Ist die Elektrokrampftherapie (die Verabreichung elektrischer Stromstöße an das Gehirn) eine effiziente Therapiemethode bei schweren Depressionen? Wie wir sehen werden, hat eine genaue Überprüfung dieser Behauptungen Psychologen dazu gebracht, die ersten beiden Fragen mit *Nein*, die dritte mit *Ja* zu beantworten.

Die praktische Umsetzung der wissenschaftlichen Haltung erfordert nicht nur Neugier und Skepsis, sondern auch *Bescheidenheit* – ein Bewusstsein für unsere eigene Anfälligkeit zu irren und eine Offenheit für Überraschungen und neue Perspektiven. Nicht meine oder Ihre Meinung zählt bei der abschließenden Analyse, sondern die Wahrheiten, die die Natur auf unser Forschen hin preisgibt. Wenn sich die Menschen oder andere Tiere nicht so verhalten, wie unsere Vorstellungen das vorhersagen, dann ist das Pech für unsere Vorstellungen. Das ist die bescheidene Einstellung, die in einem frühen Motto der Psychologie zum Ausdruck kommt: »Die Ratte hat immer Recht.«

» Ich bin ein Skeptiker, nicht weil ich nicht glauben möchte, sondern weil ich *wissen* möchte. Ich glaube, dass die Wahrheit irgendwo da draußen existiert. Aber wie können wir den Unterschied erkennen zwischen dem, was wir gerne als wahr erkennen würden und dem, das wirklich wahr ist? Die Antwort ist Wissenschaft.
Michael Shermer, »I Want to Believe«, *Scientific American*, 2009

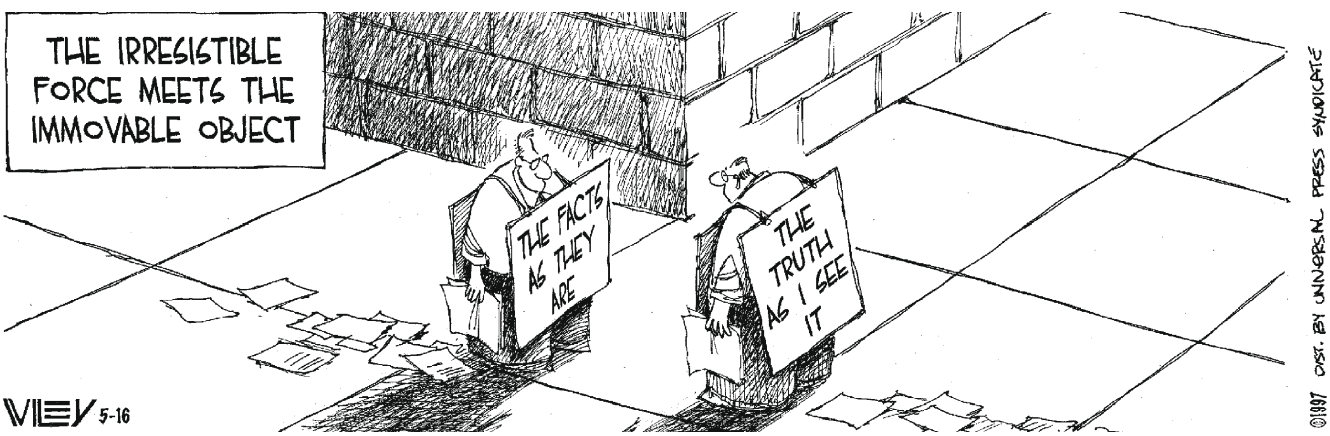

Abb. 2.7 (NON SEQUITUR © 1997 Wiley Ink, Inc.. Dist. By UNIVERSAL UCLICK. Reprinted with permission. All rights reserved.)

Wissenschaftshistoriker sagen, dass die moderne Wissenschaft überhaupt erst durch diese neugierige, skeptische und gleichzeitig bescheidene Haltung möglich wurde. Tief religiöse Menschen von heute sehen die Wissenschaft, einschließlich der wissenschaftlichen Psychologie, als Bedrohung. Und doch ist die wissenschaftliche Revolution meist von tief religiösen Menschen – dazu gehörten Kopernikus und Newton – angeführt worden, die nach der religiösen Vorstellung handelten, dass man, »um Gott zu lieben und zu achten, die Wunder seiner Schöpfung auch ganz würdigen muss« (Stark 2003a, b).

Natürlich haben Wissenschaftler, wie alle Menschen, ein großes Ego und hängen manchmal sehr an ihren vorgefassten Meinungen. Gleichwohl teilen Psychologen das Ideal der neugierigen, skeptischen und bescheidenen Haltung bei der Prüfung der Gedanken und Vorstellungen, die miteinander konkurrieren. So überprüft die Gemeinschaft der Wissenschaftler immer wieder die Befunde und Schlussfolgerungen der Kollegen.

» Ich glaube zutiefst, dass, wenn irgendein Gott der traditionellen Sorte existiert, unsere Neugier und Intelligenz durch solch einen Gott gegeben ist. Wir würden diese Geschenke nicht schätzen … wenn wir unsere Leidenschaft unterdrückten, das Universum und uns selbst zu erforschen.
Carl Sagan, *Broca's Brain*, 1979

2.1.5 Kritisches Denken

Eine wissenschaftliche Haltung befähigt uns, klüger zu denken. Kluges oder **kritisches Denken** unterzieht Vorannahmen einer Prüfung, erkennt verborgene Werte, überprüft Beweise auf ihre Richtigkeit hin und erfasst daraus resultierende Schlussfolgerungen. Ob beim Lesen eines Berichts oder beim Anhören eines Gesprächs: Ein kritischer Denker stellt Fragen. Er verhält sich wie ein Wissenschaftler: Woher weiß man das, was man da berichtet? Welche Ziele verfolgt die Person mit dem, was sie sagt? Beruht die Schlussfolgerung auf persönlichen Gefühlen und anekdotischen Berichten oder gibt es einen Beweis? Erlaubt dieser Beweis eine Schlussfolgerung auf Ursache und Wirkung? Welche alternativen Erklärungen wären möglich?

Kritisches Denken (»critical thinking«) – eine Art zu denken, die Argumente und Schlussfolgerungen nicht einfach blindlings akzeptiert. Stattdessen werden Vorannahmen einer Prüfung unterzogen, Abweichungen werden aufgedeckt, Beweise auf ihre Richtigkeit hin überprüft und daraus resultierende Schlussfolgerungen werden erfasst.

Kritisches Denken – auf wissenschaftlicher Basis – lässt unsere durch Voreingenommenheit gefärbten Linsen klar werden. Denken Sie einmal nach: Bedroht der Klimawandel unsere Zukunft und, wenn ja, wird er vom Menschen verursacht? 2009 interpretierten Klimaschützer eine australische Hitzewelle und Sandstürme als Belege für den Klimawandel. 2010 führten Klimawandelskeptiker die bittere Kälte in Nordamerika und die Schneestürme an der Ostküste an, um die globale Erwärmung in Frage zu stellen. Statt dass sie ihr Verständnis des Klimawandels vom Tageswetter beeinflussen lassen oder von ihren eigenen politischen Ansichten, verlangen kritische Denker Beweise. Erwärmt sich die Erde langfristig betrachtet tatsächlich? Schmelzen die polaren Eiskappen? Verändern sich Vegetationsmuster? Und produziert die menschliche Aktivität Gase, die uns solche Veränderungen erwarten lässt? Wenn sie über solche Themen nachdenken, werden kritische Denker die Glaubwürdigkeit von Quellen berücksichtigen. Sie werden die Beweislage anschauen (*»Werden sie von Fakten unterstützt oder erfinden sie das Ganze bloß?«*). Sie werden vielfältige Perspektiven erkennen. Und sie werden sich Informationsquellen aussetzen, die ihre vorgefassten Meinungen in Frage stellen.

» Der eigentliche Zweck der wissenschaftlichen Methode ist es, sich zu vergewissern, ob die Natur einen nicht zu der falschen Annahme verleitet hat, man wüsste etwas, was man in Wirklichkeit nicht weiß.
Robert M. Pirsig, *Zen und die Kunst ein Motorrad zu warten,* 1974

Kann man sagen, dass die kritischen Untersuchungen der Psychologie unerwartete Ergebnisse erbracht haben? Die Antwort ist ein klares *Ja*, wie auch die folgenden Kapitel zeigen werden. Ob Sie es glauben oder nicht: Ein größerer Verlust von Hirngewebe zu einem frühen Zeitpunkt des Lebens hat evtl. nur minimale Langzeiteffekte (▶ Kap. 3). Neugeborene können den Geruch und die Stimme ihrer Mutter innerhalb weniger Tage nach der Geburt erkennen (▶ Kap. 6). Nach einer Hirnverletzung kann ein Mensch neue Fähigkeiten erlernen, sich jedoch nicht bewusst sein, dass er sie erlernt hat (▶ Kap. 9). Unterschiedliche Gruppen – Männer und Frauen, Alte und Junge, Wohlhabende und Arbeiter, Behinderte und Nichtbehinderte – berichten über ein ungefähr vergleichbares Niveau persönlichen Glücks (▶ Kap. 13).

Und hat die kritische Forschung verbreitete Annahmen in überzeugender Weise erschüttert? Auch hier lautet die Antwort *Ja*, wie wir in den folgenden Kapiteln sehen werden. Es lässt sich nachweisen, dass Schlafwandler *nicht* ihre Träume in Handlungen umsetzen (▶ Kap. 4). Die Erfahrungen, die wir im Lauf unseres Lebens gemacht haben, sind *nicht* alle als Worte im Gehirn gespeichert. Man kann durch Hirnstimulation oder Hypnose *nicht* einfach »das Tonband zurückspulen« und tief vergrabene oder verdrängte Erinnerungen wieder zum Leben erwecken (▶ Kap. 9). Die meisten Menschen leiden *nicht* unter einem unrealistisch geringen Selbstwertgefühl und ein hohes Selbstwertgefühl ist nicht immer positiv (▶ Kap. 14). Gegensätze ziehen sich i. Allg. *nicht* an (▶ Kap. 15). Bei jedem dieser Beispiele und bei weiteren ist das, was man weithin glaubt, durchaus nicht das, was sich als wahr herausgestellt hat.

Prüfen Sie Ihr Wissen

- Erklären Sie, warum es das kritische Denken fördert, wenn man die Haltung eines Wissenschaftlers einnimmt.

2.2 Wie stellen und beantworten Psychologen Fragen?

Das Rüstzeug der wissenschaftlich arbeitenden Psychologen ist die *wissenschaftliche Methode* – eine sich selbst korrigierende Vorgehensweise, Ideen zu bewerten, indem man beobachtet und untersucht. Bei ihrem Versuch, die menschliche Natur zu beschreiben und zu erklären, greift die psychologische Wissenschaft gerne vage Vorstellungen und plausibel klingende Theorien auf – und testet sie. Wenn eine Theorie funktioniert – wenn die Daten mit den Vorhersagen übereinstimmen – dann umso besser für die Theorie. Wenn die Vorhersagen versagen, wird die Theorie revidiert oder verworfen.

2.2.1 Die wissenschaftliche Methode

2.3 Wie bringen Theorien die Psychologie als Wissenschaft voran?

In unserer Alltagssprache verwenden wir den Begriff *Theorie* meist im Sinn von Vorstellung oder Gedanke. Eine wissenschaftliche **Theorie** ist ein *Erklärungsmodell*, das auf bestimmten Prinzipien basiert. Mit Hilfe einer Theorie können Beobachtungen *in ein System gebracht* sowie Verhaltensweisen und Ereignisse *vorhergesagt* werden. Eine Theorie stellt Einzelbeobachtungen in einen Zusammenhang und vereinfacht damit die Arbeit. Eine Theorie ist eine hilfreiche Zusammenfassung, indem sie Einzelfaktoren zu tieferliegenden Prinzipien verbindet. Sobald wir die einzelnen Punkte unserer Beobachtung miteinander verbinden, können wir ein kohärentes Bild erkennen.

Theorie (»theory«) – auf Prinzipien gestütztes Erklärungsmodell, das Beobachtungen in einen Zusammenhang stellt und Verhalten oder Ereignisse vorhersagt.

Eine gute Theorie über die Auswirkungen von Schlafentzug auf die Gedächtnisleistung fasst beispielsweise die unzähligen Beobachtungen hinsichtlich des Schlafs in einer kurzen Liste von Prinzipien zusammen. Stellen Sie sich vor, wir beobachten immer wieder, dass Menschen mit falschen Schlafgewohnheiten Fragen im Unterricht nicht beantworten können und bei Prüfungen schlecht abschneiden. Wir könnten also daraus die Theorie ableiten, dass Schlaf die Gedächtnisleistung verbessert. So weit, so gut: Unsere Schlaf-Merkfähigkeits-Annahme fasst auf elegante Weise eine Liste von Merkmalen über die Auswirkungen von Schlafmangel zusammen.

Doch eine Theorie kann noch so vernünftig klingen – und Schlafmangel scheint eine akzeptable Erklärung für geringere Merkfähigkeit zu sein –, sie muss getestet werden. Eine gute Theorie führt zu überprüfbaren Vorhersagen, die wir **Hypothesen** nennen. Hypothesen ermöglichen es, die Theorie zu testen und dann entweder zu revidieren oder zu verwerfen und geben dadurch der Forschung die Richtung vor. Die Hypothesen geben an, welche Resultate die Theorie stützen und welche damit nicht vereinbar sind. Wenn wir unsere Theorie zum Einfluss von Schlaf auf die Merkfähigkeit testen wollen, könnten wir prüfen, wie gut sich Personen nach einer ausgiebigen oder verkürzten Nachtruhe an Kursmaterialien erinnern (Abb. 2.8).

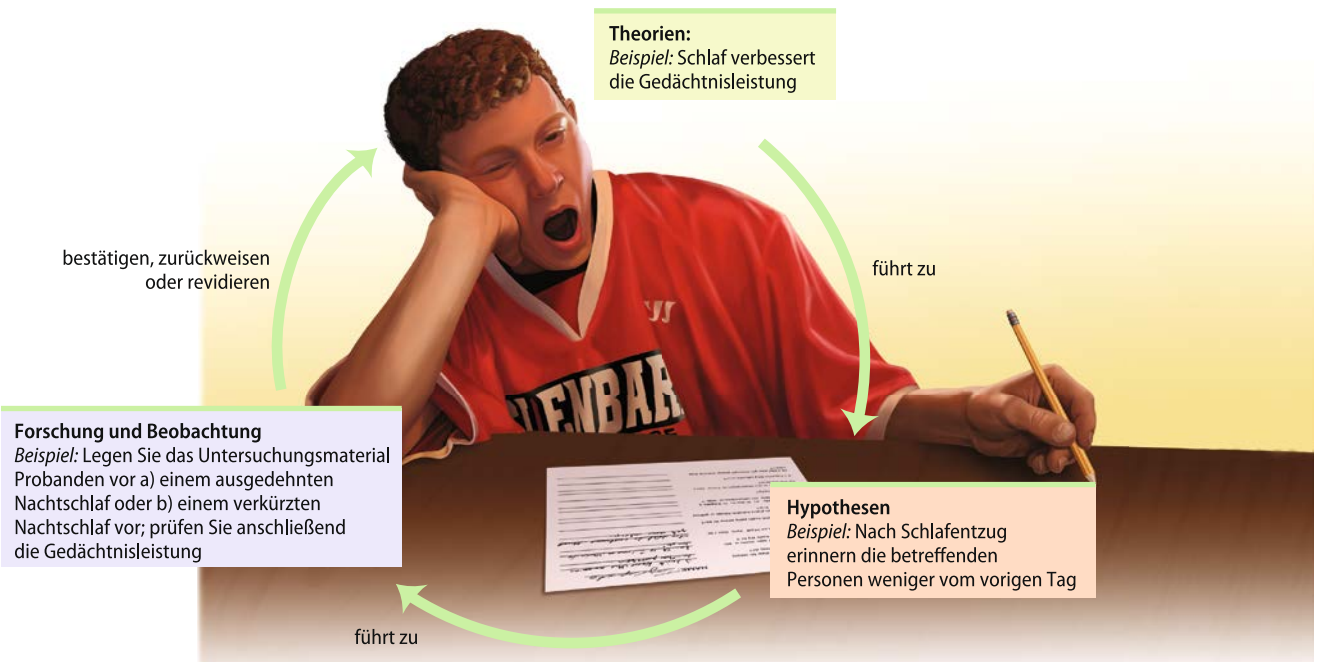

Abb. 2.8 Die wissenschaftliche Methode. Eine sich selbst korrigierende Vorgehensweise, bei der Fragen gestellt und die Antworten beobachtet werden, die die Natur gibt

Hypothese (»hypothesis«) – meist aus einer Theorie abgeleitete überprüfbare Vorhersage.

Unsere Theorien können unsere Beobachtungen verzerren (Bias, Urteilsfehler). Da unsere Theorie lautet, dass mehr Schlaf zu einem besseren Erinnerungsvermögen führt, lässt es sich nicht ausschließen, dass wir sehen, was wir erwarten zu sehen: Möglicherweise nehmen wir die Aussagen von schläfrigen Personen als weniger einsichtsvoll wahr. Der Drang zu sehen, was wir erwarten zu sehen, ist eine überall lauernde Versuchung, sowohl außerhalb als auch innerhalb des Versuchsraums. So verleiteten gemäß dem Bericht des parteienübergreifenden U.S. Senate Select Committee on Intelligence (2004) vorgefasste Erwartungen, dass der Irak Massenvernichtungswaffen besäße, Auswerter der Geheimdienste dazu, mehrdeutige Beobachtungen fälschlich so zu deuten, dass sie diese Theorie bestätigten (ähnlich wie die Meinungen, die Leute zum Klimawandel haben, ihre Interpretation von lokalen Wetterereignissen beeinflussen könnten). Diese theoriegeleitete Schlussfolgerung führte dann zur präventiven Invasion der USA im Irak.

Um diesen Bias zu kontrollieren, veröffentlichen Psychologen ihre Forschungsergebnisse mit eindeutigen **operationalen Definitionen** der Vorgehensweisen und Konzepte. *Hunger* kann beispielsweise als »Stunden ohne Essen« definiert werden, *Großzügigkeit* als »gespendetes Geld«. Werden solch sorgsam formulierte Aussagen benutzt, können andere Forscher die ursprünglichen Beobachtungen mit anderen Teilnehmern, anderem Testmaterial und in anderen Situationen **replizieren** (wiederholen). Kommen verschiedene Forscher zu ähnlichen Ergebnissen, dann wächst das Vertrauen in die Reliabilität (Zuverlässigkeit) der Ergebnisse. Die erste Untersuchung zum Hindsightbias weckte die Neugier der Psychologen. Heute, nach vielen erfolgreichen Replikationen mit jeweils anderen Menschen und anderen Fragen, sind wir uns der Bedeutung dieses Phänomens bewusst.

Operationale Definition (»operational definition«) – Festlegung der Vorgehensweise (Operation) bei der Definition der Untersuchungsvariablen. So kann Intelligenz beispielsweise operational definiert werden als »das, was ein Intelligenztest misst«.

Replikation (»replication«) – Wiederholung der wesentlichen Parameter eines Experiments, in der Regel mit anderen Versuchsteilnehmern in anderen Situationen. Mit Hilfe der Replikation kann festgestellt werden, ob sich die Grundannahmen eines Experiments auf andere Versuchsteilnehmer und andere Situationen übertragen lassen.

Unsere Theorie wird letztlich nur dann nützlich sein, wenn man mit ihrer Hilfe

1. eine Reihe von persönlichen Berichten und Beobachtungen in eine effiziente *Ordnung* bringen kann und wenn man
2. anhand der Theorie klare *Vorhersagen* machen kann, mithilfe derer jeder die Theorie überprüfen oder in der Praxis einsetzen kann. (Wenn die Leute mehr schlafen, wird sich dann die Merkfähigkeit verbessern?)

Vielleicht führt uns die Forschung zu einer revidierten Theorie, die das, was wir wissen, in einen besseren Zusammenhang bringt und genauere Vorhersagen gestattet. Unsere Forschungsergebnisse könnten aber auch repliziert und durch ähnliche Befunde unterstützt werden. (Dies war der Fall für Schlaf- und Gedächtnisstudien, wie Sie in ▶ Kap. 4 sehen werden.)

Wie wir als Nächstes erfahren werden, können wir mit Hilfe *deskriptiver* Methoden (die Verhaltensweisen beschreiben, oftmals durch Einzelfallstudien, Umfragen oder Feldbeobachtungen), *Korrelationsberechnungen* (die verschiedene Faktoren zueinander in Beziehung setzen) und *Experimenten* (die Faktoren manipulieren, um ihre Effekte zu erkennen) unsere Hypothesen überprüfen und unsere Theorien revidieren. Wenn wir verbreitete psychologische Behauptungen mit kritischem Verstand überprüfen wollen, müssen wir mit diesen Methoden vertraut sein und wissen, welche Schlussfolgerungen wir mit ihrer Hilfe ziehen können.

Prüfen Sie Ihr Wissen

- Was zeichnet eine gute Theorie aus?
- Warum ist Replikation so wichtig?

2.2.2 Beschreibung

2.4 Wie nutzen Psychologen Einzelfallstudien, Feldbeobachtungen und Umfragen, um Verhalten zu beobachten und zu beschreiben, und weshalb ist es wichtig, Zufallsstichproben zu ziehen?

Der Ausgangspunkt jeder Wissenschaft ist die Beschreibung. In unserem Alltag beobachten wir unsere Mitmenschen und beschreiben sie; daraus leiten wir ab, warum sie sich so verhalten und nicht anders. Im Wesentlichen tun professionelle Psychologen auch nichts anderes, nur gehen sie dabei systematischer vor und bemühen sich stärker um Objektivität. Dazu verwenden sie

- *Einzelfallstudien* (Untersuchungen bestimmter Individuen),
- *Feldbeobachtung* (das natürliche Verhalten vieler Individuen wird beobachtet und aufgezeichnet) sowie
- *Umfragen* und Interviews (den untersuchten Personen werden Fragen gestellt).

Einzelfallstudie

Die **Einzelfallstudie** gehört zu den ältesten Forschungsmethoden überhaupt. In der Fallstudie wird ein Individuum gründlich studiert, in der Hoffnung, dabei Dinge zu entdecken, die für alle Individuen gelten. Einige Beispiele: Viel von unserem frühen Wissen über das Gehirn stammt aus Fallstudien mit Menschen, die nach der Schädigung einer bestimmten Hirnregion in bestimmten Bereichen beeinträchtigt waren. Der Entwicklungspsychologe Jean Piaget beobachtete aufmerksam ein paar Kinder und befragte sie ausführlich; das Ergebnis waren bahnbrechende

Informationen über die Art, wie Kinder denken. Untersuchungen, die mit nur wenigen Schimpansen durchgeführt wurden, erbrachten den Beweis, dass die Tiere über die Fähigkeit zum Verstehen und über sprachliche Fähigkeiten verfügen. Intensive Fallstudien bringen manchmal sehr viel ans Licht. Sie zeigen uns, was passieren *kann*, und oftmals geben sie Richtungen für die weitere Forschung vor.

Einzelfallstudie (»case study«) – Beobachtungstechnik, bei der ein Individuum gründlich und intensiv beobachtet wird in der Hoffnung, auf diese Weise universelle Prinzipien entdecken zu können.

Individuelle Fälle können uns aber auch manchmal in die Irre führen. Vielleicht ist das untersuchte Individuum untypisch. Eine nicht repräsentative Information kann zu irrtümlichen Beurteilungen und falschen Schlussfolgerungen führen. Und auch Folgendes ist zu bedenken: Sobald ein Wissenschaftler einen Befund veröffentlicht *(»Raucher sterben jünger: 95% der Männer über 85 sind Nichtraucher«)*, findet sich sicher jemand, der widerspricht *(»Nun, ich habe einen Onkel, der täglich zwei Päckchen rauchte und 89 wurde«)*. Dramatische Geschichten und persönliche Erfahrungen (sogar psychologische Fallbeispiele) erregen unsere Aufmerksamkeit und setzen sich im Gedächtnis fest. Journalisten wissen das. Einen Artikel über Zwangsversteigerungen beginnen sie daher mit der tragischen Geschichte einer Familie, die aus ihrem Haus geworfen wurde, und nicht mit Statistiken über Zwangsversteigerungen. Geschichten bewegen uns. Aber sie können uns auch täuschen.

Welche der folgenden Aussagen finden Sie denkwürdiger?

1. »In einer Untersuchung mit 1300 Berichten über Träume, die das Schicksal eines entführten Kindes betrafen, sahen tatsächlich nur 5% den Tod des Kindes vorher und hatten damit Recht (Murray u. Wheeler 1937).«
2. »Ich kenne einen Mann, der träumte, dass seine Schwester einen Autounfall hätte, und zwei Tage später starb sie bei einer Frontalkollision!«

Zahlen können langweilig sein, aber viele Anekdoten geben noch lange keinen Beweis ab. Der Psychologe Gordon Allport (1954, S. 9) drückte es so aus: »Einen Fingerhut voll [dramatischer] Ereignisse verallgemeinern wir schleunigst auf Badewannengröße.«

» ›Nun ja, mein Lieber‹, sagte Miss Marple, ›das Wesen des Menschen ist eigentlich an allen Orten gleich, aber natürlich hat man in einem Dorf bessere Möglichkeiten, es sich genauer anzusehen.‹
Agatha Christie, *Der Dienstagabend-Klub,* 1933

Merken Sie sich: Aus Einzelfällen können fruchtbare Hypothesen abgeleitet werden. Was auf uns alle zutrifft, kann man schon am Beispiel eines Einzelnen erkennen. Um aber die allgemeinen Wahrheiten in den Einzelfällen zu entdecken, müssen wir andere Methoden anwenden.

Prüfen Sie Ihr Wissen

- Durch Einzelfallstudien erfahren wir nichts über generelle Prinzipien, die für uns alle gelten. Wieso nicht?

Beobachtung in natürlicher Umgebung (Feldbeobachtung)

Eine zweite beschreibende Forschungsmethode erfasst Verhaltensweisen in natürlicher Umgebung. **Beobachtungen in natürlicher Umgebung** oder Feldbeobachtungen reichen von der Beobachtung von Schimpansengesellschaften im Dschungel bis zu nichtreaktiven Videoaufnahmen (und der späteren systematischen Auswertung) von Eltern-Kind-Interaktionen in verschiedenen Kulturen oder der Beschreibung der Platzwahl in der Cafeteria einer Schule, die von Schülern verschiedener Kulturen besucht wird.

Die Beobachtung in einer natürlichen Umgebung *beschreibt* Verhalten, *erklärt* es aber nicht, ebenso wenig wie die Einzelfallstudie. Trotzdem können die Beschreibungen ein Licht auf bestimmte Dinge werfen. Beispielsweise dachte man einst, dass nur Menschen Werkzeuge benutzen. Die Beobachtung von Schimpansen in ihrem natürlichen Umfeld hat indessen gezeigt, dass die Tiere manchmal ein Stöckchen in einen Termitenhaufen stecken und dann ablecken. Solche nichtreaktiven Feldbeobachtungen bereiteten den Weg für spätere systematische Untersuchungen zum Denken und Fühlen von Tieren sowie ihrer Möglichkeiten zum Verständnis von Sprache. Diese Untersuchungen haben unser Verständnis für unsere Mitgeschöpfe, die Tiere, ausgedehnt. »Beobachtungen im natürlichen Habitat zeigten, dass Zusammenleben und Verhalten von Tieren weitaus komplexer sind, als wir bislang angenommen haben«, stellte die Schimpansenforscherin Jane Goodall fest (1998). Dank der Beobachtungen von Forschern wissen wir, dass Schimpansen und Paviane Täuschungsmanöver einsetzen. Die Psychologen Andrew Whiten und Richard Byrne (1988) beobachteten wiederholt, wie ein junger Pavian so tat, als sei er von einem anderen Pavian angegriffen worden, eine Taktik, mit der er seine Mutter dazu brachte, den anderen vom Futter wegzujagen. Je weiter das Gehirn einer Primatenart entwickelt ist, desto eher werden diese Tiere Täuschungsverhalten zeigen (Byrne u. Corp 2004; ◘ Abb. 2.9).

Beobachtung in natürlicher Umgebung (auch Feldbeobachtung; »naturalistic observation«) – Beobachten und Erfassen von Verhalten in natürlichen Situationen unter Verzicht auf Manipulation oder Kontrolle der Situation.

Auch bei Menschen werden Feldbeobachtungen durchgeführt. Im Folgenden finden Sie vier Beispiele dafür, an denen Sie, glaube ich, Spaß haben werden.

- *Ein witziges Ergebnis.* In sozialen Situationen lachen wir Menschen 30-mal öfter, als wenn wir allein sind. (Ist Ihnen schon einmal aufgefallen, wie selten Sie lachen, wenn Sie allein sind?) Wenn wir lachen, verziehen 17 Muskeln

Abb. 2.9 Ein geborener Beobachter. Der Schimpansenforscher Frans de Waal (2005) sagte: »Ich bin ein geborener Beobachter … Wenn ich im Restaurant einen Platz aussuche, möchte ich so viele Tische wie möglich im Blickfeld haben. Ich liebe es, die sozialen Dynamiken um mich herum – Liebe, Anspannung, Langeweile, Antipathie – zu verfolgen. Dabei konzentriere ich mich auf die Körpersprache, welche ich für informativer halte als das gesprochene Wort. Da andere zu beobachten etwas ist, das ich automatisch tue, fiel es mir sehr leicht, eine »Fliege an der Wand« einer Affenkolonie zu werden.« (Bildrechte: Jack Kearse, Emory University for Yerkes National Primate Research Center)

Tab. 2.1 Einen Penny für deine Gedanken: Das innere Erleben von Universitätsstudenten*

Innere Erfahrung	Beispiel	Häufigkeit
Innere Rede	Susan sagte zu sich selbst: »Ich muss in den Kurs gehen.«	26%
Inneres Sehen	Paul stellte sich das Gesicht seiner besten Freundin vor, inklusive ihres Halses und Kopfs.	34%
Symbolfreies Denken	Alphonse fragte sich, ob die Arbeiter die Ziegelsteine wohl fallen lassen würden.	26%
Gefühl	Courtney erlebte Wut und die dazugehörigen körperlichen Symptome.	26%
Sinneswahrnehmung	Fiona spürte die kalte Brise auf ihren Wangen und wie sich ihre Haare bewegten.	22%

* Es konnte mehr als eine Erfahrung zu einem Zeitpunkt angegeben werden.

unseren Mund und drücken unsere Augen zusammen, und wir stoßen im Abstand von jeweils einer fünftel Sekunde eine Serie von vokalähnlichen Lauten von 75 Millisekunden Dauer aus (Provine 2001).

- *Aushorchen von Studenten.* Was sagen und machen Erstsemester der Psychologie eigentlich den lieben langen Tag? Um das herauszufinden, statteten Matthias Mehl und James Pennebaker (2003) 52 Studierende der University of Texas bis zu vier Tage lang mit einem Kassettenrekorder aus, den sie an einem Gürtel trugen und der sich während der Wachstunden alle 12,5 Minuten 30 Sekunden lang für eine Aufnahme einschaltete. So waren die Forscher imstande, mehr als 10.000 eine halbe Minute lange Ausschnitte aus dem Leben der Studierenden abzuhören. Was meinen Sie, wie viel Prozent der Zeit Studierende nach diesen Befunden mit jemandem sprachen? Und wie viel Prozent der Zeit befanden sich danach die Studierenden an der Tastatur eines Computers? Die Antwort lautet: 28 bzw. 9 Prozent der Zeit. (Wie viel Prozent der Zeit, in der Sie nicht schlafen, verbringen *Sie* mit diesen Aktivitäten?)
- *An was denkst du gerade?* Um herauszufinden, was in den Köpfen von Studenten an der University of Nevada, Las Vegas, vorgeht, gaben Christopher Heavey und Russell Hurlburt (2008) ihnen Piepser. Ein halbes Dutzend Mal unterbrach ein Piepsen die täglichen Aktivitäten der Studenten, indem es ihnen signalisierte, ein Notizbuch herauszuziehen und ihr inneres Erleben zu diesem Zeitpunkt festzuhalten. Als die Forscher die Berichte später kodierten und Kategorien zuordneten, fanden sie fünf verbreitete Formen inneren Erlebens (Tab. 2.1).
- *Kultur, Klima, Lebensrhythmus.* Mit Hilfe der Feldbeobachtung konnten Robert Levine und Ara Norenzayan (1999) den Lebensrhythmus in 31 Ländern miteinander vergleichen. (Die operationale Definition von *Lebensrhythmus* beinhaltete die Gehgeschwindigkeit, das Tempo, in dem Postangestellte eine simple Bitte erfüllten und die Genauigkeit der Normaluhren.) Die Forscher kamen zu dem Ergebnis, dass der Lebensrhythmus in Japan und Westeuropa am schnellsten ist und langsamer in wirtschaftlich weniger hoch entwickelten Ländern. In einem kühlen Klima leben die Menschen tendenziell in schnellerem Rhythmus (und mehr Menschen sterben an Herzkrankheiten).

Die Beobachtung in natürlicher Umgebung (Feldbeobachtung) liefert uns interessante Schnappschüsse aus dem alltäglichen Leben. Dies geschieht jedoch ohne Kontrolle all der Faktoren, die das Verhalten beeinflussen können. Es ist eine Sache, das Lebenstempo an verschiedenen Orten zu beobachten, eine andere jedoch, zu verstehen, weshalb manche Menschen schneller gehen als andere.

Prüfen Sie Ihr Wissen

- Was sind die Vor- und Nachteile der Feldbeobachtung, wie sie Mehl und Pennebaker in dieser Studie angewandt haben (Abb. 2.10)?

Abb. 2.10 Feldbeobachtung mit Electronically Activated Recorder. Die Psychologen Matthias Mehl und James Pennebaker verwendeten Kassettenrekorder, um natürliche Ausschnitte aus dem Alltagsleben einzufangen. (Courtesy of Matthias Mehl)

Befragung

Bei der Methode der **Befragung** werden viele Fälle einbezogen, die Fragen bleiben aber eher an der Oberfläche. Bei einer Umfrage werden die Menschen gebeten, Auskunft über ihr Verhalten oder ihre Ansichten zu geben. Thema der Befragung kann alles sein, von sexuellen Praktiken bis hin zu politischen Meinungen. In neueren Studien

- gab die Hälfte aller Amerikaner an, am vorherigen Tag mehr Glück und Freude als Sorge und Stress erlebt zu haben (Gallup 2010).
- berichten Kanadier mit Internetzugang, neue Formen der elektronischen Kommunikation zu nutzen und dadurch im Jahr 2010 35% weniger E-Mails zu erhalten als im Jahr 2008 (Ipsos 2010a).
- gab quer verteilt über 22 Länder eine von fünf Personen an, zu glauben, dass Aliens zur Erde gekommen sind und nun – getarnt als Menschen – unter uns weilen (Ipsos 2010b).
- sagen 68% aller Menschen – das sind etwa 4,6 Mrd. Menschen –, dass die Religion in ihrem Alltag eine wichtige Rolle spielt (Diener et al. 2011).

Aber die Fragen richtig zu stellen, ist eine heikle Sache; und die Antworten können auch von Formulierungen und von der Auswahl der Befragten abhängen.

Befragung (»survey«) – Technik, bei der die von ihnen selbst berichteten Einstellungen oder Verhaltensweisen der Menschen einer bestimmten Gruppe ermittelt werden; i. Allg. wird eine repräsentative Zufallsstichprobe befragt.

Formulierungseffekte

Schon ganz leichte Abänderungen in der Reihenfolge oder Wortstellung der Fragen können eine große Wirkung haben. Leute sprechen sich viel eher für »Hilfe für Bedürftige« aus als für »Sozialhilfe«, für »Förderung« als für »bevorzugte Behandlung«, sie stimmen weit eher dafür, Zigarettenwerbung oder Pornografie im Fernsehen »nicht zu erlauben« als zu »zensieren« und befürworten eher eine »Ausweitung der Staatseinnahmen« als »Steuererhöhungen«. 2009 sprachen sich drei von vier Amerikanern in einer nationalen Befragung dafür aus, der Bevölkerung »eine Wahl« zu geben zwischen einer öffentlichen, staatlichen und einer privaten Gesundheitsversicherung. In einer anderen Befragung jedoch waren die meisten Amerikaner nicht dafür, »einen Plan für die öffentliche Gesundheitsvorsorge auszuarbeiten, der – verwaltet durch die Bundesregierung – direkt mit privaten Einrichtungen der Gesundheitsvorsorge in Konkurrenz stehen würde« (Stein 2009). Vom Wortlaut der Frage hängt sehr viel ab, deshalb hinterfragen kritische Denker immer, auf welche Weise die Fragestellung die Meinung der Befragten beeinflusst haben könnte.

Zufallsstichprobe

Im Alltagsdenken generalisieren wir ständig aufgrund der Stichproben, die wir beobachten, vor allem, wenn es um Fälle geht, die lebhaft vorgetragen werden. Nehmen wir folgendes Szenario: Dem Verwaltungsdirektor liegt (a) die statistische Zusammenfassung der Bewertung eines Professors durch seine Studenten vor, und gleichzeitig hört er (b) die heftigen Proteste zweier aufgebrachter Studenten, die eine nicht repräsentative Stichprobe bilden. Da kann der Eindruck des Verwaltungsdirektors ebenso von den beiden Pechvögeln beeinflusst werden wie von den vielen positiven Bewertungen dieses Professors, die in der Statistik aufgeführt sind. Der Versuchung, den Stichprobenfehler zu ignorieren und von ein paar intensiven, aber nicht repräsentativen Eindrücken zu verallgemeinern, kann man jedoch kaum widerstehen (Abb. 2.11).

Merken Sie sich: Die beste Basis für eine Generalisierung ist eine *repräsentative Stichprobe*. Aber es ist nicht immer möglich, jede Person einer Gruppe zu befragen. Wie erhalten Sie also eine repräsentative Stichprobe – sagen wir der Studenten Ihrer Hochschule? Wie könnten Sie eine Gruppe aussuchen, die repräsentativ ist für die gesamte studentische **Population** – die gesamte Gruppe, die Sie untersuchen und beschreiben wollen? Typischerweise würden Sie eine **Zufallsstichprobe** ziehen, eine Stichprobe, bei der jede einzelne Person in der Gesamtgruppe die gleichen Chancen hat teilzunehmen. Sie würden etwa die Namen in einer Auflistung der Studierenden nummerieren und dann einige Zufallszahlen generieren, um die Umfrageteilnehmerinnen und -teilnehmer auszuwählen. (Jedem Student einen Fragebogen zuzusenden, würde nicht funktionieren, denn die gewissenhaften Menschen, die ihn zurückschicken, wären keine Zufallsstichprobe.) Große repräsentative Stichproben sind besser als kleine; aber eine kleine repräsentative Stichprobe von

Abb. 2.11 (This modern world by Tom Tomorrow, © 1991)

100 Personen ist besser als eine nicht repräsentative Stichprobe von 500 Personen.

Politische Meinungsforscher erhalten ihre Stichprobe für landesweite Wählerbefragungen genau auf diesem Weg. Nur 1500 zufällig ausgewählte Leute aus allen Teilen des Landes liefern ein bemerkenswert genaues Bild von den im Land herrschenden Meinungen. Ohne Zufallsstichprobe führen große Stichproben – und dazu gehören auch Telefonbefragungen und Fernseh- oder Website-Abstimmungen – oft bloß zu irreführenden Ergebnissen.

Population (»population«) – sämtliche Fälle in einer Gruppe, aus der eine Stichprobe für eine Studie gezogen wird. (*Achtung:* Mit Ausnahme von nationalen Studien ist damit *nicht* die gesamte Population eines Landes gemeint.)

Zufallsstichprobe (»random sample«) – Stichprobe, die eine Population weitgehend repräsentiert, weil jedes Mitglied der Population mit der gleichen Wahrscheinlichkeit in die Stichprobe aufgenommen werden kann.

Merken Sie sich: Bevor man den Ergebnissen einer Umfrage Glauben schenkt, sollte man sie kritisch hinterfragen: Betrachten Sie die Stichprobe. Man kann die Nachteile einer nicht repräsentativen Stichprobe nicht dadurch wettmachen, dass man einfach weitere Personen hinzunimmt.

Schätzungen mit sehr großen Stichproben sind recht zuverlässig (reliabel). Der Buchstabe E hat schätzungsweise einen Anteil von 12,7% an allen Buchstaben in englischsprachigen schriftlichen Texten. In Melvilles *Moby Dick* stellt E tatsächlich 12,3% der 925.141 Buchstaben des Textes, 12,4% der 586.747 Buchstaben in Dickens' *A Tale of Two Cities* und 12,1% der 3.901.021 Buchstaben, aus denen 12 Werke von Mark Twain bestehen (*Chance News* 1997).

Prüfen Sie Ihr Wissen

- Was ist der *Stichprobenfehler* und wie vermeiden ihn Forscher?

2.2.3 Korrelation

2.5 Was sind positive und negative Korrelationen und wieso ermöglichen sie Vorhersagen, nicht aber Ursache-Wirkungs-Erklärungen?

Die Verhaltensbeschreibung ist der erste Schritt zur Verhaltensvorhersage. Beobachtungen in natürlicher Umgebung und Befragungen zeigen uns häufig, dass ein bestimmtes Merkmal oder ein Verhalten mit einem anderen zusammen auftritt. In solchen Fällen sprechen wir von einer **Korrelation**. Eine statistische Maßzahl (der **Korrelationskoeffizient**) zeigt uns, wie eng zwei Faktoren miteinander verknüpft sind und wie gut der eine Faktor den anderen *vorhersagt*. Wenn wir wissen, wie stark ein hoher Punktwert (Score) in einem Eignungstest mit Schulerfolg *korreliert*, dann wissen wir auch, wie gut diese Punktzahl Schulerfolg *vorhersagt*.

Korrelation (»correlation«) – Maßeinheit, welche das Ausmaß des Zusammenhangs zwischen zwei Merkmalsvariablen angibt und damit ausdrückt, wie gut eine Variable die andere Variable vorhersagt.

Korrelationskoeffizient (»correlation coefficient«) – statistische Maßzahl des Zusammenhangs zwischen zwei Variablen (von –1 bis +1).

In diesem Buch werden wir immer wieder die Frage stellen, wie stark der Zusammenhang zwischen zwei Variablen ist: Wie eng hängen beispielsweise die Persönlichkeitsscores eineiiger Zwillinge zusammen? Wie genau kann man anhand der Punktzahl in einem Intelligenztest die berufliche Leistung vorhersagen? Wie eng hängt Stress mit Krankheit zusammen? In solchen Fällen können Scatterplots bzw. Punkt- oder **Streudiagramme** Klarheit schaffen.

Jeder Punkt in einem Streudiagramm gibt den Wert von zwei Variablen an. Die drei Streudiagramme in Abb. 2.12 illustrieren die Bandbreite möglicher Korrelationen – von einer perfekten positiven bis hin zu einer perfekten negativen Korrelation. (Perfekte Korrelationen kommen in der »realen Welt« allerdings kaum vor.) Eine positive Korrelation bedeutet, dass zwei Wertereihen wie etwa Körpergröße und Gewicht jeweils gemeinsam größer bzw. kleiner werden.

Eine negative Korrelation sagt nichts darüber aus, ob der Zusammenhang zwischen zwei Variablen stark oder schwach ist; zwei negativ korrelierende Variablen bedeuten, dass hier ein umgekehrter Zusammenhang vorliegt (die Punktzahl der einen Variable steigt, während die Punktzahl der anderen sinkt). Die Studie zur inneren Rede, die an Studenten der University of

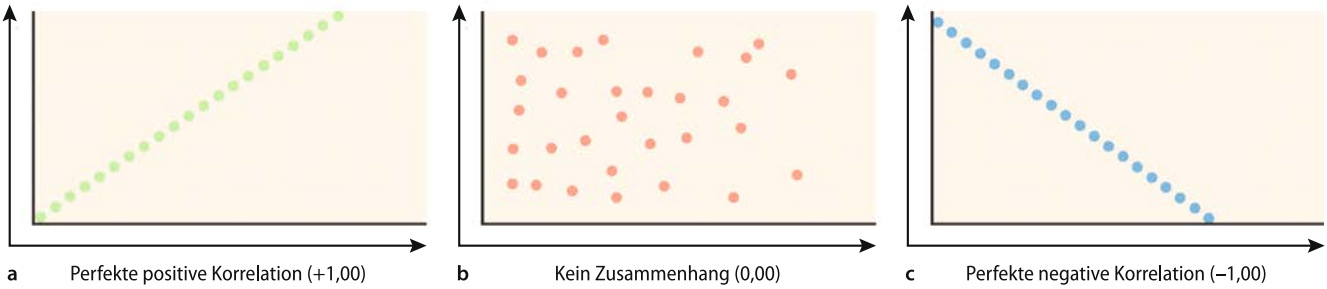

Abb. 2.12a–c Streudiagramme (Scatterplots), die verschiedene Korrelationsmuster zeigen. Korrelationskoeffizienten variieren zwischen +1,00 (die Werte einer Variablen wachsen direkt proportional mit den Werten der anderen Variablen) und –1,00 (die Werte einer Variablen sinken direkt proportional mit dem Ansteigen der anderen Variablen)

Nevada durchgeführt und früher in diesem Kapitel bereits diskutiert wurde, beinhaltete ebenfalls eine Korrelationskomponente. Die Berichte der Studenten über innere Rede korrelierten negativ (–.36) mit den Punkten auf einer anderen Variable: psychologischer Stress. Studenten, die vermehrt von innerer Rede berichteten, gaben tendenziell etwas *weniger* psychologischen Stress an.

Streudiagramm (auch Punktdiagramm; »scatterplot«) – grafisch dargestellte Punktewolke. Jeder Punkt in einem Streudiagramm gibt die Werte von zwei Merkmalsvariablen an. Der Verlauf der Verbindungslinie zwischen den Punkten zeigt die Richtung des Zusammenhangs zwischen den beiden Variablen an. Die Konzentration der Punkte verweist auf die Stärke des Zusammenhangs (eng beieinanderliegende Punkte bedeuten hohe Korrelation).

Mit Hilfe der Statistik können wir Dinge erkennen, die wir sonst nicht sehen würden. Probieren Sie es doch einmal mit einem eigenen imaginären Projekt aus. Sie könnten sich fragen, ob hochgewachsene Männer gelassener oder nervöser sind als kleine. Zu diesem Zweck sammeln Sie zwei Datensätze: die Körpergröße der Männer und ihr Temperament. Sie messen die Körpergröße von 20 Männern und bitten einen unbeteiligten Kollegen, das Temperament dieser Männer zu bewerten (dabei bedeutet 0 »sehr ruhig« und 100 »extrem nervös«).

Können Sie mit diesen Daten (Tab. 2.2) vor Augen schon sagen, ob es eine positive oder eine negative Korrelation zwischen Körpergröße und Nervosität gibt? Oder ist die Korrelation schwach oder gar nicht vorhanden?

Beim Vergleich der beiden Zahlenreihen in Tab. 2.2 können die meisten Menschen kaum einen Zusammenhang zwischen Körpergröße und Temperament erkennen. Tatsächlich ist die Korrelation in diesem fiktiven Beispiel positiv, nämlich +0,63. Das erkennen wir, wenn wir die Daten in einem Punktdiagramm anordnen. Verfolgt man die Punkte in Abb. 2.13 von links nach rechts, so zeigt der nach oben führende, ovalförmige Verlauf der Punktewolke, dass die beiden Variablen (Körpergröße und Temperament) unseres fiktiven Untersuchungsbeispiels tendenziell gemeinsam ansteigen.

Wenn wir schon bei der systematischen Darbietung der Daten in Tab. 2.2 keinen Zusammenhang erkennen können, dann verstehen wir, dass wir mögliche Zusammenhänge, die uns im Alltag begegnen, leicht übersehen. Manchmal brauchen wir die Klarheit und Deutlichkeit der Statistik, um zu erkennen, was direkt vor unseren Augen liegt. Mit Hilfe von statistisch aufbereiteten Informationen über das Anforderungsniveau der beruflichen Tätigkeit, Führungspositionen, Leistung, Geschlecht und Einkommen können wir ohne Schwierigkeiten erkennen, wo Geschlechtsdiskriminierung vorliegt. Doch wir nehmen die Diskriminierung häufig nicht wahr, wenn die gleiche Information in Form von Einzelfällen präsentiert wird (Twiss et al. 1989).

Tab. 2.2 Körpergröße und Temperament von 20 Männern

Versuchsperson	Körpergröße in cm	Temperament
1	200	75
2	158	66
3	165	60
4	195	90
5	185	60
6	173	42
7	155	42
8	183	60
9	188	81
10	151	39
11	163	48
12	189	69
13	178	72
14	165	57
15	182	63
16	175	75
17	159	30
18	180	57
19	171	84
20	175	39

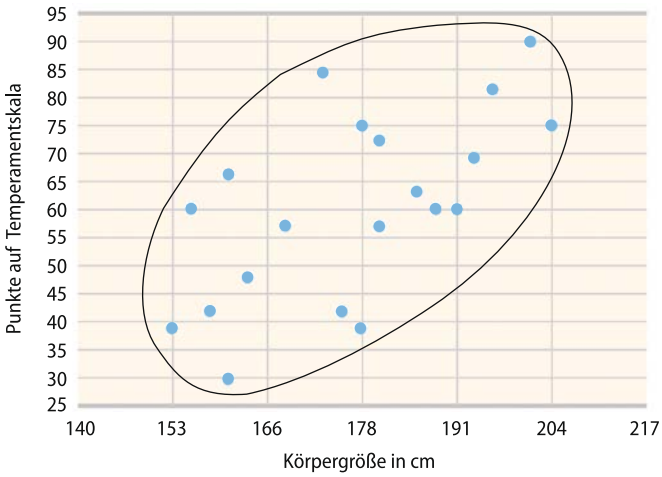

Abb. 2.13 Streudiagramm für Körpergröße und Temperament. Die dargestellten Daten von 20 fiktiven Personen (jede Person wird durch einen Punkt dargestellt) zeigen einen nach oben führenden Verlauf und weisen damit auf eine positive Korrelation hin. Die Punkte sind über das ganze Diagramm verstreut; das zeigt, dass die Korrelation deutlich unter +1,0 liegt

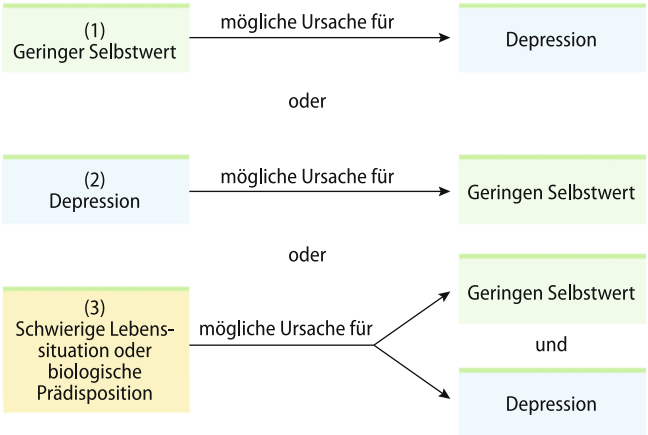

Abb. 2.14 Drei mögliche Zusammenhänge von Ursache und Wirkung. Menschen mit geringem Selbstwert werden eher von Depressionen berichten als Menschen mit hohem Selbstwert. Eine mögliche Erklärung für diese negative Korrelation könnte lauten: Ein schlechtes Selbstbild verursacht depressive Gefühle. Doch zeigt das Diagramm, dass auch andere Ursache-Wirkungs-Zusammenhänge denkbar sind

Merken Sie sich: Der Korrelationskoeffizient kann uns helfen, die Welt dadurch deutlicher wahrzunehmen, dass wir die Stärke des Zusammenhangs zwischen zwei Faktoren erkennen können.

Prüfen Sie Ihr Wissen

- Geben Sie an, welche Studien über positive Korrelationen berichten und welche über negative.
  1. Je mehr Kinder und Jugendliche verschiedene Medien nutzen, desto weniger zufrieden sind sie mit ihrem Leben (Kaiser 2010). _____
  2. Je mehr sexuell geprägte Inhalte sich Jugendliche im Fernsehen ansehen, desto wahrscheinlicher ist es, dass sie sexuell aktiv werden (Collins et al. 2004). _____
  3. Je länger Kinder gestillt werden, desto besser ist später ihre schulische Leistung (Horwood u. Fergusson 1998). _____
  4. Je stärker das Einkommen bei einer Stichprobe armer Familien anwuchs, desto weniger psychiatrische Symptome hatten ihre Kinder (Costello et al. 2003). _____

Korrelation und Kausalität

Korrelationen sind hilfreich bei der Vorhersage. Die *New York Times* berichtet, dass US-Countys mit vielen Waffenbesitzern tendenziell höhere Mordraten verzeichnen (Luo 2011). Waffenbesitz sagt Tötungsdelikte vorher. Was könnte die Korrelation zwischen Waffen und Tötungsdelikten wohl erklären?

Ich kann jemanden geradezu denken hören: »Na logisch, wenn Waffen in greifbarer Nähe sind, werden – insbesondere im Affekt – mehr Menschen getötet.« Dies wäre ein Beispiel dafür, wie A (Waffen) B (Morde) verursacht. Ich höre aber andere Leser entgegnen: »Nicht so schnell. Vielleicht kaufen Menschen, die an gefährlichen Orten leben, mehr Waffen zur Selbstverteidigung – vielleicht führt B zu A.« Oder vielleicht ist auch ein dritter Faktor C für A und B verantwortlich.

Ein anderes Beispiel: Das Selbstwertgefühl korreliert negativ mit Depression (und sagt deshalb Depression vorher). (Je niedriger der Selbstwert einer Person ist, desto größer ist ihr Risiko, an einer Depression zu erkranken.) Bedeutet das, dass ein geringes Selbstwertgefühl Depression *verursacht*? Sollten Sie – aufgrund der evidenten Korrelation – der Meinung sein, dass das der Fall ist, dann befinden Sie sich in guter Gesellschaft. Einer der fast unvermeidbaren Denkfehler ist die Annahme, dass eine Assoziation – manchmal als Korrelationskoeffizient dargestellt – der Nachweis für einen Kausalzusammenhang ist. Doch ganz gleich, wie stark der Zusammenhang auch sein mag: Korrelation ist kein Beweis für Kausalität.

Doch wie die Optionen 2 und 3 in Abb. 2.14 zeigen, bekommen wir dieselbe negative Korrelation zwischen Selbstwert und Depression, wenn wir davon ausgehen, dass die Depression die Ursache dafür ist, dass die betroffenen Menschen nicht viel von sich halten. Eine weitere denkbare Erklärung für die Korrelation wäre ein dritter Faktor: Eine ererbte Veranlagung oder eine Störung der chemischen Botenstoffe im Gehirn könnte sowohl das geringe Selbstwertgefühl als auch die Depression erklären.

Dieser Punkt ist so wichtig und zeigt so grundlegend, wie durch die Anwendung psychologischen Wissens eine genauere Beurteilung möglich ist, dass ich noch das Beispiel einer Befragung von über 12.000 Jugendlichen anführen möchte. Je mehr sich Teenager von ihren Eltern geliebt fühlen, desto geringer ist die Wahrscheinlichkeit, dass sie schädliche Gewohnheiten annehmen: frühe sexuelle Beziehungen, Rauchen, Alkohol- und

Drogenmissbrauch oder gewalttätiges Verhalten (Resnick et al. 1997). »Eltern haben großen Einfluss auf das Verhalten ihrer Kinder während der gesamten Highschool-Zeit«, schwärmte daraufhin eine Meldung der *Associated Press* (AP) über diese Studie. Doch die Korrelation beinhaltet nicht automatisch einen Ursache-Wirkungs-Zusammenhang. Die AP hätte mit ebenso viel Berechtigung titeln können: »Jugendliche, die sich vernünftig verhalten, fühlen sich von ihren Eltern geliebt und anerkannt, während Jugendliche, die zu Grenzüberschreitungen neigen, ihre Eltern für verständnislose Trottel halten.«

Prüfen Sie Ihr Wissen

- Die Dauer einer Ehe korreliert mit dem Verlust von Haaren bei Männern. Bedeutet das, dass die Ehe den Haarverlust verursacht (oder dass kahl werdende Männer die besseren Ehemänner sind)?

Merken Sie sich (passen Sie jetzt besonders gut auf): *Assoziation ist kein Beweis für Kausalität.*[3] Eine Korrelation weist auf die *Möglichkeit* eines Ursache-Wirkungs-Zusammenhangs hin, doch sie ist *kein Nachweis* für einen Kausalzusammenhang. Behalten Sie dieses Prinzip im Gedächtnis, und Sie werden, wenn Sie Berichte über wissenschaftliche Untersuchungen lesen bzw. hören, diese besser beurteilen können.

2.2.4 Experiment

2.6 Welches sind die Eigenschaften von Experimenten, die es möglich machen, Ursache und Wirkung auseinanderzuhalten?

»Glücklich ist, wer die Ursachen der Dinge erkennen kann«, sagte der römische Dichter Vergil. Wie können Psychologen Ursachen in Korrelationsstudien erkennen, z. B. die Korrelation zwischen Stillen und Intelligenz?

Forscher haben herausgefunden, dass gestillte Kinder in Intelligenztests etwas besser abschneiden als Flaschenkinder, die mit Kuhmilch ernährt wurden (Angelsen et al. 2001; Mortensen et al. 2002; Quinn et al. 2001). In England stiegen gestillte Babys später mit einer größeren Wahrscheinlichkeit in eine höhere soziale Klasse auf als Kinder, welche mit der Flasche gefüttert wurden (Martin et al. 2007). Der Effekt des Stillens auf die Intelligenz schrumpft, wenn Forscher gestillte Kinder und Flaschenkinder aus denselben Familien vergleichen (Der et al. 2006).

Was bedeuten diese Ergebnisse? Haben klügere Mütter (die in modernen Ländern häufiger stillen) klügere Kinder? Oder fördern, wie manche Wissenschaftler glauben, die Nährstoffe in der Muttermilch die Gehirnentwicklung? Um solche Fragen zu beantworten – um Ursache und Wirkung voneinander zu trennen – können Forscher **experimentieren**. Mit Hilfe eines Experiments können sich Forscher auf die möglichen Wirkungen eines oder mehrerer Faktoren konzentrieren, indem sie

1. die interessierenden Faktoren manipulieren und
2. andere Faktoren konstant halten (»kontrollieren«).

Dazu schaffen sie oft eine **Versuchsgruppe**, bei der die Teilnehmer eine Behandlung bekommen, und eine **Kontrollgruppe**, bei der dies nicht der Fall ist. Um irgendwelche vorher bestehenden Unterschiede zwischen den zwei Gruppen gering zu halten, weisen die Forscher die Teilnehmer nach dem Prinzip der **Randomisierung** den zwei Bedingungen zu. Die randomisierte Zuweisung macht die zwei Gruppen vergleichbar. Wenn ein Drittel der Personen, die freiwillig an einem Experiment teilnehmen, mit den Ohren wackeln kann, dann wird ungefähr ein Drittel der Teilnehmer in jeder Gruppe zu der Spezies der Ohrenwackler gehören. Auch in Bezug auf andere Faktoren – Alter, Einstellung und weitere Merkmale – werden die Experimental- und Kontrollgruppe ähnlich sein. Deshalb können wir, wenn sich die Gruppen am Ende des Experiments unterscheiden, davon ausgehen, dass die Behandlung einen Effekt hatte.

Experiment (»experiment«) – Forschungsmethode, bei der der Forscher einen oder mehrere Faktoren (unabhängige Variablen) manipuliert, um die Auswirkung auf eine Verhaltensweise oder einen mentalen Prozess (abhängige Variable) zu beobachten. Durch *Zufallszuweisung* der Teilnehmer können andere wichtige Faktoren kontrolliert werden.

Versuchsgruppe (»experimental group«) – Gruppe in einem Experiment, deren Teilnehmer einer Behandlung unterzogen werden, die in diesem Fall eine Ausprägung der unabhängigen Variable darstellt.

Kontrollgruppe (»control group«) – Gruppe in einem Experiment, die *keine* Behandlung erhält; die Kontrollgruppe steht der Versuchsgruppe gegenüber und wird als Vergleich herangezogen, um die Wirkung der Behandlung zu evaluieren.

Randomisierung (auch Zufallszuweisung; »random assignment«) – die Teilnehmer an der Versuchs- und an der Kontrollbedingung werden zufällig ausgewählt. Dadurch wird es höchst unwahrscheinlich, dass die beiden Gruppen sich vorher bereits unterscheiden.

Im Rahmen eines Experiments zum Stillen wies eine Forschergruppe etwa 17.000 Neugeborene aus Weißrussland und ihre Mütter per Zufallsprinzip entweder einer Stillförderungsgruppe oder einem normalen Kinderpflegeprogramm zu (Kramer et al. 2008). In einem Alter von 3 Monaten, wurden 43% der Kinder in der Experimentalgruppe ausschließlich gestillt. In der Kontrollgruppe waren es nur 6%. Im Alter von 6 Jahren wurden beinahe 14.000 der Kinder erneut untersucht und einem Intelligenztest unterzogen. Dabei erzielten die Kinder, welche in der Stillförderungsgruppe waren, im Durchschnitt um 6 Punkte höhere Werte als die Kinder aus der Kontrollgruppe.

Natürlich erlaubt ein einzelnes Experiment keine endgültige Schlussfolgerung, doch in diesem Fall konnten die Wissenschaftler durch die Zufallszuweisung der Versuchspersonen zu der

3 Da viele Assoziationen als Korrelationen angegeben werden, wird das Prinzip meist folgendermaßen ausgedrückt: »Korrelation ist kein Beweis für Kausalität.« Diese Aussage stimmt, sie gilt aber auch für Assoziationen, welche durch andere, nichtexperimentelle Statistiken überprüft wurden (Hatfield et al. 2006).

einen oder anderen Ernährungsgruppe alle Faktoren mit Ausnahme der Ernährung konstant halten. Auf diese Weise konnte man die Schlussfolgerung stützen, dass tatsächlich, soweit es um die Entwicklung der Intelligenz geht, Stillen die beste Art der Ernährung darstellt: Verändert sich ein Verhalten (beispielsweise die Testleistung), wenn wir den Experimentalfaktor variieren (so wie bei der Ernährung dieser Kinder), dann schließen wir daraus, dass dieser Faktor einen Effekt hat.

Merken Sie sich: Im Gegensatz zu Korrelationsstudien, die natürlich auftretende Zusammenhänge aufdecken, wird bei einem Experiment ein Faktor manipuliert, um seinen Effekt zu bestimmen.

Überlegen Sie sich nun, wie wir den Wert therapeutischer Interventionen bemessen könnten. Unsere Neigung, nach neuen Heilmitteln zu suchen, wenn wir krank oder niedergeschlagen sind, kann irreführende Belege hervorbringen. Wenn wir am dritten Tag einer Erkältung Vitamin C nehmen und feststellen, dass sich die Symptome bessern, dann führen wir das auf die Tabletten zurück und nicht darauf, dass die Erkältung automatisch nach ein paar Tagen abklingt. Im 18. Jahrhundert *schien* Aderlass ein wirksames Heilmittel zu sein. Manchmal ging es dem Kranken nach der Behandlung besser; war das nicht der Fall, dann schloss der Arzt daraus, dass die Krankheit schon zu weit fortgeschritten war. Ganz gleich, ob ein Heilmittel tatsächlich heilt oder nicht, es wird immer begeisterte Anhänger finden. Um seinen Effekt festzustellen, müssen wir andere Faktoren kontrollieren.

In genau dieser Weise werden neue Medikamente und neue psychologische Therapiemethoden evaluiert (▶ Kap. 17). Die Versuchspersonen werden in diesen Studien zufällig Untersuchungsgruppen zugewiesen. Eine Gruppe erhält eine Behandlung (beispielsweise ein Medikament), die andere erhält eine Pseudobehandlung – ein wirkungsloses *Placebo* (z. B. eine Tablette ohne Wirkstoff). Die Studienteilnehmer sind oftmals *»blind«*, d. h., sie wissen nicht, welche Behandlung sie erhalten oder ob sie überhaupt behandelt werden. Wenn es sich bei der Studie um einen **Doppelblindversuch** handelt, dann wissen weder die Teilnehmer noch die Assistenten des Versuchsleiters, die die Arznei verabreichen und die Daten sammeln, welche Gruppe das Medikament erhält.

Doppelblindversuch (»double-blind procedure«) – experimentelles Vorgehen, bei dem sowohl die Versuchsteilnehmer als auch die Mitarbeiter des Versuchsleiters nicht wissen (»blind« sind), ob die Teilnehmer eine Behandlung oder ein Placebo erhalten. Diese Methode wird i. Allg. bei der Evaluation von Studien zur Wirkung von Medikamenten angewandt.

Mit solchen Studien können die Wissenschaftler die tatsächliche Wirkung einer Behandlung prüfen, unabhängig davon, ob die Teilnehmer oder die Forschungsgruppe an die heilende Kraft der Behandlung glauben. Schon der bloße *Gedanke*, dass man behandelt wird, kann die Lebensgeister wieder wecken, den Körper entspannen und zu einer Verringerung der Symptome führen. Dieser **Placeboeffekt** ist bei der Reduktion von Schmerzen, Depression und Angst gut dokumentiert (Kirsch 2010). Und je teurer ein Placebo ist, desto »echter« wirkt es auf uns – eine falsche Pille, die 3 Euro kostet, bringt bessere Ergebnisse als eine für 10 Cent (Waber et al. 2008). Um herauszufinden, wie wirksam eine Therapie wirklich ist, müssen Forscher einen möglichen Placeboeffekt kontrollieren.

Placeboeffekt (Aussprache Betonung: Placebo; lateinisch für »Ich werde gefallen«; »placebo effect«) – experimentelle Ergebnisse, die nur durch Erwartungen zustande kommen; jede Auswirkung auf das Verhalten, die durch die Verabreichung einer unwirksamen Substanz hervorgerufen wird, von der der Versuchsteilnehmer jedoch annimmt, dass sie wirkt.

Prüfen Sie Ihr Wissen

- Zu welcher Maßnahme greifen Forscher, um zu verhindern, dass der *Placeboeffekt* ihre Resultate durcheinanderbringt?

Unabhängige und abhängige Variablen

Die Viagrastudie ist ein noch überzeugenderes Beispiel. Viagra erhielt die Zulassung, nachdem das Präparat in 21 klinischen Versuchen getestet worden war. Zu diesen Versuchen gehörte auch ein Experiment, bei dem die Wissenschaftler 329 Männer, die an Impotenz litten, entweder einer Experimentalgruppe (die Viagra erhielt) oder einer Kontrollgruppe (die ein Placebo erhielt) zuwiesen. Das Experiment war als Doppelblindversuch angelegt, d. h., weder die Teilnehmer noch die Assistenten, die das Präparat austeilten, wussten, wer Viagra und wer das Placebo erhielt. Das Ergebnis: Mit der Höchstdosis Viagra war der Geschlechtsverkehr in 69% der Fälle erfolgreich; bei den Männern, die das Placebo erhalten hatten, waren es nur 22% (Goldstein et al. 1998). Viagra hatte demnach einen Effekt.

Dieses einfache Experiment manipulierte nur einen Faktor: Die Dosierung des Arzneistoffes (kein Viagra vs. Höchstdosis). Diesen Experimentalfaktor nennen wir die **unabhängige Variable**. Unabhängig deshalb, weil wir sie *unabhängig* von anderen Faktoren variieren können wie etwa dem Alter der Männer, ihrem Körpergewicht oder ihrer Persönlichkeit. Diese anderen Faktoren, welche die Resultate des Experiments eventuell beeinflussen können, nennt man **Störvariablen**. Die Zufallszuweisung kontrolliert mögliche Störvariablen.

Experimente untersuchen die Wirkung von einer oder mehreren unabhängigen Variablen auf ein messbares Verhalten. Dieses Verhalten wird **abhängige Variable** genannt, weil es sich – *abhängig* von den Ereignissen im Experiment – verändern kann. Beide Variablen werden genauestens operationalisiert. Die *operationale Definition* der Variablen spezifiziert den Prozess, der die unabhängige Variable manipuliert (bei dem Viagraexperiment war das die genaue Viagradosis und das Timing) oder die abhängige Variable misst (die Fragen, mit denen die Reaktionen der Männer erfasst wurden). Diese operationalen Definitionen beantworten die Frage »Was meinen Sie damit?« so genau, dass andere die Untersuchung wiederholen (replizieren) können (s. das Design der Stillstudie in ◘ Abb. 2.15)

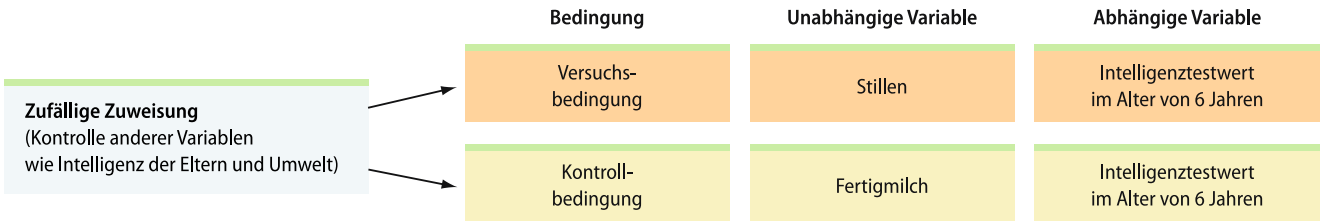

Abb. 2.15 Experiment. Um Kausalität festzustellen, weisen Psychologen einige Versuchsteilnehmer zufällig einer experimentellen Behandlung zu, andere einer Kontrollbedingung. Durch die Messung einer abhängigen Variable (Intelligenztestwert in der späteren Kindheit) wird die Wirkung der unabhängigen Variable bestimmt (Art der Milch)

Unabhängige Variable (»independent variable«) – Faktor im Experiment, der manipuliert wird und dessen Wirkung untersucht wird.

Störvariable (»confounding variable«) – ein anderer Faktor als die unabhängige Variable, der in einem Experiment eine Wirkung entfalten könnte.

Abhängige Variable (»dependent variable«) – Ergebnisfaktor; diese Variable kann sich als Reaktion auf die Manipulationen der unabhängigen Variablen verändern.

Halten wir kurz inne und überprüfen unser Wissen anhand eines einfachen psychologischen Experiments: Um den Effekt der wahrgenommenen ethnischen Zugehörigkeit auf die Bereitschaft, dieser Person eine Haus zu vermieten, zu testen, sandten Adrian Carpusor und William Loges (2006) identisch formulierte Mail-Fragebogen an 1115 Hausbesitzer aus dem Gebiet um Los Angeles. Die Forscher variierten den Namen des Senders, sodass er auf eine bestimmte ethnische Zugehörigkeit hinwies, und erfassten den Prozentsatz an positiven Antworten (Einladungen zur Wohnungsbesichtigung). »Patrick McDougall«, »Said Al-Rahman« und »Tyrell Jackson« wurden zu jeweils 89%, 66% und 56% eingeladen. (Testfrage: Welche war die unabhängige Variable in diesem Experiment? Welche die abhängige?[4])

Experimente können auch hilfreich dabei sein, sozialpolitische Programme zu evaluieren. Lassen sich durch Bildungsprogramme in der frühen Kindheit die Erfolgschancen benachteiligter Kinder erfolgreich verbessern? Worin bestehen die Wirkungen unterschiedlicher Kampagnen gegen das Rauchen? Lässt sich durch Sexualerziehung in der Schule die Anzahl der Teenager-Schwangerschaften verringern? Um diese Fragen zu beantworten, können wir Experimente entwerfen: Wenn ein Programm allgemein gutgeheißen wird, aber die Ressourcen knapp sind, können wir ein Losverfahren dazu nutzen, einige Personen (oder Regionen) nach zufälligen Kriterien einem neuen Programm zuzuweisen und andere einer Kontrollbedingung. Sollten sich die beiden Gruppen später unterscheiden, wurde der Effekt der Intervention bestätigt (Passell 1993).

Lassen Sie uns rekapitulieren: Eine *Variable* ist alles, was variieren kann (Säuglingsnahrung, Intelligenz, Fernsehkonsum – alles, was innerhalb der Grenzen des Möglichen und des ethisch Vertretbaren bleibt). Ein Experiment dient dazu, eine *unabhängige* Variable zu *manipulieren*, eine *abhängige* Variable zu *messen* und alle anderen Variablen durch die randomisierte Gruppenzuweisung zu *kontrollieren*. Ein Experiment besteht aus mindestens zwei verschiedenen Bedingungen: einer *Experimentalbedingung* und einer *Vergleichs-* oder *Kontrollbedingung*. Die *Zufallszuweisung* hat zur Folge, dass – bevor irgendwelche Behandlungseffekte auftreten – die Ausgangsbedingungen in beiden Gruppen so ähnlich wie möglich gehalten werden. Dadurch kann das Experiment Aufschluss über die Wirkung von mindestens einer unabhängigen Variable (dem, was wir manipulieren) auf mindestens eine abhängige Variable (dem Ergebnis, das wir erfassen) geben. In Tab. 2.3 werden die wichtigsten Merkmale der psychologischen Forschungsmethoden miteinander verglichen.

Prüfen Sie Ihr Wissen

- Durch eine *Zufallszuweisung* bekommen Forscher Kontrolle über _____ . Es handelt sich dabei um andere Faktoren nebst der unabhängigen Variable(n), welche die Studienergebnisse beeinflussen könnten.
- Ordnen Sie die Begriffe 1–3 den zugehörigen Beschreibungen a–c zu.
  1. Doppelblindversuch
  2. Zufallsstichprobe
  3. Zufallszuweisung
  a. hilft Forschern, die Ergebnisse der Befragung einer kleinen Stichprobe auf eine größere Population zu generalisieren
  b. hilft, vorher bestehende Unterschiede zwischen der Experimental- und den Kontrollgruppen zu minimieren
  c. kontrolliert den Placeboeffekt; weder die Forscher noch die Teilnehmer wissen, wer die echte Behandlung erhält
- Ein neues blutdrucksenkendes Mittel soll geprüft werden. Warum werden wir mehr Informationen über die Wirksamkeit dieses Mittels bekommen, wenn wir es der Hälfte der 1000 Teilnehmer verabreichen, als wenn wir es allen 1000 Teilnehmern geben?

4 Die unabhängige Variable, welche die Forscher manipulierten, war der Name, der jeweils auf eine andere Ethnie hinwies. Die abhängige Variable, welche die Forscher maßen, war die Rate an positiven Antworten.

Tab. 2.3 Forschungsmethoden im Überblick

Forschungsmethode	Forschungsziel	Praktische Durchführung	Manipuliert wird	Schwächen
Beschreibung	Verhalten beobachten und beschreiben	Fallstudien, naturalistische Beobachtungen oder Befragungen	Nichts	Keine Kontrolle der Variablen; Einzelfälle möglicherweise irreführend
Korrelationsstudie	Aufdeckung naturgegebener Zusammenhänge; Beurteilung, wie gut man eine Variable aufgrund einer anderen vorhersagen kann	Zu zwei oder mehreren Variablen werden Daten gesammelt; keine Manpulation	Nichts	Macht keine Aussage über Ursache und Wirkung
Experiment	Erkundung von Ursache-Wirkungs-Zusammenhängen	Ein oder mehrere Faktoren werden manipuliert; Zufallszuweisung	Die unabhängige(n) Variable(n)	Manchmal nicht durchführbar; Ergebnisse nicht immer über andere Kontexte hinweg generalisierbar; unethisch, bestimmte Variablen zu manipulieren

2.3 Statistische Argumentation im Alltagsleben

In der deskriptiven Forschung, in Korrelationsstudien und Experimenten sind die statistischen Methoden Werkzeuge, mit deren Hilfe man Ergebnisse sehen und interpretieren kann, die ohne statistische Aufbereitung nicht erkennbar wären. Die Forscher Michael Norton und Dan Ariely (2011) baten 5522 Amerikaner zu schätzen, wie viel Prozent des Reichtums die 20% reichsten Menschen in ihrem Land besitzen. Die durchschnittliche Schätzung einer Person lag bei 58%, womit die ungleiche Vermögensverteilung »dramatisch unterschätzt« wurde. (Die reichsten 20% besitzen 84% des Reichtums.)

Wenn man Demokraten und Republikaner fragt, wie das Vermögen in Amerika *ideal* verteilt sein müsste, so liefern sie überraschend ähnliche Antworten. In der idealen Welt der Demokraten würden die reichsten 20% der Bevölkerung 30% des Reichtums besitzen. Republikaner bevorzugten einen ähnlichen Anteil von 35% (Norton u. Ariely 2011).

Akkurates statistisches Wissen bringt allen Vorteile. Als gebildeter Mensch sollte man heutzutage in der Lage sein, einfache statistische Prinzipien im Alltag anzuwenden. Wir müssen uns dabei nicht an komplizierte Formeln erinnern, um klarer und kritischer mit den Daten umzugehen.

Oft wird die Wirklichkeit durch spontane Schätzungen verfälscht, was wiederum dazu führt, dass die Öffentlichkeit falsche Informationen erhält. Jemand wirft eine große runde Zahl in die Diskussion. Andere nehmen sie auf, und in kürzester Zeit wird aus der großen runden Zahl eine öffentliche Fehlinformation. Hier ein paar Beispiele:

- *10% der Bevölkerung sind homosexuell (Männer und Frauen).* Oder sind es nur 2–3%, wie das Ergebnis verschiedener landesweiter Befragungen lautet (▶ Kap. 12)?
- *Wir nutzen normalerweise nur 10% unseres Gehirns.* Oder doch fast 100% (▶ Kap. 3)?
- *Das menschliche Gehirn enthält 100 Milliarden Nervenzellen.* Oder sind es eher etwa 40 Milliarden, wie sich – beruhend auf der Untersuchung eines Gewebeausschnittes – hochrechnen lässt (▶ Kap. 3)?

Merken Sie sich: Große runde Zahlen, die nicht belegt sind, sollten Sie immer mit Vorsicht betrachten.

Statistisches Unwissen löst auch unnötige Gesundheitssorgen aus (Gigerenzer et al. 2008, 2009, 2010). In den 1990er Jahren berichtete die britische Presse über eine Studie zu einer bestimmten empfängnisverhütenden Pille. Es war zu lesen, dass Frauen, die diese Pille einnehmen, ein um 100% erhöhtes Risiko für Blutgerinnsel haben und dass dies zu Schlaganfällen führen kann. Als Folge setzten Tausende von Frauen die Pille ab. Es kam zu einer Welle von ungewollten Schwangerschaften und ungefähr 13.000 Abtreibungen (die ebenfalls mit einem erhöhten Thromboserisiko in Verbindung stehen). Doch was hat die erwähnte Studie eigentlich gefunden? Tatsächlich ein um 100% erhöhtes Risiko – es handelte sich dabei aber um einen Anstieg von 1 von 7000 Fällen auf 2 von 7000 Fällen. Solche falschen Alarme zeigen, wie wichtig es ist, den Umgang mit Statistik zu lehren und statistische Informationen transparenter zu präsentieren.

2.3.1 Datenbeschreibung

2.7 Wie können wir mit Hilfe der Maße der zentralen Tendenz und der Variabilität Daten beschreiben?

Sind die Rohdaten gesammelt, besteht die erste Aufgabe der Wissenschaftler darin, sie sinnvoll zu ordnen (Abb. 2.16). Eine Möglichkeit dafür ist die Verwendung eines einfachen *Säulendiagramms* wie in Abb. 2.17, die die Verteilung von Lastwagen verschiedener Marken zeigt, die nach 10 Jahren noch in Betrieb sind. Bei derartigen Diagrammen ist allerdings Vorsicht geboten.

Abb. 2.16 (© Patrick Hardin / Search ID: phan123, Rights Available from CartoonStock.com)

Je nachdem, was betont werden soll, kann das Diagramm so gezeichnet werden, dass Unterschiede groß (Abb. 2.17a) oder gering (Abb. 2.17b) ausfallen. Das Geheimnis liegt darin, wie man die Ordinate (die *y-Achse*) beschriftet.

Merken Sie sich: Lassen Sie sich nichts vormachen. Wenn Ihnen in Zeitschriften oder im Fernsehen Abbildungen präsentiert werden, lesen Sie die Bezeichnung der Achsen und achten Sie auf den dargestellten Bereich (Variationsbreite, »range«).

Maße der zentralen Tendenz

Der nächste Schritt bei der Bearbeitung der Rohdaten ist die Zusammenfassung mit Hilfe der Maße zur »zentralen Tendenz«, eines einziges Wertes, mit dessen Hilfe die Gesamtmenge der Werte dargestellt wird. Die einfachste Maßzahl ist der **Modalwert**, das ist der Wert, der am häufigsten auftritt (bzw. die Werte, die am häufigsten auftreten). Die bekannteste Maßzahl ist der **Mittelwert**, das arithmetische Mittel. Dabei werden sämtliche Werte addiert und durch die Anzahl der Werte dividiert. Der **Median** teilt die Menge der Daten genau in der Mitte; er entspricht also dem 50. Perzentil. Wenn Sie die Werte der Größe nach grafisch anordnen, liegt die eine Hälfte unter dem Median, die andere Hälfte darüber.

Modalwert (»mode«) – der Wert oder die Werte, die in einer Verteilung am häufigsten auftreten.

Mittelwert (»mean«) – das arithmetische Mittel wird berechnet durch die Addition sämtlicher Werte; diese Summe wird durch die Gesamtzahl der Werte dividiert.

Median (»median«) – der mittlere Wert in einer Verteilung; eine Hälfte der Werte liegt unterhalb, die andere Hälfte oberhalb des Medianwertes.

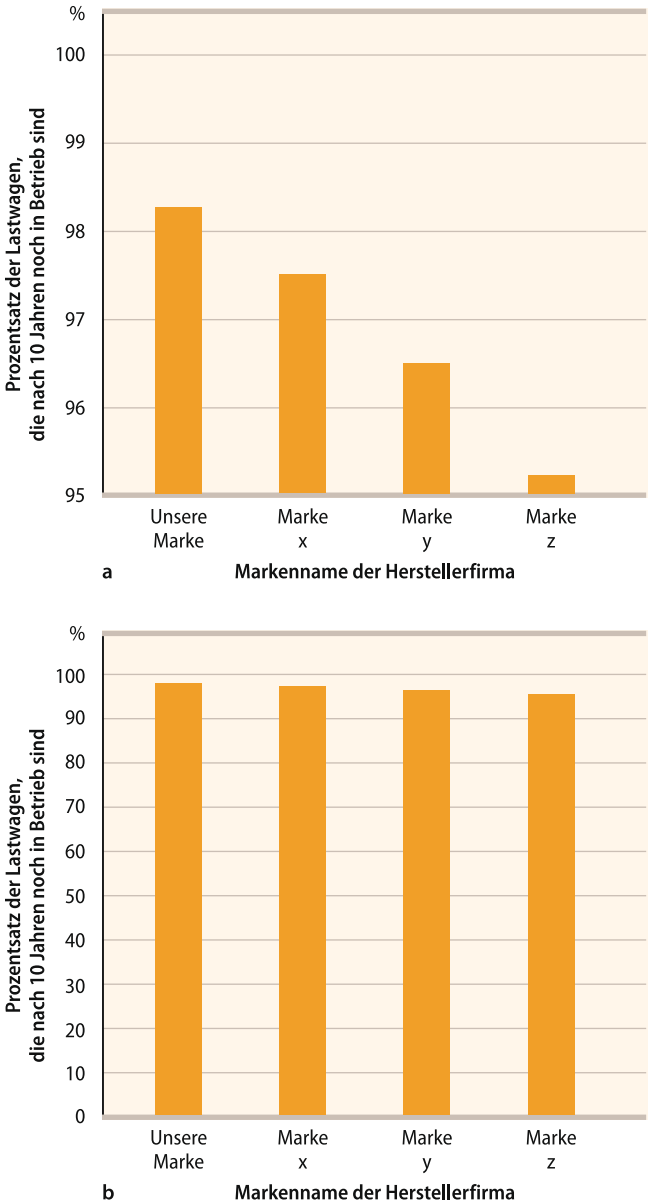

Abb. 2.17a,b Achten Sie auf die Skaleneinteilung

Die Maße der zentralen Tendenz bringen die Daten in eine gewisse Ordnung. Doch achten Sie darauf, was mit dem Mittelwert geschieht, wenn die Verteilung wegen ein paar extremer Werte nach einer Seite hängt, also *schief* ist. Wenn wir Daten betrachten, die das Einkommen beschreiben, dann berichten Median, Modal- und Mittelwert jeweils sehr unterschiedliche Geschichten (Abb. 2.18). Das liegt daran, dass der Mittelwert durch ein paar Extremwerte verfälscht wird. Wenn sich Bill Gates, der Mitbegründer von Microsoft, in ein kleines Café setzt, wird der Durchschnittsgast (Mittelwert) mit einem Mal zum Milliardär. Das mittlere Einkommen (Median) im Café hat sich jedoch nicht verändert. Wenn Sie das verstanden haben, können Sie nachvollziehen, weshalb eine englische Tageszeitung richtigerweise mit folgender Schlagzeile erscheinen konnte: »Das Einkommen von

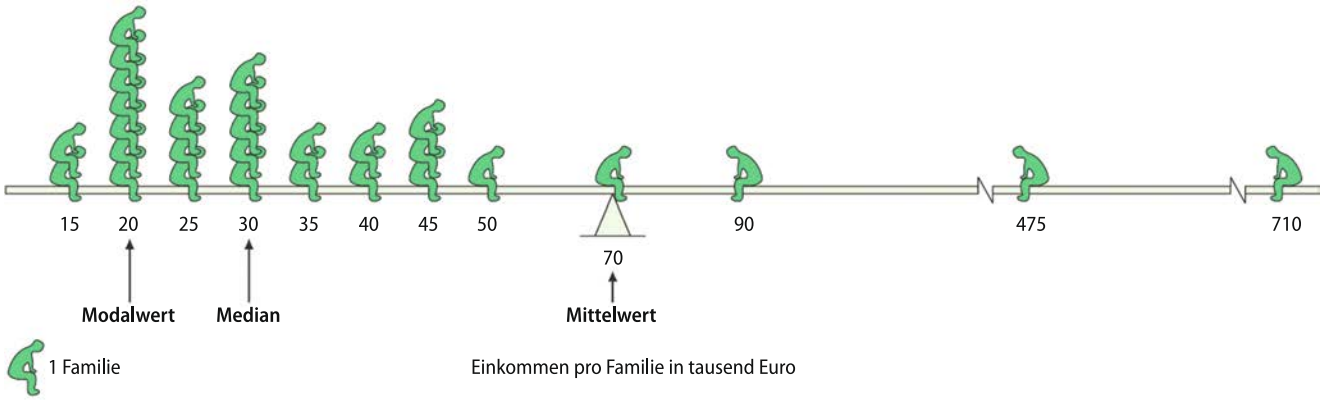

Abb. 2.18 Schiefe Verteilung. Diese grafische Darstellung der Einkommensverteilung in einem Dorf illustriert die drei Maße der zentralen Tendenz: Modalwert, Median und Mittelwert. Beachten Sie, wie einige wenige hohe Einkommen den Mittelwert – den Dreh- und Angelpunkt, der die Einkünfte oberhalb und unterhalb austariert – nach oben verschieben

62% der Bevölkerung liegt unter dem Durchschnitt« (Waterhouse 1993). Weil die untere *Hälfte* der englischen Arbeitnehmer nur ein *Viertel* des gesamten Volkseinkommens erhält, liegt das Einkommen der meisten Engländer unter dem Mittelwert. Das gilt aber nicht nur für die Engländer, sondern ist in allen Ländern der Erde so. Mittelwert und Median besagen beide etwas Wahres, aber eben auch etwas Unterschiedliches.

Der Durchschnittsmensch hat einen Eierstock und einen Hoden.

Merken Sie sich: Achten Sie immer darauf, welches Maß der zentralen Tendenz berichtet wird. Handelt es sich dabei um den Mittelwert, dann schauen Sie nach, ob nicht ein paar untypische Werte den Mittelwert verzerren.

Prüfen Sie Ihr Wissen

- Ein amerikanischer Lastwagenhersteller präsentierte die Grafik in Abb. 2.17a mit tatsächlichen Markennamen, um die viel längere Haltbarkeit seiner Lastwagen hervorzuheben. Was macht die Grafik in Abb. 2.17b in Bezug auf die verschiedenen Haltbarkeiten klar und wie schafft sie das?

Maße der Variabilität

Die Maße der zentralen Tendenz können uns bereits eine große Menge an Informationen vermitteln. Doch in einer einzigen Zahl sind nicht alle Informationen enthalten. Es ist hilfreich, über die *Variabilität* (Streuung) innerhalb der Datenmenge Bescheid zu wissen, beispielsweise, ob die Werte dicht beieinander liegen oder sehr unterschiedlich ausfallen. Mittelwerte von Daten mit geringer Variabilität sagen mehr aus als Mittelwerte von Daten mit hoher Variabilität. Nehmen wir einmal die Handballer als Beispiel: Von einer Spielerin, die in den ersten 10 Spielen der Saison jeweils zwischen 4 und 6 Treffer erzielte, erwarten wir, dass sie in ihrem nächsten Spiel etwa 5-mal trifft. Hätte ihre Trefferquote zwischen 2 und 10 geschwankt, dann könnten wir unserer Prognose nicht allzu sehr vertrauen. Die **Variationsbreite** der Daten, d. h. der Abstand zwischen dem höchsten und dem niedrigsten Wert, liefert nur eine grobe Schätzung der Variation, denn schon ein paar »Ausreißer« können eine sonst durchaus einheitliche Gruppe sprengen. In Abb. 2.18 ist zu sehen, wie zwei Einkommen von 475.000 Euro und 710.000 Euro für eine fälschlich viel zu große Variationsbreite sorgen.

Variationsbreite (»range«) – Differenz zwischen dem höchsten und dem niedrigsten Wert einer Verteilung.

Ein besserer Standard, um zu erfassen, wie stark die Werte voneinander abweichen, ist die **Standardabweichung**. In die Berechnung der Standardabweichung fließen Informationen über jeden einzelnen Messwert ein. Dadurch lässt sich besser erkennen, ob die Daten eng beieinander liegen oder über die gesamte Variationsbreite verstreut sind. Bei der Berechnung der Standardabweichung wird berücksichtigt, um wie viel sich die Messwerte der einzelnen Individuen vom Mittelwert unterscheiden (Tab. 2.4). Ein Beispiel: Wenn eine Universität nur Studenten mit einem bestimmten Intelligenz- und Bildungsniveau aufnimmt, dann wird die Standardabweichung der Ergebnisse eines Intelligenztests mit dieser Population deutlich geringer sein als die Standardabweichung bei einem Intelligenztest, der mit einer heterogeneren Population außerhalb der Universität durchgeführt wird.

Standardabweichung (»standard deviation«) – berechnete Maßzahl, die die Streuung der Daten um den Mittelwert angibt.

Die Bedeutung der Standardabweichung wird klar, wenn Sie sich vor Augen führen, wie Werte in einem natürlichen Umfeld verteilt sind. Große Datenmengen – Angaben zu Größe oder Gewicht, Intelligenzwerte, Noten (jedoch nicht Einkommensbezifferungen) – bilden oft eine symmetrische, *glockenförmige*

Tab. 2.4 Berechnung der Standardabweichung anhand von Stichproben-Testwerten aus zwei Klassen. Die Standardabweichung lässt sich mithilfe der Formel berechnen, die zuunterst in der Tabelle zweimal notiert ist. Hier berechnen und vergleichen wir die Standardabweichungen der Testwerte aus zwei verschiedenen Klassen. Beachten Sie, dass die Testwerte in Klasse A und Klasse B denselben Mittelwert (80) aufweisen, aber ganz unterschiedliche Standardabweichungen. Letztere sagen mehr darüber aus, welche Leistungen die Schülerinnen und Schüler in jeder Klasse wirklich zeigen

Testwerte in Klasse A			Testwerte in Klasse B		
Testwert	Abweichung vom Mittelwert	Quadrierte Abweichung	Testwert	Abweichung vom Mittelwert	Quadrierte Abweichung
72	–8	64	60	–20	400
74	–6	36	60	–20	400
77	–3	9	70	–10	100
79	–1	1	70	–10	100
82	+2	4	90	+10	100
84	+4	16	90	+10	100
85	+5	25	100	+20	400
87	+7	49	100	+20	400
Summe = 640	Summe der (Abweichungen)^2 =	204	Summe = 640	Summe der (Abweichungen)^2 =	2000
Mittelwert = 640 : 8 = 80			Mittelwert = 640 : 8 = 80		
Standardabweichungsformel: $\sqrt{\frac{\text{Summe der (Abweichungen)}^2}{\text{Anzahl der Werte}}} = \sqrt{\frac{204}{8}} = 5.0$			Standardabweichungsformel: $\sqrt{\frac{\text{Summe der (Abweichungen)}^2}{\text{Anzahl der Werte}}} = \sqrt{\frac{2000}{8}} = 15.8$		

Verteilung. Die meisten Fälle ordnen sich um den Mittelwert an; an den beiden Extremen finden wir weniger Fälle. Diese glockenförmige Verteilung ist so typisch, dass wir sie als **Normalverteilung** bezeichnen.

Wie Abb. 2.19 zeigt, besteht eine nützliche Eigenschaft der Normalverteilung darin, dass ungefähr 68% der Fälle im Bereich von einer Standardabweichung um den Mittelwert (in beide Richtungen) zu liegen kommen. Ungefähr 95% der Fälle befinden sich innerhalb von zwei Standardabweichungen. Darum werden – wie in ▶ Kap. 11 erwähnt wird – bei einem Intelligenztest ungefähr 68% der Menschen einen Wert zwischen 85 und 115 Punkten erreichen, der also im Bereich ±15 Punkte von 100 liegt. Ungefähr 95% erzielen einen Wert im Bereich ±30 Punkte.

Normalverteilung (»normal curve«, »normal distribution«) – symmetrische, glockenförmige Kurve, mit der die Verteilung vieler Datentypen beschrieben wird. Die meisten Werte finden sich in der Nähe des Durchschnitts (ungefähr 68% liegen im Bereich einer Standardabweichung links oder rechts des Durchschnitts). Je weiter man sich zu den Extremen hin bewegt, desto weniger Werte findet man.

Prüfen Sie Ihr Wissen

- Der Durchschnitt einer Werteverteilung ist der _____ . Der Wert, der am häufigsten auftritt, ist der _____ . Den Wert, der sich genau in der Mitte einer Verteilung befindet (die Hälfte der Werte liegt über, die andere Hälfte unter ihm) nennt man _____ . Wir bestimmen, wie die Messwerte um den Mittelwert streuen, indem wir die _____ berechnen. Dabei erhalten wir auch Informationen zur _____ von Werten (Unterschied zwischen dem größten und dem kleinsten Wert).

2.3.2 Signifikante Unterschiede

2.8 Wie können wir herausfinden, ob sich ein beobachteter Unterschied auf eine andere Population übertragen lässt?

Daten sind mehrdeutig. Wenn sich das Durchschnittsergebnis einer Gruppe (Muttermilchbabys) von dem der anderen Gruppe (Fertigmilchbabys) unterscheidet: Wie können wir sicher sein, dass der Unterschied tatsächlich besteht und nicht durch zufällige Fluktuationen in der Stichprobe hervorgerufen wurde? Wie viel Vertrauen können wir in unsere Schlussfolgerung setzen,

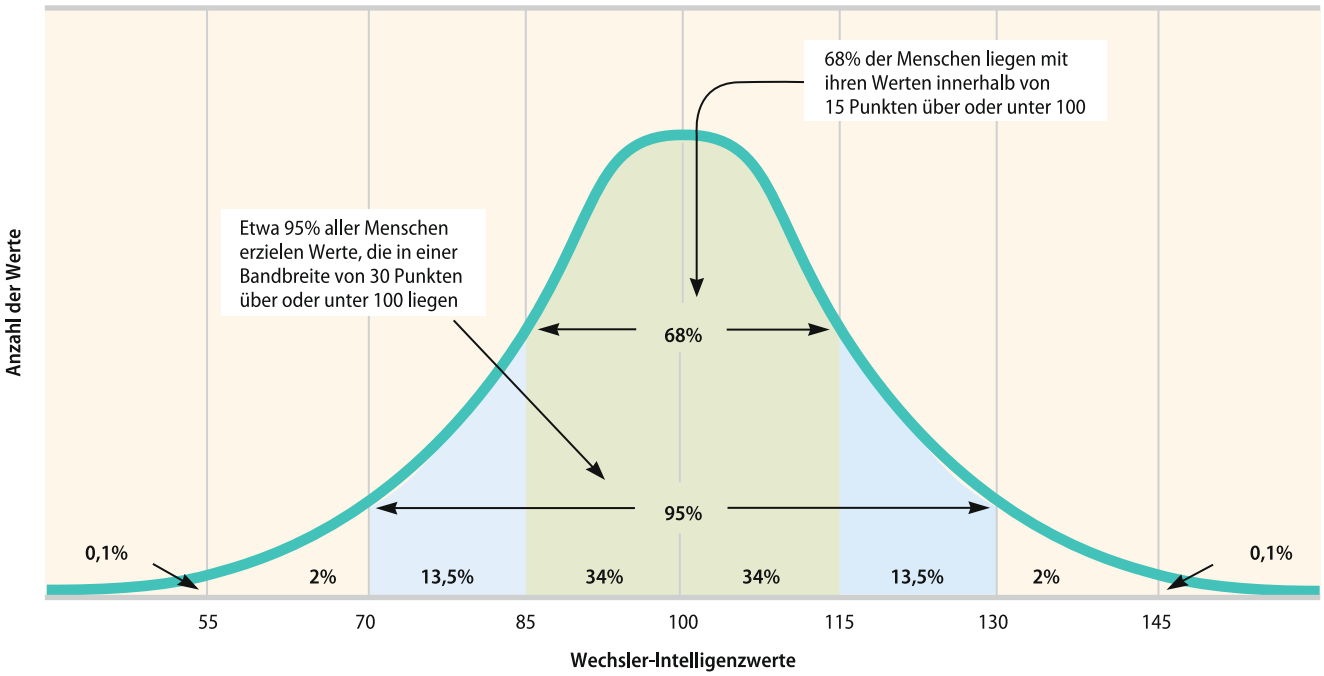

Abb. 2.19 Normalverteilung. Die bei einem Eignungstest erzielten Werte haben die Tendenz, eine glockenförmige Normalverteilung zu bilden. Beim »Wechsler Intelligenztest für Erwachsene« etwa ist der Mittelwert 100

dass die beobachtete Differenz nicht bloß ein Zufallsergebnis ist? Zur Orientierung können wir fragen, wie reliabel (zuverlässig) und signifikant die Unterschiede sind.

Wann ist ein beobachteter Unterschied reliabel (zuverlässig)?

Wenn wir entscheiden müssen, ob unser Ergebnis zuverlässig genug ist, um eine Generalisierung auf Basis der Stichprobe zu erlauben, müssen wir drei Grundsätze beachten.

1. **Repräsentative Stichproben sind besser als verzerrte Stichproben**. Nicht die außergewöhnlichen und denkwürdigsten Einzelfälle, die man an den Rändern einer Verteilung findet, sind die beste Grundlage für eine Verallgemeinerung, sondern eine repräsentative Stichprobe von Fällen. Keine Studie umfasst eine repräsentative Stichprobe der gesamten Menschheit. Deshalb muss man darauf achten, aus welcher Population eine Studie ihre Stichprobe genommen hat.
2. **Weniger variierende Beobachtungen sind zuverlässiger als jene, die eine größere Variation aufweisen.** Wir haben dazu das Beispiel der Handballspielerin gesehen, deren Punktzahl nur geringfügig variierte: Ein Durchschnittswert vermittelt zuverlässigere Informationen, wenn er aus Scores mit geringer Streuung berechnet wird.
3. **Mehr Fälle sind besser als wenige.** Ein eifriger künftiger Student besucht zwei Universitäten und verbringt an jeder einen Tag. In der ersten Universität besucht er zwei zufällig ausgewählte Vorlesungen und findet die Professoren witzig und anregend. In der zweiten Uni sind die zufällig gewählten Professoren langweilig und wenig anregend. Zu Hause berichtet der Student (er lässt die kleine Stichprobengröße von nur zwei Dozenten in jeder Hochschule außer Acht) dann seinen Freunden von den »tollen Lehrern« der ersten Uni und den »schrecklichen Langweilern« der zweiten. Hier haben wir wieder den Fall, den wir so gut kennen und doch immer wieder vergessen: Durchschnittswerte, die auf der Grundlage von vielen Einzelfällen berechnet werden, sind zuverlässiger (weniger variabel) als Durchschnittswerte auf der Basis einiger weniger Fälle.

Merken Sie sich: Clevere Köpfe lassen sich von ein paar Einzelfällen nicht übermäßig beeindrucken. Verallgemeinerungen auf der Basis einiger unrepräsentativer Fälle sind nicht reliabel.

Wann ist ein Unterschied signifikant?

Vielleicht haben Sie die Werte von Männern und Frauen bei einem Aggressionstest im Labor verglichen und einen Geschlechtsunterschied gefunden. Aber Individuen unterscheiden sich immer. Wie wahrscheinlich ist es, dass der Unterschied nur ein Zufallsbefund war? Eine statistische Überprüfung kann die Wahrscheinlichkeit abschätzen, mit der das Ergebnis zufällig entstanden ist.

Die Logik, die hinter diesem Vorgehen liegt, ist die folgende: Wenn die Durchschnittswerte (Mittelwerte) von zwei Stichproben reliable (zuverlässige) Maßzahlen für die jeweilige Population sind (wenn also jeder Mittelwert auf vielen Beobachtungen

Abb. 2.20 (© Peanuts Worldwide LLC./Dist. By Universal Uclick/Distr. Bulls)

beruht und nur wenig Variabilität aufweist), dann ist der *Unterschied* zwischen den Gruppen wahrscheinlich gleichfalls reliabel. (Beispiel: Je geringer die Variabilität bei den jeweiligen Scores zu Aggression bei Männern und Frauen ausfällt, desto sicherer können wir sein, dass der beobachtete Gender- oder Geschlechtsunterschied reliabel ist.) Und wenn die Differenz zwischen den Mittelwerten der Stichproben *groß* ist, dann können wir noch sicherer sein, dass diese Differenz eine echte Differenz in den jeweiligen Populationen spiegelt.

Also: Wenn die Mittelwerte der Stichproben reliabel und der Unterschied zwischen den Mittelwerten relativ groß ist, dann sagen wir: Der Unterschied ist **statistisch signifikant**. Das bedeutet, dass der Unterschied, den wir beobachtet haben, wahrscheinlich nicht auf eine zufällige Variation zwischen den Stichproben zurückzuführen ist.

Statistische Signifikanz (»statistical significance«) – statistische Aussage über die Wahrscheinlichkeit, mit der das Ergebnis einer Untersuchung dem Zufall zuzuschreiben ist.

Psychologen sind konservativ, wenn es darum geht, statistische Signifikanz zu bewerten. Sie sind wie ein Geschworenengericht, das so lange von der Unschuld des Angeklagten auszugehen hat, bis die Schuld bewiesen ist. Für die meisten Psychologen fängt der Beweis über einen vernünftigen Zweifel hinaus erst dann an, wenn die Chance, dass das Ergebnis durch Zufall entstanden ist, unter 5% liegt (Abb. 2.20).

Wenn Sie Forschungsberichte lesen, sollten Sie daran denken, dass bei ausreichend großen oder homogenen Stichproben die Differenz zwischen den Stichproben zwar »statistisch signifikant« sein, aber nur wenig praktische Signifikanz haben kann. So ergab zum Beispiel der Vergleich der Werte bei einem Intelligenztest von Hunderttausenden erstgeborenen Kindern und ihren nachgeborenen Geschwistern eine hoch signifikante Tendenz, dass die Erstgeborenen höhere Werte erzielten als ihre jüngeren Geschwister (Kristensen u. Bjerkedal 2007; Zajonc u. Markus 1975). Da die Scores aber nur um 1 bis 3 Punkte differieren, hat dieser Unterschied wenig praktische Bedeutung.

Merken Sie sich: Statistische Signifikanz drückt die *Wahrscheinlichkeit* aus, mit der ein Ergebnis auf Zufall zurückzuführen ist. Sie sagt nichts über die *Bedeutung* des Ergebnisses aus.

Prüfen Sie Ihr Wissen

- Können Sie dieses Rätsel lösen? – Das Sekretariat der University of Michigan hat festgestellt, dass durchschnittlich 100 Studierende der Geistes- und Naturwissenschaften am Ende ihres ersten Semesters an der Universität hervorragende Noten haben. Doch nur 10–15 Studenten schließen ihr Studium mit hervorragenden Noten ab. Welches ist Ihrer Meinung nach die wahrscheinlichste Erklärung für die Tatsache, dass es am Ende des ersten Semesters mehr hervorragende Noten gibt als beim Studienabschluss (Jepson et al. 1983)?

2.4 Häufig gestellte Fragen zur Psychologie

Wir haben darüber nachgedacht, wie eine wissenschaftliche Vorgehensweise Verzerrungen (Biases) gering halten kann. Wir haben gesehen, wie Verhalten mit Hilfe von Fallstudien, Befragungen und Beobachtungen in einer natürlichen Umgebung beschrieben werden kann. Wir haben auch bemerkt, dass Korrelationsstudien den Zusammenhang zwischen zwei Faktoren erfassen und damit anzeigen, wie gut man einen Faktor aufgrund eines anderen vorhersagen kann. Dann sind wir auf die Logik von Experimenten mit Kontrollbedingung und Zufallszuweisung der Versuchsteilnehmer eingegangen, bei denen die Wirkung einer unabhängigen Variable auf eine abhängige Variable isoliert betrachtet werden kann. Und wir haben darüber nachgedacht, wie uns statistische Werkzeuge helfen können, die Welt um uns herum zu sehen und zu interpretieren.

Aber auch vor dem Hintergrund dieses Wissens werden Sie vielleicht mit einer Mischung aus Neugier und Vorsicht an die Psychologie herangehen. Wir wollen uns deshalb, bevor wir

richtig anfangen, näher mit einigen typischen Fragen beschäftigen.

2.9 Können Laborversuche etwas über den Alltag aussagen?

Wenn Sie einen Artikel über psychologische Forschung lesen oder etwas darüber hören, fragen Sie sich vielleicht, ob das Verhalten von Menschen im Labor wohl als Vorhersage für ihr Verhalten im realen Leben dienen kann. Kann bei einem Versuch, bei dem das Blinken eines schwachen roten Lichts in einem dunklen Raum wahrgenommen werden soll, irgendetwas Hilfreiches für Nachtflüge herauskommen? Oder folgender Versuch: Einem Mann wird ein Film mit sexuellen Gewaltszenen gezeigt, der Mann ist erregt, und seine Bereitschaft steigt, auf einen Knopf zu drücken, mit dem er, wie er glaubt, einer Frau einen Elektroschock verpasst. Sagt so ein Versuch tatsächlich etwas darüber aus, ob Gewaltpornografie bei Männern die Wahrscheinlichkeit erhöht, eine Frau zu vergewaltigen?

Doch lassen Sie uns, ehe Sie antworten, über ein paar Dinge nachdenken. Die *Absicht* des Versuchsleiters ist es, im Labor eine vereinfachte Realität herzustellen, eine Realität, in der wichtige Merkmale des Alltags simuliert und kontrolliert werden. Flugzeugdesigner benutzen einen Windkanal, um atmosphärische Kräfte unter kontrollierten Bedingungen zu untersuchen. Genauso können Psychologen beim Laborversuch psychische Kräfte unter kontrollierten Bedingungen untersuchen.

Es ist nicht Zweck eines Experiments, alltägliche Verhaltensweisen genau zu reproduzieren, sondern die *theoretischen Prinzipien* zu testen (Mook 1983). Was Aggressionsstudien betrifft: Einen Knopf zu drücken, der einen Elektroschock auslöst, ist vielleicht nicht dasselbe, wie jemanden ins Gesicht zu schlagen; doch das Prinzip ist das gleiche. *Es sind die zugrunde liegenden Prinzipien – nicht die spezifischen Befunde –, mit deren Hilfe Alltagsverhalten erklärt werden kann.*

Wenn Psychologen beispielsweise die Ergebnisse der Aggressionsforschung im Labor auf tatsächliches Gewaltverhalten anwenden, dann verwenden sie ihre Kenntnisse über die theoretischen Grundlagen, auf denen aggressives Verhalten beruht, und damit Prinzipien, die sie in zahlreichen Versuchen verfeinert haben. So verhält es sich auch mit den Erkenntnissen zu Grundlagen des visuellen Systems, die mit Hilfe von Experimenten in einer künstlichen Umgebung entwickelt wurden (wie beispielsweise auf rote Lichter in einem dunklen Raum schauen) und dann auf komplexeres Verhalten (Nachtflug) angewandt werden. Zahlreiche Untersuchungen zu diesem Thema haben gezeigt, dass sich die im Labor herausgearbeiteten Prinzipien tatsächlich auf Alltagsbedingungen anwenden lassen (Anderson et al. 1999).

Merken Sie sich: Als Psychologen interessieren uns weniger bestimmte Verhaltensweisen, uns interessieren die allgemeinen Prinzipien, mit deren Hilfe viele Verhaltensweisen erklärt werden können.

2.10 Ist Verhalten kulturabhängig und geschlechtsspezifisch?

Psychologische Studien werden zu einer bestimmten Zeit an einem bestimmten Ort durchgeführt. Oftmals untersuchen sie Menschen aus Kulturen, die Joseph Henrich, Steven Heine und Ara Norenzayan (2010) als »WEIRD«, also seltsam, bezeichnen. »WEIRD« steht für »*W*estern, *E*ducated, *I*ndustrialized, *R*ich and *D*emocratic«. Die meisten Studienteilnehmer stammen aus solchen Kulturkreisen, jedoch nur 12% der gesamten Menschheit. Was können uns psychologische Studien also wirklich über die Menschen i. Allg. sagen? Wir werden es immer wieder erleben: Die **Kultur**, d. h. die Gemeinschaft von Menschen mit gleichen Vorstellungen und Verhaltensweisen, die von Generation zu Generation weitergegeben werden, ist ein wichtiger Faktor. Unsere jeweilige Kultur formt unser Verhalten. Sie beeinflusst unsere Reaktionsschnelligkeit und Offenheit, unsere Einstellung gegenüber vorehelichem Sex und unsere wechselnden Vorstellungen von der Idealfigur, unsere Tendenz zur Förmlichkeit oder zur Formlosigkeit, unseren Blickkontakt, unsere Distanz beim Gespräch und vieles andere mehr. Sobald wir uns dieser Unterschiede bewusst sind, können wir uns von der Annahme trennen, dass alle anderen Menschen genauso denken und handeln wie wir. Angesichts der ständig steigenden Durchdringung und Vermischung der Kulturen ist eine bewusste Haltung gegenüber diesen Unterschieden dringend erforderlich.

Kultur (»culture«) – dauerhafte Verhaltensweisen, Vorstellungen, Einstellungen, Werte und Traditionen einer Gruppe von Menschen, die von einer Generation zur nächsten überliefert werden.

Unser gemeinsames biologisches Erbe macht uns indessen zu einer großen Familie von Menschen. Es sind dieselben zugrunde liegenden Prozesse, die Menschen überall auf der Welt steuern.

- Ganz egal, ob es sich um Italiener, Franzosen oder Engländern handelt: Bei Menschen, die an Dyslexie (einer Lesestörung) leiden, ist überall auf dieser Welt dieselbe Hirnfunktion gestört (Paulesu et al. 2001).
- Unterschiedliche Sprachen mögen die Kommunikation über die kulturellen Grenzen hinweg behindern. Doch alle Sprachen gehorchen den Prinzipien der Grammatik und Menschen aus verschiedenen Erdteilen können sich mit einem Lächeln oder einem Stirnrunzeln verständigen.
- Menschen verschiedener Kulturen leiden nicht auf die gleiche Art und Weise unter Einsamkeit. Doch über die kulturellen Grenzen hinweg verstärken Schüchternheit, geringes Selbstwertgefühl und Unverheiratetsein die Einsamkeit (Jones et al. 1985; Rokach et al. 2002).

In mancher Hinsicht sind wir wie alle anderen, in mancher Hinsicht sind wir wie manche anderen und in mancher Hinsicht sind wir wie sonst niemand auf der Welt. Die Beschäftigung mit anderen Völkern und Kulturen hilft uns, Ähnlichkeiten und Unterschiede, Gemeinsamkeit und Andersartigkeit zu erkennen (Abb. 2.21).

Abb. 2.21 Ein kulturspezifischer Gruß. Soziales Verhalten wird von der jeweiligen Kultur geprägt, deshalb können Handlungen, die uns ganz normal vorkommen, bei Besuchern aus fremden Ländern Verwunderung hervorrufen. Doch hinter diesen Unterschieden verbergen sich große Ähnlichkeiten. Überall auf der Welt begrüßen sich die Menschen – wenn auch nicht unbedingt so förmlich und respektvoll wie in Japan. (© photos.com)

Auch die Themen »Geschlechtsrolle« und »Geschlechtszugehörigkeit« werden in diesem Buch immer wieder angesprochen. Wissenschaftler fanden Geschlechtsunterschiede in unseren Träumen, in der Art, wie wir Emotionen ausdrücken und bei anderen erkennen, und in dem Risiko, Alkoholiker zu werden, depressiv zu werden oder eine Essstörung zu entwickeln. Geschlechtsunterschiede faszinieren uns. Sie zu studieren, ist außerdem potenziell nützlich. So glauben Wissenschaftler beispielsweise, dass Frauen eine Unterhaltung führen, um eine Beziehung herzustellen, Männer dagegen eher, um Informationen und Ratschläge zu geben (Tannen 2001). Wenn uns dieser Unterschied bewusst ist, können wir Konflikte und Missverständnisse in unseren alltäglichen Beziehungen leichter vermeiden.

Psychologisch und biologisch sind Frauen und Männer allerdings einander sehr ähnlich. Männliche und weibliche Kinder lernen ungefähr zur gleichen Zeit gehen. Ob Mann oder Frau, wir haben die gleiche Licht- und Klangempfindung. Auch das Gefühl von Hunger, Begehren und Angst ist für beide Geschlechter gleich. Intelligenz und Wohlbefinden sind ähnlich.

» Alle Menschen sind gleich, nur ihre Gewohnheiten unterscheiden sich.
Konfuzius, 551–479 v. Chr.

Merken Sie sich: Auch wenn bestimmte Einstellungen und Verhaltensweisen über die kulturellen Grenzen hinweg und je nach Geschlecht variieren, wie es häufig der Fall ist, so sind doch die ihnen zugrunde liegenden Prozesse im Wesentlichen die gleichen.

? **2.11** Warum machen Psychologen Experimente mit Tieren und welche ethischen Richtlinien schützen menschliche und tierische Versuchsteilnehmer?

Viele Psychologen arbeiten mit Tieren, weil sie sie faszinierend finden. Sie wollen verstehen, auf welche Weise die verschiedenen Arten lernen, denken und sich verhalten. Experimente mit Tieren dienen aber auch dazu, mehr über Menschen zu erfahren. Wir Menschen sind nicht *wie* Tiere; wir *sind* Tiere, denn wir teilen dieselbe Biologie. Tierversuche haben daher schon häufig zur Entwicklung von Medikamenten gegen menschliche Krankheiten geführt: Insulin für Diabetiker oder Impfstoffe gegen Polio und Tollwut. Auch Organtransplantationen wurden zunächst an Tieren durchgeführt.

Menschen sind komplex. Aber die gleichen Prozesse, die menschlichem Lernen zugrunde liegen, finden wir auch bei Ratten, Affen, ja sogar bei Meeresschnecken. Gerade der einfache Aufbau des Nervensystems der Meeresschnecke macht sie zu einem guten Modell für die neuronalen Abläufe beim Lernen. Sollten wir Tiere nicht mehr achten, wenn wir einander doch so ähnlich sind? »Wir können unsere wissenschaftliche Arbeit mit Tieren nicht einerseits mit der Ähnlichkeit zwischen ihnen und uns verteidigen und andererseits moralisch mit der Unterschiedlichkeit rechtfertigen«, sagte Roger Ulrich (1991). Die Tierschutzbewegung protestiert gegen die Verwendung von Tieren in der psychologischen, biologischen und medizinischen Forschung.

» Ratten und Menschen sind sich ziemlich ähnlich,
nur sind die Ratten nicht so dumm, Lotterielose zu kaufen.
Dave Barry, 2. Juli 2002

In dieser hitzigen Debatte zeichnen sich zwei Themen ab. Die grundsätzliche Frage lautet, ob wir berechtigt sind, menschliches Wohlbefinden über das der Tiere zu stellen. Haben wir das Recht, Experimente zu Stress und Krebs zu machen und Mäuse Tumore bekommen zu lassen in der Hoffnung, dass wir Menschen keine bekommen? Dürfen wir ein paar Affen einem HIV-ähnlichen Virus aussetzen, wenn wir nach einem Impfstoff gegen Aids suchen? Wenn wir Tiere für unsere Zwecke opfern, verhalten wir uns da so natürlich wie fleischfressende Falken, Katzen oder Wale? Die Antworten auf derartige Fragen sind von Kultur zu Kultur unterschiedlich. Nach einer Umfrage von Gallup in Kanada und den Vereinigten Staaten halten ungefähr 6 von 10 Erwachsenen medizinische Tests an Tieren für »moralisch akzeptabel«. In Großbritannien ist dies nur für 37% der Befragten der Fall (Mason 2003).

Wenn wir dem menschlichen Leben oberste Priorität zugestehen, dann ist das zweite Thema das Wohlergehen der Tiere, mit denen wir Forschung betreiben. Mit welchen Maßnahmen können wir sie schützen? Eine Befragung von Forschern, die mit Tieren arbeiten, lieferte hierzu eine Antwort. Etwa 98% unterstützten die von der Regierung erlassenen Vorschriften zum Schutz von Primaten, Hunden und Katzen; 74% traten dafür ein, den humanen Umgang mit Ratten und Mäusen durch entsprechende Vorschriften zu gewährleisten (Plous u. Herzog 2000). Viele Berufsverbände und Institutionen der Forschungsförderung haben bereits Richtlinien für den humanen Umgang mit Tieren. Auch die Deutsche Gesellschaft für Psychologie (DGPs) und der Berufsverband Deutscher Psychologinnen und Psychologen (BDP) haben

Abb. 2.22 Forschung mit Tieren nützt Tieren. Diese Gorillas im Zoo der Bronx in New York verdanken ihre bessere Lebensqualität teilweise den Untersuchungen zu den positiven Auswirkungen von Anregung, Kontrolle und dem Kontakt mit neuen, unbekannten Dingen. (D. Shapiro, © Wildlife Conservation Society)

gemeinsame ethische Richtlinien zum Einsatz von Tieren in Forschung und Lehre formuliert. Sie stellen eine grundsätzliche Verpflichtung der Psychologen zur Achtung des Lebens fest und fordern, Schmerzen, Leiden und Schäden für Versuchstiere soweit möglich zu verhindern. Die Richtlinien der American Psychological Association fordern, dass Forscher das »Wohl, die Gesundheit und die humane Behandlung« von Tieren gewährleisten und »Infektionen, Krankheit und Schmerzen« auf ein Minimum reduzieren (APA 2002). Das Europäische Parlament ist nun dabei, die Ausarbeitung von Standards für die Behandlung und Unterbringung von Tieren in Auftrag zu geben (Vogel 2010).

» Vergessen Sie bitte nicht jene von uns, die an unheilbaren Krankheiten oder an Behinderungen leiden und die auf Heilung hoffen; hier hilft eine Forschung, bei der Tiere eingesetzt werden.
Der Psychologe Dennis Feeney (1987)

» Die Größe einer Nation kann man an ihrem Umgang mit Tieren ablesen.
Mahatma Gandhi (1869–1948)

Die Tiere selbst sind auch Nutznießer der Tierexperimente (Abb. 2.22). Ein psychologisches Forscherteam in Ohio maß die Ausschüttung von Stresshormonen bei Stichproben von Millionen von Hunden, die jedes Jahr in Tierheime gebracht werden. Die Psychologen entwickelten Methoden, wie man die Tiere behandeln und streicheln müsste, um den Stress zu mildern und ihnen den Übergang in eine neue Familie zu erleichtern (Tuber et al. 1999). Andere Studien trugen dazu bei, die Betreuung und die Behandlung von Tieren in ihrer natürlichen Umgebung zu verbessern. In Versuchen wurde auch die bemerkenswerte Intelligenz von Schimpansen, Gorillas und anderen Tieren nachgewiesen; ebenso wurde die Ähnlichkeit des Verhaltens bei Mensch und Tier deutlich, was wiederum bei den Menschen zu verstärkter Empathie mit ihren Mitgeschöpfen sowie zu deren Schutz führte. Im besten Fall kommt eine Psychologie, die Interesse am Menschen und Einfühlungsvermögen für Tiere hat, beiden zugute.

Wie steht es mit Versuchen an Menschen? Erschreckt Sie das Bild von Wissenschaftlern in weißen Kitteln, welche Elektroschocks verabreichen? Wenn ja, dann werden Sie mit Erleichterung zur Kenntnis nehmen, dass die meisten psychologischen Untersuchungen ohne derartigen Stress durchgeführt werden. In der Regel geht es bei Versuchen mit Menschen um blinkende Lichter, flackernde Wörter und angenehme soziale Interaktionen. Außerdem sind psychologische Experimente harmlos verglichen mit dem Stress und der Demütigung, die in Reality-TV-Shows erzeugt werden. In einer Folge von *Der Bachelor* gab ein Mann – auf Wunsch des Regisseurs – seiner neuen Verlobten vor laufender Kamera den Laufpass und entschied sich für eine Frau, die zuvor auf dem zweiten Platz gelandet war (Collins 2009).

Gelegentlich allerdings müssen Wissenschaftler die Menschen hinters Licht führen oder sie einem kurzen Stress aussetzen, doch nur, wenn es ihnen zur Klärung einer wichtigen Frage unabdingbar erscheint. Ich denke dabei an Versuche zum Verständnis und zur Kontrolle gewalttätigen Verhaltens oder Untersuchungen zu Stimmungsschwankungen. Manche Experimente erbrächten kein Ergebnis, wenn die Teilnehmer vorher wüssten, was auf sie zukommt. (In der Absicht, hilfreich zu sein, würden die Teilnehmer möglicherweise versuchen, die Erwartungen des Versuchsleiters zu erfüllen.)

Die von der American Psychological Association wie auch die von der Deutschen Gesellschaft für Psychologie und dem Berufsverband Deutscher Psychologinnen und Psychologen erarbeiteten ethischen Grundsätze fordern von den Versuchsleitern, dass sie

1. die **informierte Einwilligung** (»informed consent«) der potenziellen Teilnehmer einholen,
2. die Teilnehmer vor schädlichen Einflüssen und Unannehmlichkeiten schützen,
3. Informationen über einzelne Teilnehmer vertraulich behandeln und
4. die Versuchspersonen in einer **Nachbesprechung** (»debriefing«) vollständig aufklären (hinterher die Untersuchung erklären).

Außerdem haben viele Universitäten heute einen Ethikrat eingesetzt, der Forschungsanträge prüft und das Wohlergehen der Teilnehmer sicherstellt.

Informierte Einwilligung (»informed consent«) – ein ethischer Grundsatz, der darin besteht, dass Versuchsteilnehmer genügend informiert werden, um entscheiden zu können, ob sie an einem Versuch teilnehmen möchten.

Nachbesprechung (»debriefing«) – Aufklären der Versuchsteilnehmer nach Abschluss des Experiments über die Studie, inklusive ihres Ziels und der verwendeten Täuschungen.

Abb. 2.23a,b Was sehen Sie auf dem rechten Bild? Unsere Erwartungen beeinflussen unsere Wahrnehmung. Haben Sie hier eine Ente oder ein Kaninchen gesehen? Zeigen Sie dieses Bild einigen Freunden, wobei Sie das nebenstehende Foto des Kaninchens abdecken. Finden Sie heraus, ob Ihre Freunde unter dieser Bedingung eher eine Ente sehen, die auf dem Rücken liegt. (Nach Shepard 1990; a: © Eric Isselée / Fotolia; b: © Roger Shepard)

2.12 Ist die Psychologie frei von Werturteilen?

Die Psychologie ist eindeutig nicht wertfrei. Unser Wertesystem bestimmt, was wir untersuchen, wie wir es untersuchen und wie wir die Ergebnisse interpretieren. Die Wertvorstellungen der Psychologen beeinflussen ihre Wahl des Forschungsthemas. Sollen wir die Produktivität der Arbeiter oder ihre Arbeitsmoral untersuchen? Geschlechtsdiskriminierung oder Geschlechtsrollenunterschiede? Sozial angepasstes oder unabhängiges Verhalten? Die »nackten Tatsachen« sind auch von unseren Überzeugungen beeinflusst. Wie zuvor erwähnt, können unsere Vorannahmen unsere Beobachtungen und Interpretationen verzerren; manchmal sehen wir genau das, was wir sehen wollen oder zu sehen erwarten (Abb. 2.23).

Sogar in den Worten, mit denen wir ein Phänomen beschreiben, können sich unsere Wertvorstellungen spiegeln. Handelt es sich bei sexuellen Praktiken, die uns nicht vertraut sind, um »Perversionen« oder um »Varianten«? Mit Bezeichnungen beschreiben wir nicht nur, sondern bewerten gleichzeitig – in der Psychologie wie auch in unserer Alltagssprache: Was wir *rigide* nennen, nennen andere *unbeugsam*; was dem einen als *fester Glaube* gilt, ist für den anderen *Fanatismus*. Die »*erweiterten Befragungsmethoden*« eines Landes, wie Eintauchen in kaltes Wasser, werden zur *Folter*, wenn sie von dessen Feinden angewandt werden. Wenn wir jemanden als *standhaft* oder *stur*, *vorsichtig* oder *mäkelig*, *verschwiegen* oder *heimlichtuerisch* bezeichnen, kommen darin unsere eigenen Einstellungen zum Ausdruck.

Auch in den gängigen praktischen Anwendungen enthält die Psychologie verborgene Wertvorstellungen. Wenn Sie bei der Gestaltung Ihres Lebens »professionellen« Rat berücksichtigen, etwa, wie Sie Ihre Kinder erziehen, persönliche Erfüllung finden, mit sexuellen Gefühlen umgehen oder mit Ihrer Arbeit vorankommen sollen, dann akzeptieren Sie wertebeladene Antworten. Eine Wissenschaft vom Verhalten und von mentalen Prozessen kann uns helfen, unsere Ziele zu erreichen, doch sie kann nicht entscheiden, welche Ziele wir erreichen sollten.

» Es ist völlig unmöglich, an irgendein menschliches Problem mit unverstelltem Blick heranzugehen.
Simone de Beauvoir, *Das andere Geschlecht*, 1953

Manche Menschen halten Psychologie für nichts weiter als gesunden Menschenverstand, doch andere betrachten sie mit Sorge: Sie befürchten, die Psychologie könnte eine gefährliche Macht erlangen. Ist es Zufall, dass die Astronomie die älteste Wissenschaft ist, die Psychologie dagegen die jüngste? Manche Menschen halten die Erforschung des äußeren Universums für weit sicherer als die Erkundung unseres eigenen inneren Universums. Sie fragen sich, ob die Psychologie dazu benutzt werden könnte, Menschen zu manipulieren.

Wie jede Macht kann die Macht des Wissens für gute oder schlechte Zwecke eingesetzt werden. Atomkraft wird genutzt, um Städte zu beleuchten – und auch, um sie zu zerstören. Überzeugungskraft dient dazu, Menschen zu erziehen – und dazu, sie zu täuschen. Obwohl die Psychologie die Macht hat zu täuschen, liegt ihr Zweck doch darin aufzuklären. Tag für Tag suchen Psychologen nach neuen Wegen, um Lernen, Kreativität und Menschlichkeit zu fördern. Die Psychologie spricht viele der großen Probleme unserer Welt an: Krieg, Überbevölkerung, Vorurteile, zerfallende Familien und Verbrechen; denn alle diese Probleme haben mit Einstellungen und Verhaltensweisen zu tun. Und die Psychologie spricht unser tiefes inneres Verlangen an, unser Bedürfnis nach Nahrung, Liebe und Glück. Freilich kann die Psychologie nicht alle großen Lebensfragen ansprechen, wohl aber einige, die uns auf den Nägeln brennen (Abb. 2.24).

Prüfen Sie Ihr Wissen

- Wodurch werden Versuchsteilnehmer (Mensch oder Tier) geschützt?

Abb. 2.24 Die Psychologie sagt aus. 1954 fällte das oberste Gericht der USA, der Supreme Court, seine historische Entscheidung, dass nach Rassen getrennte Schulen verfassungswidrig sind. Es berief sich dabei auf die Expertenaussage und Forschung der Psychologen Kenneth Clark und Mamie Phipps Clark (1947). Die Clarks berichteten, dass, wenn man afroamerikanischen Kindern die Wahl zwischen schwarzen und weißen Puppen ließ, sich die meisten von ihnen für die weißen entschieden. Dies war ein Hinweis auf internalisierte Vorurteile gegenüber Schwarzen. (Image courtesy of University Archives, Columbia University in the City of New York)

2.5 Kapitelrückblick

2.5.1 Verständnisfragen

Die Notwendigkeit der Psychologie als Wissenschaft

2.1 Wie illustrieren der Verzerrungseffekt durch nachträgliche Einsicht (Hindsightbias), die übertriebene Selbstsicherheit und die Tendenz, in zufälligen Ereignissen eine Ordnung wahrzunehmen, weshalb wissenschaftlich fundierte Antworten der Intuition und dem gesunden Menschenverstand überlegen sind?

2.2 Wie hängen die drei Hauptkomponenten der wissenschaftlichen Haltung mit dem kritischen Denken zusammen?

Wie stellen und beantworten Psychologen Fragen?

2.3 Wie bringen Theorien die Psychologie als Wissenschaft voran?

2.4 Wie nutzen Psychologen Einzelfallstudien, Feldbeobachtungen und Umfragen, um Verhalten zu beobachten und zu beschreiben, und weshalb ist es wichtig, Zufallsstichproben zu ziehen?

2.5 Was sind positive und negative Korrelationen und wieso ermöglichen sie Vorhersagen, nicht aber Ursache-Wirkungs-Erklärungen?

2.6 Welches sind die Eigenschaften von Experimenten, die es möglich machen, Ursache und Wirkung auseinanderzuhalten?

Statistische Argumentation im Alltagsleben

2.7 Wie können wir mit Hilfe der Maße der zentralen Tendenz und der Variabilität Daten beschreiben?

2.8 Wie können wir herausfinden, ob sich ein beobachteter Unterschied auf eine andere Population übertragen lässt?

Häufig gestellte Fragen zur Psychologie

2.9 Können Laborversuche etwas über den Alltag aussagen?

2.10 Ist Verhalten kulturabhängig und geschlechtsspezifisch?

2.11 Warum machen Psychologen Experimente mit Tieren und welche ethischen Richtlinien schützen menschliche und tierische Versuchsteilnehmer?

2.12 Ist die Psychologie frei von Werturteilen?

2.5.2 Schlüsselbegriffe

- Abhängige Variable
- Befragung
- Beobachtung in natürlicher Umgebung
- Doppelblindversuch
- Einzelfallstudie
- Experiment
- Hindsightbias
- Hypothese
- Informierte Einwilligung
- Kontrollgruppe
- Korrelation
- Korrelationskoeffizient
- Kritisches Denken
- Kultur
- Median
- Mittelwert
- Modalwert
- Nachbesprechung
- Normalverteilung
- Operationale Definition
- Placeboeffekt
- Population
- Randomisierung
- Replikation
- Standardabweichung
- Statistische Signifikanz
- Störvariable
- Streudiagramm
- Theorie
- Unabhängige Variable
- Variationsbreite
- Versuchsgruppe
- Zufallsstichprobe

http://www.springer.com/978-3-642-40781-9

Psychologie
Myers, D.G.
2014, Etwa 1000 S. 900 Abb. in Farbe. Mit Online-Extras.,
Hardcover
ISBN: 978-3-642-40781-9