



Markus Antonius
Wirtz

Basiswissen Statistische Methoden

Grundlagen und Anwendung
in den Therapie- und
Gesundheitswissenschaften



Basiswissen Statistische Methoden

Basiswissen Statistische Methoden

Markus Antonius Wirtz

Programmbereich Gesundheitsberufe

Wissenschaftlicher Beirat Programmbereich Gesundheitsberufe

Sophie Karoline Brandt, Bern; Jutta Berding, Osnabrück; Heidi Höppner, Berlin; Sinje Gehr, Göttingen;
Heike Kubat, Feldbach; Christiane Mentrup, Zürich; Sascha Sommer, Bochum; Birgit Stubner,
Regensburg; Ursula Walkenhorst, Osnabrück; Claudia Winkelmann, Berlin

Markus Antonius Wirtz

Basiswissen Statistische Methoden

Grundlagen und Anwendung in den Therapie-
und Gesundheitswissenschaften

Wichtiger Hinweis: Der Verlag hat gemeinsam mit den Autoren bzw. den Herausgebern große Mühe darauf verwandt, dass alle in diesem Buch enthaltenen Informationen (Programme, Verfahren, Mengen, Dosierungen, Applikationen, Internetlinks etc.) entsprechend dem Wissensstand bei Fertigstellung des Werkes abgedruckt oder in digitaler Form wiedergegeben wurden. Trotz sorgfältiger Manuskripterstellung und Korrektur des Satzes und der digitalen Produkte können Fehler nicht ganz ausgeschlossen werden. Autoren bzw. Herausgeber und Verlag übernehmen infolgedessen keine Verantwortung und keine daraus folgende oder sonstige Haftung, die auf irgendeine Art aus der Benutzung der in dem Werk enthaltenen Informationen oder Teilen davon entsteht. Geschützte Warennamen (Warenzeichen) werden nicht besonders kenntlich gemacht. Aus dem Fehlen eines solchen Hinweises kann also nicht geschlossen werden, dass es sich um einen freien Warennamen handelt.

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://www.dnb.de> abrufbar.

Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Kopien und Vervielfältigungen zu Lehr- und Unterrichtszwecken, Übersetzungen, Mikroverfilmungen sowie die Einspeicherung und Verarbeitung in elektronischen Systemen.

Anregungen und Zuschriften bitte an:

Hogrefe AG
Lektorat Gesundheitsberufe
z.Hd. Barbara Müller
Länggass-Strasse 76
3012 Bern
Schweiz
Tel. +41 31 300 45 00
info@hogrefe.ch
www.hogrefe.ch

Lektorat: Barbara Müller
Herstellung: Daniel Berger
Umschlagabbildung: © Klaus Vedfelt, GettyImages.com
Umschlag: Claude Borer
Satz: Claudia Wild, Konstanz
Druck und buchbinderische Verarbeitung: Finidr s.r.o., Český Těšín
Printed in Czech Republic

1. Auflage 2024
© 2024 Hogrefe Verlag, Bern

(E-Book-ISBN_PDF 978-3-456-96280-1)
ISBN 978-3-456-86280-4
<https://doi.org/10.1024/86280-000>

Nutzungsbedingungen

Der Erwerber erhält ein einfaches und nicht übertragbares Nutzungsrecht, das ihn zum privaten Gebrauch des E-Books und all der dazugehörigen Dateien berechtigt.

Der Inhalt dieses E-Books darf von dem Kunden vorbehaltlich abweichender zwingender gesetzlicher Regeln weder inhaltlich noch redaktionell verändert werden. Insbesondere darf er Urheberrechtsvermerke, Markenzeichen, digitale Wasserzeichen und andere Rechtsvorbehalte im abgerufenen Inhalt nicht entfernen.

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

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

Das Anfertigen von Vervielfältigungen, das Ausdrucken oder Speichern auf anderen Wiedergabegeräten ist nur für den persönlichen Gebrauch gestattet. Dritten darf dadurch kein Zugang ermöglicht werden. Davon ausgenommen sind Materialien, die eindeutig als Vervielfältigungsvorlage vorgesehen sind (z. B. Fragebögen, Arbeitsmaterialien).

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

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

Inhaltsverzeichnis

Vorwort	11
Danksagung	17
1 Statistik und Forschungsmethoden in der gesundheitswissenschaftlichen Forschung	19
1.1 Angewandte wissenschaftliche empirische Forschung	20
1.1.1 Alltagstheorien und -überzeugungen	21
1.1.2 Wesentliche Merkmale von Alltagstheorien und -überzeugungen	22
1.1.3 Kriterien verlässlichen Wissens in den empirischen Wissenschaften	24
1.2 Empirische Forschung und Statistik	32
1.2.1 Inhaltliche Ebene und statistische Ebene	34
1.2.2 Der empirische Forschungsprozess	38
1.2.3 Phasen des Forschungsprozesses	40
1.2.4 Grundlegende Merkmale der Anwendung statistischer Verfahren in den Gesundheitswissenschaften	44
1.3 Zusammenfassung	49
1.4 Blitzlichter aus der Forschungspraxis	50
2 Erfassung von Merkmalen: Messen und Skalen	53
2.1 Klärung des Begriffs „Messen“	53
2.2 Skalenniveau von Messergebnissen	56
2.2.1 Nominalskalenniveau	58
2.2.2 Ordinalskalenniveau	61
2.2.3 Intervallskalenniveau	63
2.2.4 Verhältnisskalenniveau	67
2.2.5 Absolutskalenniveau	68
2.2.6 Die Rangordnung der Skalentypen	68
2.3 Zusammenfassung	68
2.4 Blitzlichter aus der Forschungspraxis	70
3 Darstellung der Verteilung eines Merkmals: Tabellen und Grafiken	75
3.1 Empirische Merkmalsverteilungen	75
3.2 Häufigkeitsverteilungen kategorialer Merkmale	78

3.2.1	Absolute und relative Häufigkeiten	78
3.2.2	Tabellarische Darstellung nominalskaliertter Merkmale	78
3.2.3	Grafische Darstellung nominalskaliertter Merkmale	81
3.2.4	Tabellarische und grafische Darstellung mindestens ordinalskaliertter Merkmale mit wenigen diskreten Merkmalsstufen . .	82
3.2.5	Verteilung stetig verteilter bzw. fein gestufter mindestens intervallskaliertter Merkmale	83
3.3	Zusammenfassung	87
3.4	Blitzlichter aus der Forschungspraxis	87
4	Kennzahlen der Verteilung eines Merkmals: Zentrale Tendenz, Variabilität und Schiefe	91
4.1	Maße der zentralen Tendenz: Modus, Median, arithmetisches Mittel	92
4.2	Maße der Schwankungsbreite oder Variabilität eines Merkmals: Spannweite, Varianz, Standardabweichung	93
4.3	Die Schiefe einer Verteilung	96
4.3.1	Ausreißerwerte	99
4.4	Standardisierung der Verteilung eines Merkmals	102
4.5	Zusammenfassung	107
4.6	Blitzlichter aus der Forschungspraxis	108
5	Modellierung von Zufallsschwankungen: Wahrscheinlichkeiten, statistische Verteilungen und Wahrscheinlichkeitsbereiche	111
5.1	Wahrscheinlichkeiten	112
5.2	Verteilung einer Zufallsvariable bei einmaliger Beobachtung: Bernoulli-Verteilung, Gleichverteilung	115
5.3	Verteilung der Häufigkeit eines Merkmals bei mehrmaliger Beobachtung: Binomialverteilung	116
5.4	Bestimmung der Wahrscheinlichkeitswerte einer Binomialverteilung	119
5.5	Verteilung eines Merkmals, dessen Ausprägung durch viele unabhängige Zufallsprozesse zustande kommt: Die Gauß'sche Normalverteilung	122
5.6	Bestimmung der Wahrscheinlichkeitswerte einer Normalverteilung	123
5.7	Zusammenfassung	125
5.8	Blitzlichter aus der Forschungspraxis	126
6	Unsicherheit geschätzter Merkmalsausprägungen: Konfidenzintervalle für Mittelwerte	129
6.1	Untersuchung von Stichproben als Basis der Schätzung von Populationseigenschaften	129
6.1.1	Repräsentative Stichprobenauswahl	130
6.1.2	Bestimmung deskriptiver Statistiken in einer Stichprobe	131

6.2	Schätzung des Standardfehlers der Kennwerteverteilung der Mittelwerte . . .	132
6.3	Form der Kennwerteverteilung der Mittelwerte	134
6.4	Bestimmung des Konfidenzintervalls für den Stichprobenmittelwert	135
6.5	Bestimmung des Konfidenzintervalls für den Erwartungswert in der Population	137
6.6	Zusammenfassung	142
6.7	Blitzlichter aus der Forschungspraxis	143
7	Prüfung des Merkmalsunterschieds zwischen zwei Gruppen:	
	Prinzip des Signifikanztests	145
7.1	Grundprinzip der Analyse von Gruppenunterschieden	145
7.2	Der t-Test für unabhängige Stichproben	147
7.2.1	Formulierung der Hypothesen	147
7.2.2	Festlegung des Signifikanzniveaus	148
7.2.3	Voraussetzungen für die Anwendung des t-Tests	148
7.2.4	Bestimmung der statistischen Prüfverteilung unter der Annahme, dass die Nullhypothese gilt	149
7.2.5	Berechnung des t-Werts	151
7.2.6	Ergebnisse bei Analyse mittels der Statistiksoftware SPSS	152
7.3	Vergleich zweier Gruppen bei ordinalskalierten Daten: Der Wilcoxon-Rangsummen-Test bzw. Mann-Whitney-U-Test	155
7.4	Ausblick: t-Test für abhängige Stichproben, Wilcoxon-Rangsummen-Test, Varianzanalyse	158
7.5	Allgemeines Ablaufschema für Signifikanztests	160
7.6	Zusammenfassung	160
7.7	Blitzlichter aus der Forschungspraxis	161
8	Anwendung und Interpretation des Signifikanztests und das Effektstärkemaß Cohens d	165
8.1	Zwei Arten möglicher Fehlentscheidungen: Der α - (alpha-) und der β - (beta-) Fehler	165
8.2	Cohens d als Maß der Effektstärke und der Einfluss der Stichprobengröße . .	169
8.3	Zusammenfassung	172
8.4	Blitzlichter aus der Forschungspraxis	172
9	Der Zusammenhang zweier Merkmale: Die Korrelationsanalyse	175
9.1	Korrelationskoeffizienten: Maßzahlen der Stärke und Richtung von Merkmalszusammenhängen	175
9.2	Deterministische und stochastische lineare Zusammenhänge	177

9.3	Der lineare Zusammenhang intervallskalierter Merkmale: Die Kovarianz und Pearsons Produkt-Moment-Korrelationskoeffizient r	179
9.3.1	Die Kovarianz zweier intervallskalierter Merkmale	179
9.3.2	Die standardisierte Kovarianz: Pearsons Produkt-Moment-Korrelation	184
9.4	Prüfung der Signifikanz von Korrelationen	187
9.4.1	Signifikanz und Effektstärke	190
9.4.2	Voraussetzungen des Signifikanztests	192
9.5	Maße des Zusammenhangs zweier ordinalskalierter Merkmale: Die Rangkorrelationskoeffizienten Spearmans ρ (rho) und Kendalls τ (tau)	193
9.6	Berechnung der Produkt-Moment-Korrelation r , Spearmans ρ und Kendalls τ mittels SPSS	197
9.7	Überblick über wichtige Korrelationskoeffizienten	199
9.8	Zusammenfassung	199
9.9	Blitzlichter aus der Forschungspraxis	201
10	Fallstricke bei der Interpretation von Merkmalszusammenhängen: Korrelation und Kausalität	205
10.1	Die Frage nach Ursache und Wirkung: Merkmalszusammenhänge vs. Kausalität bzw. kausale Wirkmodelle	205
10.2	Das Simpson-Paradox: Konfundierungen als Ursache sich widersprechender Studienbefunde	213
10.3	Bedeutung des Unterschieds zwischen Korrelation und Kausalität für die Evidenzbasierung in den Gesundheitswissenschaften	218
10.4	Zusammenfassung	222
10.5	Blitzlichter aus der Forschungspraxis	223
11	Schätzung der Ausprägung eines Merkmals aufgrund der Ausprägung anderer Merkmale: Regressionsanalyse	227
11.1	Einfache lineare Regressionsanalyse	227
11.1.1	Bestimmung der linearen Regressionsgleichung	228
11.1.2	Varianzaufklärung, Determinationskoeffizient, Standardschätzfehler	232
11.1.3	Lineare Regressionsanalyse mit SPSS	236
11.2	Multiple lineare Regression: Schätzung der Ausprägung einer Kriteriumsvariablen auf Basis mehrerer Prädiktoren	238
11.3	Zusammenfassung	242
11.4	Blitzlichter aus der Forschungspraxis	243
12	Analyse des Zusammenhangs zweier kategorialer Merkmale mittels des χ^2-Werts und daraus abgeleiteter Maßzahlen	247
12.1	Darstellung von Häufigkeitsverteilungen in Kontingenztabellen	248

12.2	Analyse des Zusammenhangs zweier dichotomer Merkmale	249
12.2.1	Grundbegriffe aus der Mengenlehre	249
12.2.2	Beobachtete und erwartete Häufigkeiten	252
12.2.3	Prüfung der Signifikanz des Zusammenhangs zweier Merkmale: Der χ^2 -Test (Chi-Quadrat-Test)	252
12.2.4	Der ϕ -Koeffizient als Korrelationskoeffizient für zwei dichotome Variablen	256
12.2.5	Einfluss der Grundrate auf das Ergebnis des Signifikanztests und die Zusammenhangsstärke	258
12.3	Analyse des Zusammenhangs polytomer kategorialer Merkmale	259
12.4	Berechnung der Maßzahlen mittels SPSS	261
12.5	Zusammenfassung	263
12.6	Blitzlichter aus der Forschungspraxis	264
13	Analyse des Zusammenhangs von Risiko- und Schutzmerkmalen sowie Testergebnissen mit dem Gesundheitsstatus: Epidemiologische Maßzahlen . .	267
13.1	Maßzahlen epidemiologischer und klinischer Studien	268
13.1.1	Prävalenz, Inzidenz, Morbidität und Mortalität	269
13.1.2	Risiko und Odds (Ereignisquote)	271
13.1.3	Relatives Risiko (RR), Absolute Risikoveränderung (ARV) und Number-Needed-to-Treat (NNT)	274
13.1.4	Odds Ratio (OR), logarithmiertes Odds Ratio und Yules Y	277
13.1.5	Die Berechnung und Prüfung der Signifikanz von OR und $\ln(\text{OR})$ mittels SPSS	281
13.1.6	Vertiefender Ausblick: Integration von Zusammenhangsständen in Metaanalysen	283
13.2	Logistische Regression	284
13.2.1	Einfache logistische Regression aufgrund einer dichotomen Prädiktorvariable	285
13.2.2	Einfache logistische Regression aufgrund einer intervallskalierten Prädiktorvariable	287
13.2.2	Multiple logistische Regression aufgrund mehrerer Prädiktorvariablen	292
13.3	Analyse des Vorhersagewerts von Testergebnissen	294
13.3.1	Sensitivität, Spezifität, Positiv und Negativ Prädiktiver Wert	294
13.3.2	Das Bayes-Problem: Hohe Fehlerrate von positiven Testergebnissen bei seltenen Erkrankungen	298
13.4	Zusammenfassung	303
13.5	Blitzlichter aus der Forschungspraxis	305

14 Schlussbemerkungen 309

Literatur 317

Glossar 317

Abkürzungsverzeichnis 335

Autorenverzeichnis 339

Sachwortverzeichnis 341

Hinweise zu Zusatzmaterialien 347

Vorwort

Studierende der Therapie- und Gesundheitswissenschaften benötigen in ihrer zukünftigen beruflichen Tätigkeit Fähigkeiten und Fertigkeiten, die sie in die Lage versetzen, den Gesundheitszustand der betreuten und behandelten Menschen – gemeinsam mit diesen – positiv zu beeinflussen. Damit die Gesundheit von Menschen gezielt gefördert werden kann, bedarf es soliden Wissens, wodurch deren Gesundheit gekennzeichnet ist und von welchen Merkmalen oder Behandlungsmaßnahmen der Gesundheitszustand beeinflusst wird.

Gesundheitswissenschaftliche Forschung dient der Gewinnung und der kritischen Analyse dieses Wissens. In empirischen gesundheitswissenschaftlichen Forschungsstudien werden Menschen hinsichtlich Eigenschaften und Merkmalen untersucht, die für eine Forschungsfragestellung bedeutsam sind: z.B.: „Welche Erkrankungssymptome treten typischerweise wann im Falle einer bestimmten Erkrankung auf?“, „Welche Gesundheitsmerkmale stehen mit einer hohen Lebenszufriedenheit in Zusammenhang?“ oder „Hat ein verbessertes Ernährungs- und Bewegungsverhalten ein höheres Wohlbefinden oder ein verringertes Erkrankungsrisiko zur Folge?“

Bei der Beforschung und Analyse solcher Fragestellungen sind empirische Forschungsmethoden und statistische Analyseverfahren von grundlegender Bedeutung. Dies wird schnell deutlich, wenn man sich zum Stand der Forschung zu einem Thema kundig machen will. Studienbefunde werden in Fachzeitschriften und auf Fachkonferenzen so präsentiert, dass diese ohne Kenntnis der methodischen und

statistischen Hintergründe nicht angemessen verstanden und eingeordnet werden können: Die Studieninformationen und -ergebnisse werden mittels statistischer Koeffizienten, wie Mittelwerten, Standardabweichungen oder Korrelationskoeffizienten, kommuniziert. Ob eine Maßnahme den Menschen hilft, wird daraus abgeleitet, ob ein ominöser sogenannter p-Wert den Wert ,05 unterschreitet. Und nicht nur die Zahlen, sondern auch die Begriffe wirken zum Teil kryptisch, wenn man sich nicht fundiert mit den forschungsmethodischen und statistischen Hintergründen auseinandergesetzt hat.

Gesundheitswissenschaftlerinnen und Gesundheitswissenschaftler sollten in der Lage sein, solche statistischen Informationen und Argumentationen richtig zu verstehen, kritisch zu beurteilen und daraus angemessene Schlussfolgerungen zu ziehen. Dieses Lehrbuch verfolgt das Ziel, Sie zu befähigen, diese Informationen zu verstehen und entsprechende Methoden einzuordnen sowie auch eigenständig umzusetzen.

Jedoch gehören statistische Methoden bei Studierenden erfahrungsgemäß in der Regel nicht zu den Inhalten des Studiums, denen sie mit dem größten Interesse und mit dem größten Wissensdurst begegnen. Zum einen, weil die Inhalte vielleicht nicht den Kern dessen betreffen, was sie zur Wahl eines gesundheitswissenschaftlichen Studiums bewogen hat. Mit Menschen arbeiten zu wollen und deren Gesundheit und Wohlbefinden zu fördern, hat – zumindest auf den ersten und wahrscheinlich auch auf den zweiten Blick – nichts mit der Anwendung von Formeln und Rechenvorschriften

gemeinsam. Zum anderen löst „Mathematik“ und „Statistik“ bei vielen nicht unmittelbar positive Emotionen aus. Mathematikangst oder ein aversionsbegleiteter Respekt vor Mathematik stellen eine nicht seltene Hürde dar, um das Studium mit unvoreingenommener Neugierde, Freude und Offenheit zu beginnen.

Aber spätestens bei der genaueren Beschäftigung mit der Inhaltsstruktur des Studiums wird deutlich, dass Statistik und eher technisch anmutende Forschungsmethoden einen Kernbestandteil ausmachen. Die bittere Einsicht mag dann unvermeidbar scheinen: „Es nützt alles nichts, da muss man durch!“

Als Autor dieses Lehrbuchs ist es mir ein dringendes Anliegen, eine positivere Sichtweise und Haltung bei Ihnen, den Leserinnen und Lesern, zu unterstützen. Ein wesentliches Ziel dieses Lehrbuchs besteht darin, dass Sie verstehen, *weshalb* den Forschungsmethoden und der Statistik eine so hohe Bedeutung in der Ausbildung und in der Beschäftigung mit gesundheitswissenschaftlicher Forschung beigemessen wird. Verstehen Lernende den Sinn dessen, so lernen sie tendenziell unbefangener und interessierter Statistik – und damit auch wesentlich erfolgreicher. Eine pragmatische Einstellung, die die Sinnhaftigkeit unterstützt, besteht darin, dass Statistik und Forschungsmethoden unerlässliche Werkzeuge von Gesundheitswissenschaftlerinnen und Gesundheitswissenschaftlern darstellen. In jeder Profession gibt es Wissen, Fähigkeiten und Fertigkeiten, über die man verfügen und die man beherrschen muss, um den Beruf professionell betreiben zu können – unabhängig davon, in welchem Maße diese an sich als spannend oder interessant empfunden werden. Also, auch wenn Sie Statistik und Forschungsmethoden als „notwendiges Übel“ betrachten würden, wäre bereits eine Basis geschaffen, sich dem Themengebiet in interessierter Weise zu nähern: Nämlich dann, wenn Sie die „Notwendigkeit“ erkennen und für sich ernst nehmen, weil diese Kernmerkmale gesundheitswissenschaftlicher Professionalität darstellen.

Erstrebenswerter ist es jedoch, dass Sie für sich das Bewusstsein entwickeln, dass Statistik und Forschungsmethoden Ihr Wissen, Ihre Kompetenzen sowie Ihre Entscheidungsfähigkeit positiv bereichern. Statistik und Forschungsmethoden besitzen nämlich deswegen einen so hohen Stellenwert, weil sie es in gesundheitswissenschaftlichen Forschungs- und Handlungsfeldern ermöglichen, Sachverhalte aussagekräftiger, transparenter, fehlerfreier und zuverlässiger zu beurteilen. Nur wer die entsprechenden Informationen versteht und kritisch einordnen kann, kann Forschungserkenntnisse adäquat nutzen. Möchte man beispielsweise beurteilen, ob und in welchem Maße eine Fördermaßnahme die Gesundheit unterstützt, so muss man sich fragen: „Woran kann ich erkennen, dass die Wirksamkeit als nachgewiesen gelten kann?“, „Woran kann ich erkennen, welche gesundheitlichen Merkmale in stärkerem oder schwächerem Maße gefördert werden können?“. Wie wir in diesem Lehrbuch sehen werden, bedarf es mehrerer begründeter Entscheidungen zur Studienanlage („Welche Personen müssen wann wie untersucht werden?“) und der statistischen Auswertung, damit belastbare Aussagen resultieren. Denkt man darüber nach, wie sich eine gesundheitswissenschaftliche Fragestellung beantworten lässt, so bedeutet dies unmittelbar, dass man sich Gedanken darüber machen muss, wie man das Problem strukturiert untersuchen (Domäne der Forschungsmethoden) und wie man die Daten statistisch auswerten sollte (Domäne der Statistik), damit eine möglichst klare Antwort erwartet werden kann.

Die öffentlichen Informationen und Diskussionen zur COVID-19-Pandemie haben besonders nachdrücklich gezeigt, wie eng verwoben individuelle und öffentliche gesundheitsbezogene Entscheidungen mit statistischen Informationen und statistisch sorgfältig ausgewerteten Studienergebnissen sind: Maßzahlen wie die Inzidenz (Neuerkrankungsrate) oder der R-Wert (als Maß der Infektionsdynamik) waren entscheidend, um die Gefährdung des Einzel-

nen und der Gesellschaft beurteilen zu können. Wie sich Bürgerinnen und Bürger verhalten sollten, um ihre Gesundheit zu schützen, wurde von diesen Kennziffern oder von statistischen Vorhersagemodellen abhängig gemacht. Ob eine Impfung gegen das Virus effektiv schützt, wurde aus dem statistischen Vergleich kontrolliert ermittelter Prozentzahlen abgeleitet. Aber auch Falsch-Nachrichten („Fake news“) wurden durch Zahlenangaben und Studien begründet. Als Gesundheitswissenschaftlerinnen und Gesundheitswissenschaftler müssen Sie mit diesen Informationen und deren Hintergründen kompetent umgehen können.

Aber auch für die direkte Kommunikation mit Menschen sind forschungsmethodische und statistische Kenntnisse wichtig. Dies wird insbesondere durch das moderne Konzept der gemeinsamen oder partizipativen Entscheidungsfindung nachvollziehbar. Weder die beratende bzw. betreuende noch die beratene bzw. betreute Person ist für gesundheitsbezogene Entscheidungen allein verantwortlich. Vielmehr ist es die Aufgabe beider, sich kritisch über die möglichen Alternativen zu informieren und sich vor dem Hintergrund der individuellen Lebenssituation und Bedürfnisse für die im Einzelfall passendste Alternative zu entscheiden. Ihre Aufgabe wird dabei insbesondere darin bestehen, die beratenen und betreuten Menschen, die Hilfe und Unterstützung in Gesundheitsfragen suchen, kompetent über den Stand der Forschung zu informieren. Dies setzt voraus, dass Sie aktuelle Erkenntnisse identifizieren, verstehen, kritisch beurteilen und anwendungsbezogen reflektieren können. Dies stellt die Basis dar, um Menschen angemessen informieren und sich mit Menschen professionell austauschen zu können.

Gesundheitswissenschaftliches Denken und gesundheitswissenschaftliche Erkenntnisgewinnung gehen also Hand in Hand mit forschungsmethodisch fundiertem Handeln und statistischer Datenanalyse. Je vernetzter und integrierter Sie sich parallel mit der inhaltlichen Ebene und der forschungsmethodisch-statisti-

schen Ebene beschäftigen, desto besser können Sie Ihre eigenen Kompetenzen als Gesundheitswissenschaftlerin und Gesundheitswissenschaftler entwickeln und in der Praxis zielgerichtet und lösungsorientiert anwenden.

Um Ihnen hierzu einen hilfreichen Einstieg zu ermöglichen können, werden in diesem Lehrbuch, verschiedene **didaktische Leitgedanken** und **Methoden** berücksichtigt:

- *Beispielbasierte Einführung der Inhalte:* Alle statistischen Inhalte werden in Zusammenhang mit einem Anwendungsbeispiel aus der Gesundheitswissenschaft eingeführt. Die kombinierte Darstellung unterstützt Sie dabei, die statistischen Informationen verknüpft mit anschaulichen Anwendungsfragen zu verstehen.
- *Anwendungsbeispiele:* Zusätzlich werden vertiefende gesundheitswissenschaftliche Beispiele aus der angewandten Forschung dargestellt, die das Verständnis für die gesundheitswissenschaftliche Bedeutung der Verfahren fördern.
- *Blitzlichter aus der Forschungspraxis:* Als letzter Unterabschnitt werden in jedem Kapitel drei Blitzlichter aus der Forschungspraxis samt Quellenangabe berichtet. Hierbei handelt es sich in der Regel um eindrucksvolle Anwendungsfälle aus der Literatur, komplexere Studien oder spezialisierte Vertiefungstexte, in denen die Kapitelinhalte wieder auftauchen. Hierdurch wird die Verbindung der Kapitelinhalte zu anspruchsvolleren wissenschaftlichen Anwendungen, Diskursen und Textformen hergestellt.
- *Einbindung der Informationen in Forschungsprozesse:* Statistische Inhalte und Analyseschritte werden als Teil von Forschungsprozessen eingeführt, sodass die Bedeutung in Bezug auf gesundheitswissenschaftliche Inhalte („Wieso bzw. mit welchem Ziel wird diese Analyse so ausgeführt?“) im Blick behalten werden kann. Insbesondere das einführende Kapitel 1 widmet sich vorwiegend der Rahmung statistischer Verfahren in vollständigen

gesundheitswissenschaftlichen Forschungsprozessen – von der theoretischen Fundierung bis zur Implementierung in der Anwendungspraxis.

- *So wenig Mathematik wie erforderlich, aber nicht weniger:* Ihnen stehen Programme und Internettools zur Verfügung, die nahezu jede konkrete Rechenarbeit und statistische Analyse für Sie erledigen werden. Es wäre also auch möglich, die Grundlagen der Statistik so einzuführen, dass Sie nur noch entscheiden müssen, welche Zahlen Sie wo eintragen müssen und welcher Buttonclick Ihnen dann die fertigen Ergebnisse liefert. Für das wirkliche Verständnis der Verfahren und deren flexible Anwendung ist es aber wichtig, einmal an exemplarischen Fällen genau nachzuvollziehen, wie die statistische Berechnung erfolgt: Woher soll man wissen, was gerechnet werden muss, wenn man nicht weiß, was da eigentlich wie verrechnet wird? Zudem werden Sie auch im Studium und bei der Beschäftigung mit Literatur mit abgewandelten Formen der im Buch behandelten statistischen Inhalte und darauf aufbauenden weiterführenden Verfahren konfrontiert werden. Damit Sie ein solides Fundament für das „Denken in Bezug auf statistische Methoden“ entwickeln, ist es wichtig, die elementaren Rechenprinzipien durchdacht und nachvollzogen zu haben.

Die in diesem Buch eingeführten statistischen Inhalte sind prinzipiell aufgrund elementarer mathematischer Kenntnisse nachvollziehbar: Über die Bildung von Summen, Differenzen, Produkte und Brüchen sowie die Quadrierung von Werten gehen die Inhalte nur selten hinaus. Die Exponentialfunktion und der Logarithmus lassen sich leider nicht ganz vermeiden. Auch wenn der Autor dies aufgrund seines Interesses und der längerfristigen Beschäftigung mit statistischen Inhalten wahrscheinlich nicht mehr unvoreingenommen beurteilen kann, so werden zwar detaillierte, aber keine komplexen mathematischen Betrachtungen notwendig

sein. Bei der Darstellung wurde darauf geachtet, dass alle Schritte logisch und auf Basis konkreter Beispiele, Zahlen und Erläuterungen eingeführt werden. Es würde mich sehr freuen, wenn dadurch bei Ihnen das sehr hilfreiche Bewusstsein unterstützt werden kann, dass die Basisstatistik kein mathematisches Hexenwerk darstellt und diese prinzipiell auch inhaltlich-logisch sinnvollen Überlegungen folgt.

- *Kurzeinführung in die computergestützte Datenverarbeitung:* Für wichtige Testverfahren wird gezeigt, wie diese in der Analysepraxis mittels des Statistikprogramms SPSS oder online frei verfügbarer Internettools angewendet werden. So können Sie unmittelbar nachvollziehen, wie die von Ihnen selbst berechneten Ergebnisse bei der Berechnung per Computer angezeigt werden. Dies erleichtert den Transfer der Inhalte in die praktische Anwendung bei der Durchführung eigener statistischer Analysen im Studium und in der eigenständigen Forschungspraxis.
- *Übungsaufgaben:* In den einzelnen Kapiteln werden die statistischen Verfahren stets erläutert und begründet. Am Ende des Kapitels finden Sie dann Aufgaben, die Sie selbstständig allein oder in einer Lerngruppe bearbeiten sollten. Inhalte zunächst angeleitet oder kommentiert durchzuarbeiten, und sie dann möglichst unmittelbar selbstständig an Aufgaben anzuwenden, die einen gewissen Transfer erfordern, sichert den Lernerfolg nachhaltig. Ihre Lösungen können Sie dann mit den Musterlösungen abgleichen, die als Zusatzmaterialien in der Mediathek beigelegt sind. Sie können dort unter dem Link hgf.io/download und dem Code B-158199 abgerufen werden.
- *Aufgliederung in gut abgrenzbare, strukturierte Informationseinheiten:* Die einzelnen Kapitel wurden möglichst kurz gehalten, sodass sie sich mit gut abgrenzbaren Inhalten kompakt beschäftigen können. Die wesentlichen Inhalte und Ziele werden zu Beginn jedes Kapi-

tels als Ausblick (engl. Advance Organizer) explizit vorangestellt. Am Ende jedes Kapitels findet sich eine strukturierte Zusammenfassung der zentralen Kapitelinhalte.

Ich wünsche Ihnen, liebe Leserinnen und Leser, eine erkenntnisreiche Beschäftigung mit den Inhalten des Lehrbuchs. Es würde mich sehr freuen, wenn sich diese – verknüpft mit den übrigen Inhalten Ihres Studiums oder Ihrer gesundheitsberuflichen Tätigkeit – als hilfreich und nützlich erweisen.

Freiburg, im Juni 2023
Markus Antonius Wirtz

Anmerkung des Lektorats

Das Lektorat dankt der Hamburger Fernhochschule für die Übertragung der Nutzungsrechte an den Studienbriefen GLS 01-1705-001-1, GLS 01-1705-002-1, GLS 01-1705-003-1, GLS 01-1705-004-1, GLS 01-1705-005-1.

Danksagung

Die inhaltliche Aufbereitung dieses Lehrbuchs basiert auf einer Reihe von Lehrmaterialien, die ich gemeinsam mit Frau Dr. Anja Alexandra Schulz als Studienbriefe für den Studiengang Therapie- und Pflegewissenschaften für die Fernhochschule Hamburg verfasst habe. Anja Schulz hat auch die Darstellung der Inhalte in diesem Buch durch kritische Rückmeldungen unterstützt. Die Aufbereitung der Inhalte dieses Lehrbuchs orientiert sich an dem an der Fernhochschule Hamburg etablierten Konzept für die didaktische Gestaltung von Studienbriefen. Sehr herzlich möchte ich mich deswegen bei

Frau Prof. Dr. Cathleen Gaede-Illig und Stephan Müthel für die unkomplizierte und angenehme Kommunikation bei Erstellung der Studienbriefe bedanken. Zudem sei Frau Bettina Scharpenberg und Frau Jutta Boye vom Verlag Handwerk und Technik gedankt, die die Erstellung eines Textes für das Oberstufenlehrbuch „Mensch im Fokus I – Gesundheitswissenschaften und Gesundheitsförderung“, an dem sich einzelne Inhalte des vorliegenden Lehrbuchs orientieren, hilfreich begleitet haben. Wesentliche Teile des Textes wurden durch Stephan Müthel und Katharina Wirtz sehr sorgfältig lektoriert.

1 Statistik und Forschungsmethoden in der gesundheitswissenschaftlichen Forschung

Statistische Methoden sind von elementarer Bedeutung für alle wissenschaftlichen Disziplinen, wenn Phänomene systematisch beschrieben und Fragestellungen mittels empirischer Sachverhalte untersucht werden. In den Therapie- und Gesundheitswissenschaften interessieren wir uns beispielsweise für Merkmale des Gesundheitssystems, Behandlermerkmale oder Eigenschaften von Patientinnen und Patienten. Es sollen Informationen gewonnen werden, unter welchen Bedingungen (z. B. Ernährungsstil, Stressbelastung) Erkrankungen (wahrscheinlicher) auftreten, oder welche Behandlungsmethoden unter welchen Bedingungen dazu führen, dass sich der Gesundheitszustand der Behandelten günstig entwickelt bzw. Behandlungsziele erreicht werden. Statistische oder quantitative Methoden sind ein wichtiges Werkzeug, um solche Forschungsinhalte in strukturierter Weise und auf Basis begründeter Wissenschaftsstandards untersuchen zu können.

Um die Prinzipien des angewandten Arbeitens mit statistischen Verfahren in den Therapie- und Gesundheitswissenschaften verstehen zu können, ist es nicht nur wichtig, sich damit auseinanderzusetzen, wie Statistik funktioniert oder Statistiken berechnet werden können. Vielmehr bedarf es eines vertieften Verständnisses, weshalb die quantitativen Methoden einen solch hohen Stellenwert bei der Gewinnung verlässlicher wissenschaftlicher Erkenntnisse haben. Quantitative Methoden müssen begründet in komplexere Forschungsprozesse eingebunden sein, damit sich ein optimaler Erkenntnisnutzen für die therapie- und gesund-

heitswissenschaftliche Forschung und Praxis ergeben kann.

In diesem Kapitel werden einführend grundlegende *Merkmale wissenschaftlichen Arbeitens* geklärt. Dies wird an den Anfang gestellt, da die Anwendung und Aussagekraft statistischer Methoden wesentlich davon abhängen, wie ein mehr oder weniger komplexer Forschungsansatz begründet werden kann. Forschung muss immer so erfolgen, dass das gewählte Vorgehen bestmöglich geeignet ist, die Forschungsfrage angemessen zu beantworten. Die Bearbeitung von Forschungsfragen erfolgt dabei stets in mehreren Schritten. Diese Schritte bilden den sogenannten Forschungsprozess.

Nach Durcharbeiten dieses Kapitels

- kennen Sie die Grundmerkmale gesundheitswissenschaftlicher Forschung.
- verstehen Sie, wie in der wissenschaftlichen Forschung die Zuverlässigkeit und Aussagekraft von Befunden sichergestellt wird.
- können Sie die Grenzen wissenschaftlicher Erkenntnisse einschätzen.
- wissen Sie, wie statistische Methoden im Rahmen empirischer Erkenntnisprozesse eingebunden werden.
- kennen Sie typische Fragestellungen der gesundheitswissenschaftlichen Forschung.
- kennen Sie die Parallelen statistischer Forschungsansätze und inhaltlicher Forschungsziele.
- kennen Sie die Parallelen von empirischen und theoretischen Forschungselementen.

1.1 Angewandte wissenschaftliche empirische Forschung

Wissenschaftliche Erkenntnisse gelten im Vergleich zu alltäglichen Erfahrungen als besonders aussagekräftig, verlässlich und nützlich. Auf persönlichen Erfahrungen gründende Meinungen und Überzeugungen gelten in der Regel als individuell, d. h. durch die spezifische Sichtweise der jeweiligen Person geprägt. Zudem sind geäußerte Meinungen von der jeweiligen Kommunikationssituation abhängig und verändern sich je nach Verlauf der Kommunikation. Wird hingegen eine Erkenntnis als „wissenschaftlich“ oder „wissenschaftlich geprüft“ bezeichnet, so kann man sich in stärkerem Maße darauf verlassen, dass diese Erkenntnis auch wirklich zutrifft.

Aber was bedeutet „wissenschaftlich“ tatsächlich? Sind wissenschaftliche Erkenntnisse „wahr“? Sind Aussagen, die nicht mit einer wissenschaftlich geprüften Aussage im Einklang stehen, notwendigerweise „unwahr“ oder „falsch“? Ist etwas nur dann „wissenschaftlich“, wenn eine Wissenschaftlerin oder ein Wissenschaftler, die bzw. der eine entsprechende Ausbildung durchlaufen hat, diese Erkenntnis gewonnen hat?

Dass der Begriff der „Wahrheit“ kein guter Ausgangspunkt zur Charakterisierung von „Wissenschaftlichkeit“ ist, wird bereits dadurch kenntlich, dass sich nicht alle Wissenschaftlerinnen und Wissenschaftler einig sind. Ganz im Gegenteil bestehen auch in der Wissenschaft unterschiedliche Ansätze zur Erklärung zu demselben Problem. So existieren Befunde und Theorien, dass beispielsweise psychische Störungen eher genetisch, eher neurophysiologisch, eher durch individuelle und soziale Lernerfahrungen oder durch die systemische Lebensumwelt bedingt sein können. Verschiedene Wissenschaftlerinnen bzw. Wissenschaftler können den möglichen Ursachen eine unterschiedliche Bedeutung zuschreiben und ihre

eigenen Forschungsaktivitäten daran ausrichten. Die Erkenntnisse, die die Forschenden gewinnen, sind somit in gewissem Maße auch davon abhängig, welche Forschungsinhalte sie als besonders wichtig oder „beforschenswert“ halten und wie sie die Befunde argumentativ einordnen.

Zudem verändern sich wissenschaftliche Erkenntnisse. Bestehende Erkenntnisse werden erweitert oder es wird deutlich, dass diese – anders als ursprünglich angenommen – nur unter bestimmten Bedingungen gültig sind. Auch als gesichert angenommene Erkenntnisse können durch neue Erfahrungen oder neue wissenschaftliche Untersuchungsmethoden widerlegt werden. Wissenschaftliche Erkenntnisse sind also nicht für immer gültig: Vielmehr ist gerade das kritische Hinterfragen etablierten Wissens ein wichtiger Ansatzpunkt für wissenschaftlichen Fortschritt.

BOX – Merke

Wissenschaftliche Erkenntnisse gelten als besonders aussagekräftig, verlässlich und nützlich. Trotzdem sind diese nicht als „wahr“, sondern im Vergleich zu nicht wissenschaftlichen Erkenntnissen als besonders gründlich und begründet untersucht sowie kritisch geprüft zu betrachten.

Um die Kernmerkmale wissenschaftlichen Arbeitens und wissenschaftlicher Erkenntnisse verstehen und einordnen zu können, ist es hilfreich, sich mit dem Unterschied zu **Alltagsüberzeugungen** und **Alltagserkenntnissen** zu beschäftigen. Wissenschaftliche Erkenntnisgewinnung unterscheidet sich wesentlich von der Gewinnung von Alltagserkenntnissen. Um die Bedeutung der Statistik für die Gesundheitswissenschaften verstehen zu können, sollen Alltagswissen und wissenschaftliches Wissen zunächst kontrastiert werden, bevor der Fokus auf die Statistik an sich gelegt wird.

1.1.1 Alltagstheorien und -überzeugungen

Alltägliches menschliches Wahrnehmen und Denken verfolgt ähnliche Ziele wie wissenschaftliches Arbeiten. Wir müssen im Alltag erkennen, welche Informationen in unserer Umwelt wichtig sind und müssen deren Bedeutung verstehen. Wir bilden uns eine Vorstellung von der Welt, durch die für uns verständlich wird, welche Bedingungen oder Ursachen welche Konsequenzen zur Folge haben. Auf Basis dieses Weltverständnisses entwickeln wir Handlungsstrategien, um uns in der Welt zurechtzufinden und diese im Sinne von uns angestrebten Zielen zu gestalten bzw. zu beeinflussen.

So hat jeder Mensch zu wesentlichen Fragen des alltäglichen Lebens bestimmte Überzeugungen oder Meinungen. Fragt man Menschen zum Beispiel, was für ihn in Bezug auf seine Gesundheit wichtig ist, so wird jeder hierzu ein Gefüge von Meinungen und Überzeugungen äußern können. Menschen können für sich persönlich in der Regel äußern und auch begründen, in welchem Maße körperliche Bewegung im Alltag als wichtig erachtet wird, was eine gesunde Ernährung kennzeichnet, welche Bedeutung psychischen Belastungen (z. B. Stress) beigemessen wird oder in welchem Maße die Kontrolle von Gesundheitsparametern (z. B. Gewicht, Inanspruchnahme von Vorsorgeuntersuchungen) als

bedeutsam erachtet wird. Menschen haben darüber hinaus implizite Annahmen oder Überzeugungen, ob Gesundheitsvorsorge überhaupt für sie persönlich bedeutsam ist (Bedrohungserleben) und in welchem Maße sich die Gesundheitsentwicklung überhaupt beeinflussen lässt (Kontrollüberzeugungen; Schüz, 2018).

Im Kopf jedes Menschen existiert also ein mehr oder weniger starkes „Bild“ oder Informationsmodell von einem Gegenstandsbereich. Dies hilft dem Menschen zu verstehen, weshalb die Dinge so sind wie sie sind, und wie er sich selbst verhalten kann, damit er sein Leben gemäß der eigenen Bedürfnisse und Ziele gestalten kann. Ein solches Informationsmodell kann man als Alltagstheorie verstehen. Denn eine solchen Theorie wird durch Erfahrungen im natürlichen alltäglichen Leben geprägt und erzeugt.

BOX – Definition

Alltagstheorien bzw. **Alltagsüberzeugungen** sind Auffassungen, die sich Menschen über ihre Lebensumwelt aufgrund ihrer Erfahrung in ihrer natürlichen Lebensumwelt gebildet haben. Sie „haben die wichtige Funktion, sowohl das eigene Leben als auch das Verhalten anderer Menschen verstehbar, subjektiv voraussagbar und scheinbar kontrollierbar zu machen.“ (Fahrenberg, 2022, S. 136)

BOX – Merke

Überzeugungen und **Meinungen** kennzeichnen die Sichtweise und Bewertungen des einzelnen Menschen in Bezug auf ein bestimmtes Objekt oder einen bestimmten Sachverhalt. Sie sind abhängig von individuellen Erfahrungen und subjektiv geprägt. Zu demselben Sachverhalt unterscheiden sich Überzeugungen und Meinungen zwischen Menschen (interindividuell) systematisch. Sie verändern sich aber auch ggf. für jede Person (intraindividuell) je nach Situation und Erfahrungen.

Je genauer und zutreffender eine Alltagstheorie die tatsächlichen Sachverhalte in der Welt abbildet, desto besser gelingt es dem Menschen, die Welt zu verstehen und sich in ihr zurechtzufinden. Alltagstheorien bilden die Sachverhalte in der Welt jedoch in der Regel nur annähernd und gegebenenfalls nicht angemessen ab. Angenommen, eine Person äußert die Überzeugung, dass Rauchen nicht schädlich sei. Dies begründet sie damit, dass ein ehemaliger Bundeskanzler trotz langzeitigem und intensivem Tabakkonsum ein beneidenswert hohes Lebensalter bei gleichzeitig außergewöhnlich hoher kognitiver Leistungsfähigkeit erreicht habe. Obwohl diese Per-

son aufgrund eines wahren Sachverhaltes ihre Überzeugung begründen kann, kommt sie zu einer falschen Schlussfolgerung.

In der Statistik lernen wir, dass eine allgemeine Gesetzmäßigkeit (z.B. Rauchen erhöht die Wahrscheinlichkeit für Gesundheitsschäden; Rauchen hat eine verringerte Lebenserwartung zur Folge) durch Einzelfälle nicht widerlegt werden kann. In der Statistik beschäftigen wir uns wesentlich mit der Bedeutung von *Zufallseffekten*, die in Einzelfällen zu Beobachtungen führen können, die für sich genommen einer allgemeinen Gesetzmäßigkeit zu widersprechen scheinen. Erst dadurch, dass wir Techniken und Methoden entwickeln, solche Zufallseffekte abschätzen zu können und zu kontrollieren, kommen wir zu belastbaren und stabilen Aussagen.

BOX – Definition

Allgemeine Gesetzmäßigkeit: Regel oder Prinzip, die bzw. das empirischen Sachverhalten unterliegt (z.B. Frauen haben eine höhere Lebenserwartung als Männer; gesunde Ernährungsweise geht mit einem besseren Gesundheitsstatus einher).

Zufallseffekt: Unsystematische Abweichungen, die nicht erklärt oder vorhergesagt werden können (z.B. gleichaltrige Menschen gleichen Geschlechts mit identischem Ernährungsstil unterscheiden sich dennoch unsystematisch bzw. zufallsbedingt in ihrem Gesundheitsstatus).

1.1.2 Wesentliche Merkmale von Alltagstheorien und -überzeugungen

Dass Alltagstheorien und -überzeugungen – obwohl sie auf den Erfahrungen von Menschen gründen – eingeschränkt belastbar und zudem fehleranfällig sind, wird durch folgende Aspekte deutlich:

- *Jeder Mensch hat eine eigene, subjektiv geprägte Theorie (mangelnde Intersubjektivität):*

Die Sichtweise jedes Menschen wird insbesondere durch die Erziehung, durch die soziale Umwelt, durch individuelle Erfahrungen und die eigene Persönlichkeit geprägt. Diskutieren zwei beliebige Menschen über einen Sachverhalt (z.B. „Was kennzeichnet ein gesundes Leben?“), so werden immer auch Unterschiede in den Sichtweisen kenntlich. Wenn aber Menschen die Welt für sich unterschiedlich verstehen oder verstehbar machen, so wird deutlich, dass das *individuelle Bild der Realität* in der Regel nicht deckungsgleich mit der Realität an sich sein kann (**Abbildung 1-1**).

- *Dieselbe Person äußert in verschiedenen Situationen oder in unterschiedlichen Lebensphasen andere Überzeugungen (Mangelnde Zuverlässigkeit oder Konstanz):* Obwohl Menschen in der Regel zu einem bestimmten Zeitpunkt fest von der Richtigkeit ihrer Weltansicht überzeugt sind, verändert sich diese, wenn neue Erfahrungen gesammelt werden, sich die Lebenssituation verändert oder auch, wenn in einer Diskussionssituation überzeugende Argumente vorgebracht werden, die der eigenen Weltansicht widersprechen. Obwohl auch wissenschaftliche Theorien sich verändern, sind Alltagstheorien oft durch eine besonders hohe situative Instabilität gekennzeichnet.
- *Erfahrungen werden unsystematisch gesammelt und ausgewertet (Mangelnde Transparenz des Erkenntnisprozesses):* Die Sichtweise von Menschen beruht auf Erfahrungen aus unterschiedlichen Quellen. Was ein Mensch erfährt, hängt von seinen spezifischen Lebensumständen (z.B. familiäre und soziale Umwelt) ab. Fordert man Menschen auf, zu begründen, weshalb sie einer bestimmten Überzeugung sind, so werden häufig typische oder anekdotische Erlebnisse berichtet, die sich im Gedächtnis besonders plastisch eingeprägt haben. Besonders emotional erlebte Situationen wird dabei ein hohes Gewicht verliehen. In der Psychologie charakterisiert man das Erzählen oder Erinnern

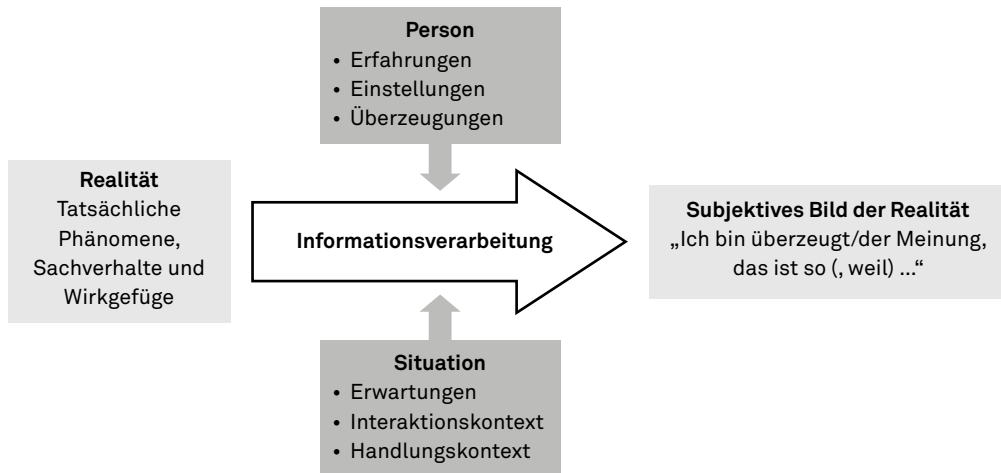


Abbildung 1-1: Aufgrund von individuellen Merkmalen der Person und Merkmalen der Situation, in der eine Person über einen Sachverhalt nachdenkt, verändert sich das (geäußerte) subjektive Bild der Realität.

von Erlebnissen oft als *rekonstruktiven Prozess*: Menschen erinnern sich dabei selektiv nur an bestimmte Aspekte. Und die Art und Weise, wie die Dinge erinnert werden, wird durch den Sinn und die Bedeutung, die man dem Erlebten im Nachhinein beimisst, gefärbt oder verzerrt.

Menschliches Erinnern und Bewerten wird durch viele kognitive Verzerrungen beeinflusst. So werden beispielsweise Risiken für Gefahren (z. B. Verkehrsunfälle, Flugzeugabstürze, Terroranschläge) systematisch höher bewertet, wenn diese in den Medien häufiger berichtet werden (Verfügbarkeitsheuristik) oder mit besonders negativen Emotionen (z. B. Ängsten) verbunden sind (Kahneman, 2012).

- *Alltagstheorien haben häufig eine hohe Wertigkeit für Personen (Mangelnde Neutralität; Identitätsfunktion)*: Sind Menschen von bestimmten Dingen fest überzeugt und sind diese Überzeugungen für sie persönlich von besonderer Bedeutung (z. B. „Menschen einer anderen religiösen Überzeugung sind weniger vertrauenswürdig“; „Schutzimpfungen haben schädliche Wirkungen“), so werden Gegenargumente abgewertet und abge-

lehnt. Menschen immunisieren sich gegen Argumente, die ihre eigene Überzeugung infrage stellen, in dem sie aktiv deren Gültigkeit anzweifeln. Informationen, die in das eigene Weltbild passen, werden tendenziell unhinterfragt akzeptiert. Informationen, die dem eigenen Weltbild widersprechen, werden ausgeblendet oder besonders kritisch hinterfragt.

- *Begriffe sind oft unpräzise und unterschiedlich definiert (Mangelnde Klärung von Bedeutungen)*: In der Alltagssprache meinen Menschen manchmal unterschiedliche Dinge, obwohl sie denselben Begriff verwenden. Ob ein Mensch mit seinem Leben zufrieden ist, wird von manchen primär am materiellen Besitz, von anderen hingegen an der psychischen oder aber körperlichen Gesundheit festgemacht. Andere sehen die familiäre, soziale oder berufliche Situation als maßgebende Merkmale, die die Lebenszufriedenheit von Menschen determiniert. Spricht man von der Gesundheit eines Menschen, so verstehen manche Menschen darunter ausschließlich die körperliche Gesundheit. Andere berücksichtigen psychische und soziale Aspekte als wesentliche Elemente des Ge-

sundheitszustands. Je abstrakter Begriffe sind, desto größer sind im Allgemeinen die Unterschiede in den Bedeutungen, die verschiedene Menschen mit diesen verknüpfen.

BOX – Merke

Merkmale von Alltagstheorien und -überzeugungen

Mangelnde Intersubjektivität: Menschen beurteilen denselben Sachverhalt verschieden.

Mangelnde Zuverlässigkeit: Subjektive Urteile verändern sich ggf. von Situation zu Situation.

Mangelnde Transparenz: Die Bedingungen für die Bildung und Veränderungen subjektiver Urteile bleibt in der Regel unklar.

Mangelnde Neutralität: Menschen identifizieren sich mit ihren bzw. durch ihre Meinungen und Überzeugungen (Identitätsfunktion).

Mangelnde Klarheit von Begriffen: Die genaue Bedeutung von Begriffen ist oft nicht hinreichend definiert.

1.1.3 Kriterien verlässlichen Wissens in den empirischen Wissenschaften

Wissenschaftliche Forschung versucht die im vorangegangenen Abschnitt dargestellten Einschränkungen der Verlässlichkeit von Alltagswissen und Alltagstheorien zu vermeiden oder bestmöglich zu minimieren. Dabei stehen folgende Merkmale wissenschaftlichen Arbeitens im Mittelpunkt:

- **Transparenz des Erkenntnisprozesses:** Forschende legen offen, was sie zu ihrem Forschungsvorhaben angeregt hat, weshalb sie sich für eine bestimmte Untersuchungsmethode entschieden haben und wie die gewonnenen Befunde und die daraus gezogenen Schlussfolgerungen begründet werden. Für wissenschaftliche Befunde genügt es nicht, dass Wissenschaftlerinnen oder Wissenschaftler die Überzeugung gewonnen ha-

ben, dass bestimmte Dinge plausibel erscheinen. Vielmehr ist es notwendig, dass sie unabhängigen Personen alle Informationen zugänglich machen, die ein angemessenes Verständnis des tatsächlichen Untersuchungsgeschehens ermöglichen. Außenstehenden wird Einblick in alle Details der Studie gewährt, sodass diese sich einen eigenen kritischen Eindruck der Befundlage bilden können.

- **Systematische Dokumentation des Vorgehens:** Forschende gewähren Einblick in die verwendeten Untersuchungsmaterialien und dokumentieren/begründen ihre Entscheidungen im Verlauf des Forschungsprozesses. Die Forschungsdaten müssen gesichert und gegebenenfalls anderen Forschenden zur kritischen unabhängigen Analyse zur Verfügung gestellt werden. In einem Forschungsbericht wird der gesamte Prozess vollständig dargelegt.
- **Einhaltung von vereinbarten Forschungsstandards:** Sind die Methoden geeignet, um eine Fragestellung aussagekräftig zu beantworten? In der Wissenschaft existieren Standards, die eingehalten werden müssen, damit die angestrebte Erkenntnis bestmöglich erreicht werden kann. Welche Methode zur Beantwortung der Forschungsfragestellung eingesetzt wird, wird wesentlich davon bestimmt, welche Methoden in Lehrbüchern oder Forschungsleitlinien als angemessen empfohlen werden. Möchte man zum Beispiel untersuchen, ob eine Behandlungsmethode ursächlich zur Verbesserung des Patientenzustandes führt, so sind die Forschenden dazu angehalten, ein bestimmtes Design zu verwenden – nämlich ein Experiment oder eine randomisierte kontrollierte Studie (Wirtz & Schulz, 2020). Wissenschaftliche Forschungsstandards legen fest, welche Schlussfolgerungen aus Daten logisch zulässig sind und welche nicht. Hierdurch soll jede Beliebigkeit, die durch die individuelle Person des oder der Forschenden entstehen könnte, vermieden werden.

- **Präzision der Terminologie** (*Definition von Begriffen*): In Forschungsstudien muss jedes Merkmal, das analysiert werden soll, genau definiert werden. In Kapitel 2 wird in Zusammenhang mit dem „Messen“ und „Datenerhebungsmethoden“ dargestellt, dass in der quantitativen Forschung Messvorschriften und Messmethoden (Operationalisierung) explizit definiert werden. Nur wenn Merkmale, wie zum Beispiel „Gesundheitszustand“, „Behandlungsmotivation“ oder „Behandlungszufriedenheit“, im Zahlenformat aussagekräftig abgebildet werden, sind die Voraussetzungen für eine solide quantitative Datenanalyse gegeben. Ein erheblicher Teil der gesundheitswissenschaftlichen Forschung beschäftigt sich vorwiegend damit, Messinstrumente und diagnostische Methoden zu entwickeln und zu prüfen (Jerusalem et al., 2018). Hierdurch wird der Eindeutigkeit und Klarheit der Interpretation von Begriffen ein besonders hoher Stellenwert eingeräumt.
- **Einhaltung von Gütekriterien**: Sowohl für die Erfassung von Merkmalsausprägungen bzw. Messergebnissen als auch die verwendete Studienanlage (Welche Gruppen müssen zu welchen Messzeitpunkten untersucht werden?) sind Gütestandards oder -kriterien definiert. Nur wenn diese erfüllt sind, sind die Studienbefunde angemessen interpretierbar und verwertbar.
- **Kontrolle von Störgrößen**: Standardisierte Untersuchungsdesigns ermöglichen es den Forschenden, die Studienanlage so zu wählen, dass die gewonnenen Daten und Informationen möglichst eindeutig interpretiert werden können. Die Forschenden können sich darauf verlassen, dass durch die Anwendung der Untersuchungsanlage verzerrende Einflüsse, die die eigentlich interessierenden Informationen überlagern, bestmöglich kontrolliert werden (Wirtz, 2018b).

BOX – Beispiel

Kontrolle von Störgrößen: Angenommen, alkoholabhängige Patientinnen und Patienten können sich freiwillig für ein Abstinenzprogramm entscheiden. Dann könnte es sein, dass sich motiviertere und veränderungsbereitere Alkoholabhängige eher für die Teilnahme am Programm entscheiden. Gelingt es nun denjenigen, die am Abstinenzprogramm teilnehmen, zukünftig besser, ihren Alkoholkonsum zu kontrollieren, als denjenigen, die nicht am Programm teilnehmen, so spiegelt sich darin nicht notwendigerweise die Wirksamkeit oder der Effekt des Programms eindeutig wider. Es kann nicht eindeutig abgeleitet werden, dass ausschließlich die Teilnahme am Behandlungsprogramm bzw. das im Programm Erlernte für diesen positiven Effekt verantwortlich ist. Es gibt nämlich eine alternative Erklärung: Diejenigen, die veränderungsbereiter und motivierter sind, verändern ihr Verhalten eher in günstigerer Weise eher als diejenigen, die generell unmotivierter und weniger veränderungsbereit sind. Möchte man die Wirksamkeit des Behandlungsprogramms aussagekräftig prüfen, so muss man also darauf achten, dass andere möglicherweise syste-

BOX – Definition

Gütekriterien: Die wichtigsten Gütekriterien in wissenschaftlichen Studien beziehen sich darauf, dass die Ergebnisse nicht von der Person des Forschenden beeinflusst werden dürfen (Objektivität), die Ergebnisse genau und stabil sein müssen (Reliabilität/Zuverlässigkeit) und die Interpretation von Daten und Ergebnissen tatsächlich das widerspiegelt, was auch tatsächlich untersucht werden sollte (Validität/Gültigkeit) (Bühner, 2021).

matischen Einflüsse (Störgrößen: Motivation, Veränderungsbereitschaft) diesen Effekt nicht überlagern. In einer Forschungsstudie sollte also möglichst sichergestellt werden, dass sich die Veränderungsbereitschaft und Motivation zwischen der Behandlungsgruppe und der unbehandelten Vergleichsgruppe nicht unterscheidet. Sind Motivation und Veränderungsbereitschaft in beiden Gruppen gleich, so lässt sich eine Verbesserung in der Behandlungsgruppe im Vergleich zur Vergleichsgruppe eindeutiger auf die Behandlung zurückführen. Die Wirksamkeit der Behandlung an sich kann eindeutiger belegt werden, wenn zwei ansonsten vergleichbare Gruppen unter den Studienbedingungen (Teilnahme vs. keine Teilnahme am Behandlungsprogramm) untersucht werden.

- **Systematische Auswertung der Daten:**

Die letztendliche Entscheidung, wie die untersuchten Forschungsfragen auf Basis der gewonnenen Daten beantwortet werden können, erfolgt durch die Anwendung *statistischer Testverfahren*. Je nach Untersuchungsanlage und Forschungsfragestellungen, muss dasjenige statistische Verfahren ausgewählt werden, das die angemessenste und aussagekräftigste Beantwortung der Fragestellung ermöglicht. Es existieren genau definierte Entscheidungskriterien, ob Datenstrukturen (z.B. Unterschiede zwischen Gruppen oder Zusammenhänge zwischen Merkmalen) als rein zufallsbedingt betrachtet werden müssen (Forschungshypothese widerlegt) oder aber als systematisch gelten können (Forschungshypothese bestätigt) (Kapitel 7). Die Statistik bietet vielfältige Möglichkeiten, die Stärke von Effekten vergleichend und einheitlich zu beurteilen. Hierdurch wird ein einheitlicher statistischer Vergleichsmaßstab zur Verfügung gestellt, der eine begründete Beurteilung und Einordnung von Effekten aus unterschiedlichen Studien auf rein statistischer Basis ermöglicht (Kapitel 8).

- **Explizit begründete Interpretation von Informationen:** Interpretation von Befunden einer empirischen Studie bedeutet vor allem, dass die konkreten Studienergebnisse in Bezug auf das übergreifende Forschungsinteresse, die Merkmale des verwendeten Studienansatzes und die Literatur reflektiert und eingeordnet werden. Hierbei ist eine enge Verknüpfung der ermittelten statistischen Ergebnisse und der Interpretationsergebnisse sehr wesentlich. Die Forschenden müssen verdeutlichen, was möglichst eindeutig aus den Studienbefunden für die übergreifende Forschungsthematik hervorgeht. Angenommen, ein Behandlungskonzept besteht aus mehreren Elementen (z.B. Vermittlung von Erkrankungswissen, Vermittlung von Behandlungsprinzipien, Stärkung der Patientenkompetenz, Umgang mit der Erkrankung, aktives Behandlungsmanagement im Alltag). Dann sollte bei der Interpretation deutlich werden, für welche Elemente welche Effekte nachgewiesen werden konnten. Dabei sollte betont werden, welche Befunde das Konzept des Behandlungsprogramms bestätigen und welche Befunde unerwarteterweise nicht im Einklang mit den Annahmen der Forschenden stehen. Die Rezipientinnen und Rezipienten der Studie müssen genau verstehen können, was die Daten und die Analyseergebnisse für die Wirksamkeit und gegebenenfalls für die Optimierungsnotwendigkeiten des Behandlungsprogramms bedeuten. Jedes Behandlungsprogramm muss sowohl empirisch (d.h. aufgrund der Ergebnisse vorangehender Forschungsstudien) als auch konzeptionell und theoretisch (d.h. aufgrund von theoretischen Modellen, die die zugrunde liegenden Wirkprozesse beschreiben) gut begründet sein. Es ist grundsätzlich anzustreben, die Bedeutung der Befunde für die theoretischen Annahmen zu klären: Kann davon ausgegangen werden, dass die theoretischen Annahmen für das untersuchte Forschungsfeld Gültigkeit besitzen oder

müssen die zugrunde gelegten Annahmen verworfen oder aber angepasst werden?

- **Wiederholbarkeit und Stabilität:** Ergebnisse sind nur dann als gesichert zu akzeptieren, wenn davon ausgegangen werden kann, dass sich diese bei wiederholten Untersuchungen immer wieder in ähnlicher Weise zeigen werden. Einen wissenschaftlich zuverlässigen Befund gewinnt man nicht dadurch, dass man in einer bestimmten Gruppe von Personen eine Struktur oder einen Zusammenhang zwischen Merkmalen erkennt. Erst wenn sich diese Struktur gezielt vorhersagen lässt, kann von der Tragfähigkeit und Allgemeingültigkeit der Aussage ausgegangen werden.

BOX – Beispiel

Wiederholbarkeit und Stabilität von Beobachtungen: Angenommen, eine Person vermutet, dass in einer Spielbank die Ergebnisse beim Roulettespiel systematischen Regeln folgen. Sie beobachtet am Roulettetisch 8 aufeinanderfolgende Spiele. Es ergeben sich folgende Zahlen- und Farbenreihen (in Klammern; jede Zahl ist beim Rouletterad mit einer Farbe kombiniert).

1. Spiel: 26 (schwarz)
2. Spiel: 35 (schwarz)
3. Spiel: 22 (schwarz)
4. Spiel: 23 (rot)
5. Spiel: 17 (schwarz)
6. Spiel: 9 (rot)
7. Spiel: 5 (rot)
8. Spiel: 16 (rot)

Obwohl die Ergebnisse das Produkt reinen Zufalls sind, vermeint die Person, Regeln erkennen zu können. Zu Beginn erscheinen eher schwarz markierte Zahlen, am Ende eher rot markierte Zahlen. Dieselbe Farbe erscheint maximal in 3 aufeinanderfolgenden Spielen. Die Farbe grün, mit der beim Rouletterad die „0“ markiert ist, erscheint grundsätzlich nicht. In den ersten 4 Durchgängen erscheinen Zahlen im oberen Werte-

spektrum, in den letzten 4 Durchgängen Zahlen des unteren Wertespektrums. Die Differenz zweier aufeinanderfolgende Zahlen beträgt maximal 11.

Man beachte, dass alle diese Aussagen richtig sind, wenn man sie auf die bekannten Daten bezieht. Trotzdem sind die abgeleiteten Regeln völlig wertlos und falsch, wenn daraus zukünftige Ergebnisse vorhergesagt werden sollen. Denn diese „Regeln“ sind nicht gültig für die Ergebnisse beim Roulettespiel im Allgemeinen, sie beschreiben lediglich die bekannten Daten. In den Forschungsmethoden würde man von einer *A-posteriori-Analyse* („im Nachhinein“ bzw. nach Kenntnis der Daten) sprechen: Wir betrachten Daten und versuchen, eine Systematik zu erkennen. Dies reicht aber nicht, um die erkannten Auffälligkeiten als generell gültige Systematik abzuleiten. Einen wirklichen Erkenntniswert würden diese „Regeln“ nur besitzen, wenn diese uns erlauben würden, zukünftige Ereignisse gezielt vorherzusagen (*A-priori-Vorhersage* bzw. „im Vorhinein“).

Entsprechend müsste diese Person die von ihr postulierten Regeln erst prüfen, bevor sie diesen einen wirklichen Erkenntniswert beimisst. Angenommen, sie wiederholt ihre Beobachtung am nächsten Tag: Es ergebe sich dann folgende Ergebnisreihe:

1. Spiel: 24 (schwarz)
2. Spiel: 21 (rot)
3. Spiel: 13 (rot)
4. Spiel: 13 (schwarz)
5. Spiel: 6 (schwarz)
6. Spiel: 17 (schwarz)
7. Spiel: 15 (schwarz)
8. Spiel: 16 (rot)

Die am ersten Tag vermuteten Regeln bestätigen sich zumeist nicht: Zu Beginn erscheint nicht auffällig gehäuft schwarz, am Ende nicht auffällig gehäuft rot. Dieselbe Farbe erscheint nun vier Spiele nacheinander (4. bis 7. Runde). Das obere bzw. untere Zahlenspektrum ist nun gleichverteilt über die ersten vs. letzten vier Runden. Nur zwei der

„Regeln“ scheinen sich zu bestätigen: Die Farbe grün erscheint nicht und zwei aufeinanderfolgende Zahlen unterscheiden sich nie mehr als 11 Punkte.

Hier zeigt sich ein weiteres Problem, dessen man sich in der Forschung bewusst sein muss: Untersucht man kleine Stichproben und viele potenzielle Gesetzmäßigkeiten, so läuft man Gefahr, dass sich einzelne vermutete falsche Gesetzmäßigkeiten rein zufallsbedingt erneut zeigen. In der Statistik legt man deswegen darauf Wert, dass möglichst *große Stichproben* untersucht werden (bei 8 Spielrunden treten viel eher ins Auge springende Auffälligkeiten auf als bei z.B. 80 Spielrunden) und Befunde häufig *repliziert* werden: Beim Roulettespiel zeigt sich vielleicht zufallsbedingt an zwei aufeinanderfolgenden Tagen dieselbe zufallsbedingte Auffälligkeit bei den ersten 8 Spielen des Tages, an 10 aufeinanderfolgenden Tagen ist dies deutlich unwahrscheinlicher und nahezu auszuschließen). Zudem sollten möglichst gezielt wenige Fragestellungen an einem Datensatz überprüft werden: Fokussiert man sich auf wenige Fragestellungen und untersucht diese gezielt, ist die Gefahr geringer, dass Zufallsbefunde auftreten.

Und was hat nun das Roulettespiel mit Gesundheit zu tun? Stellen Sie sich vor, ein Hausarzt oder eine Hausärztin untersucht 20 Personen, die sich vor zwei Wochen mit dem COVID-19-Virus infiziert haben. Unter den 10 Personen, die an einer Allergie leiden, berichten 6 – also 60% – über immer noch vorliegende deutliche Erschöpfungssymptome. Bei den 10 Personen, die an keiner Allergie leiden, geben lediglich 2 – also 20% – an, noch an deutlichen Erschöpfungssymptomen zu leiden. Obwohl zwischen den Gruppen ein augenscheinlicher Unterschied zu bestehen scheint (60% vs. 20%), wäre es völlig unzulässig zu behaupten, dass ein Zusammenhang zwischen Allergieerkrankung und dem Vorliegen von Erschöpfungssymptomen als geprüft und nachgewiesen gelten kann. Diese A-posteriori-Vermutung des

Hausarztes bzw. der Hausärztin müsste erst an neuen Daten einer deutlich größeren Stichprobe a priori geprüft werden. Also auch bei Gesundheitsdaten ist die Frage: „Ist eine auffällig scheinende Struktur als systematisch (bzw. informativ) oder zufallsbedingt zu betrachten?“, von grundlegender Bedeutung, um die Verlässlichkeit empirischer Erkenntnisse sicherzustellen.

- **Überprüfung der Gültigkeit von Annahmen und Befunden:** Werden statistische Methoden eingesetzt, um nach auffälligen Strukturen zu suchen, handelt es sich um ein *informationssuchendes* oder *exploratorisches Vorgehen*. Wie unter dem vorangegangenen Punkt dargestellt, besteht hier das Problem, dass die Wiederholbarkeit bzw. Stabilität der gefundenen Strukturen nicht sichergestellt werden kann. Nur wenn die Studie erneut durchgeführt wird (*Replikationsstudie*), lässt sich die Wiederholbarkeit bzw. Stabilität sicherstellen.

Den Ergebnissen *prüfender* oder *konfirmatorischer Forschungsansätze* ist grundsätzlich ein höherer Erkenntniswert beizumessen, als den Befunden exploratorischer Forschungsansätze. Bei konfirmatorischen Ansätzen wird zu Beginn der Studie genau formuliert, welche Effekte theoriebasiert erwartet werden. Die empirische Studie liefert dann die Antwort, ob diese Erwartung gerechtfertigt war oder verworfen werden muss. Die Anwendung dieses Prinzips ist wichtig, um die Stabilität, Genauigkeit und Verlässlichkeit von Studienbefunden zu gewährleisten.

- **Wissenschaftliche Aussagen müssen prinzipiell widerlegbar oder falsifizierbar sein:** Aussagen können nur wissenschaftlich sein, wenn explizit angegeben werden kann, woran erkenntlich würde, dass sie falsch sind oder nicht zutreffen. Dieses Kriterium erscheint auf den ersten Blick überraschend, da die Bewertung einer Aussage genau daran festgemacht wird, dass sie auch nicht zutref-