

Lernziele

- Die unterschiedlichen Vorgehensweisen bei der Bildung theoretischer Konzepte in der qualitativen Forschung und der Operationalisierung theoretischer Konzepte in der quantitativen Forschung kennen.
- Wissen, was man unter der Konzeptspezifikation versteht und wie man dabei vorgeht.
- Latente und manifeste Variablen voneinander abgrenzen können.
- Die vier Skalenniveaus definieren und an Beispielen erläutern können.
- Die Ratingskala als Antwortformat von der psychometrischen Skala als Messinstrument unterscheiden können.
- Die Messung theoretischer Konzepte mit Einzelindikatoren, psychometrischen Skalen und Indizes als drei zentralen Operationalisierungsvarianten in ihren Vor- und Nachteilen abwägen können.

Angenommen, wir möchten empirisch die Forschungshypothese prüfen, dass die Nutzung gewalthaltiger Computerspiele die Aggressivität steigert. Als Untersuchungsdesign kommt hierfür eine explanative Studie infrage, also eine theorieprüfende quantitative empirische Primärstudie, die (sofern kein Auftraggeber existiert) der Grundlagenforschung zuzurechnen wäre (zum Untersuchungsdesign ► Kap. 7). Die explanative Studie hat dabei den größten kausalen Erklärungswert, wenn sie als echtes Experiment realisiert wird, also wenn die Nutzung gewalthaltiger Computerspiele experimentell variiert wird und die Versuchspersonen den experimentellen Bedingungen zufällig zugeordnet werden, ehe am Ende der Effekt auf die Aggressivität gemessen wird. Entsprechende Laborexperimente sind in der Computerspielforschung recht verbreitet. Sie fokussieren dabei kurzfristige Effekte (zum kausalen Nachweis langfristiger Effekte wären Längsschnittstudien wie z. B. prospektive Kohortenstudien indiziert, ► Abschn. 7.8.2).

Bevor wir nun eine Stichprobe von Versuchspersonen ziehen und die Daten im Labor tatsächlich erheben können, müssen wir uns noch ganz genau darüber klar werden, was wir mit den in der Forschungshypothese verwendeten **theoretischen Konzepten** bzw. **Variablen** „Nutzung gewalthaltiger Computerspiele“ und „Aggressivität“ im Einzelnen meinen bzw. wie wir diese Konzepte definieren. In der **quantitativen Forschung** erfolgt die Definition von theoretischen Konzepten grundlegend anders als in der qualitativen Forschung (► Abschn. 8.1). Entsprechende inhaltliche Festlegungen theoretischer Konzepte werden als **Konzeptspezifikation** bezeichnet (► Abschn. 8.2). An die Konzeptspezifikation schließt sich in der quantitativen Forschung die **Operationalisierung** an, die durch die Auswahl von Indikatoren

und Datenerhebungsinstrumenten vorgibt, wie die Konzepte zu messen sind, d. h., wie Objekten mit unterschiedlicher Variablenausprägung aussagekräftige quantitative Messwerte zuzuordnen sind (► Abschn. 8.3). Die Messung kann generell auf vier unterschiedlichen **Skalenniveaus** erfolgen und erzeugt dabei unterschiedlich aussagekräftige numerische Daten (► Abschn. 8.4). Theoretische Konzepte können über **Einzelindikatoren** gemessen werden (► Abschn. 8.5), häufiger werden komplexe Konzepte jedoch durch eine Zusammenfassung mehrerer Indikatoren erfasst, nämlich über sog. psychometrische **Skalen** (► Abschn. 8.6) und **Indizes** (► Abschn. 8.7). Von standardisierten bzw. psychometrischen Messinstrumenten wird in der quantitativen Forschung erwartet, dass sie die **Gütekriterien** der Objektivität, Reliabilität und Validität erfüllen (zum Qualitätskriterium der methodischen Strenge in der Wissenschaft ► Abschn. 3.1.2; zu Gütekriterien von psychometrischen Skalen ► Abschn. 10.4.1).

8.1 Theoretische Konzepte in der qualitativen und quantitativen Forschung

Jede Forschungsfrage oder Forschungshypothese enthält **theoretische Konzepte**, **Konstrukte** bzw. **Begriffe** („theoretical construct“, „theoretical concept“), die im Rahmen empirischer Forschung mit der beobachtbaren Erfahrungswirklichkeit verknüpft werden müssen. Dabei geht man in der qualitativen Forschung induktiv, in der quantitativen Forschung dagegen deduktiv vor (Neumann 2003, S. 177f.).

8.1.1 Induktive Bildung theoretischer Konzepte

Auf der Basis relativ offener Forschungsfragen und nicht-standardisierter Erhebungsinstrumente wird in der qualitativen Forschung reichhaltiges nicht-numerisches (z. B. textuelles, audiovisuelles, multimediales) Datenmaterial erhoben und dieses im Rahmen der qualitativen Datenanalyse und weiteren Datenerhebung schrittweise zu theoretischen Konzepten verdichtet. Die Spezifikation der theoretischen Konzepte sowie Angaben zu ihren Relationen untereinander sind somit das **Ergebnis** einer qualitativen Studie (► Tab. 8.1). Die Konzeptbildung erfolgt primär induktiv bzw. „bottom-up“, d. h. auf der Basis von empirischen Daten werden durch Abstraktion übergeordnete theoretische Begriffe gebildet (zur induktiven Konzeptbildung in der qualitativen Forschung, speziell im Grounded-Theory-Ansatz Glaser 2002). In der qualitativen Sozialforschung entfällt somit die Arbeitsphase

■ **Tabelle 8.1** Induktive Bildung theoretischer Konzepte in der qualitativen Forschung

3. Theoretische Ebene	Detaillierte Darstellung der theoretischen Konzepte (Konzeptspezifikation) und ihrer Relationen untereinander (Theoriebildung) als Ergebnis einer qualitativen Studie
↑	
2. Operationalisierung	Interpretation und Verdichtung des reichhaltigen nicht-numerischen Datenmaterials im Rahmen der qualitativen Datenanalyse
↑	
1. Empirische Ebene	Datenerhebung mithilfe nicht-standardisierter Erhebungsinstrumente anhand von Forschungsfragen

■ **Tabelle 8.2** Deduktive Definition und Messung theoretischer Konzepte in der quantitativen Forschung

1. Theoretische Ebene	Präzise Definition der theoretischen Konzepte (Konzeptspezifikation) und Angaben zu ihren Relationen untereinander in Form von theoretisch abgeleiteten Hypothesen als Ausgangspunkt einer quantitativen Studie
↓	
2. Operationalisierung	Auswahl vorhandener und/oder Entwicklung eigener standardisierter Messinstrumente (z. B. Einzelindikatoren, psychometrische Skalen und Indizes), um die theoretischen Konzepte in messbare Variablen zu überführen
↓	
3. Empirische Ebene	Erhebung von Messwerten für alle Objekte in der Stichprobe und statistische Datenanalyse zu den Relationen der Variablen, auf dieser Basis Rückschlüsse auf die Gültigkeit der Hypothesen als Ergebnis der quantitativen Studie

der Operationalisierung vor der Datengewinnung. Vielmehr verlangt es das qualitative Wissenschaftsverständnis ausdrücklich, **theoretische Offenheit** zu bewahren, damit sich die Bedeutung von theoretischen Konzepten im Zuge der Datenerhebung und Datenanalyse noch verändern bzw. dem Untersuchungsgegenstand annähern kann. Präzise Definitionen von theoretischen Konstrukten und deren Festschreibung in Messinstrumenten durch Operationalisierung werden im qualitativen Paradigma abgelehnt (► [Abschn. 2.3.2](#)).

8.1.2 Deduktive Definition theoretischer Konzepte

In der quantitativen Forschung werden die theoretischen Konzepte im Vorfeld der Datenerhebung ganz genau definiert, so dass mithilfe standardisierter Messinstrumente Merkmalsausprägungen in Form von sinnvoll interpretierbaren numerischen Messwerten erhoben werden können. Es ist also vor allem in der quantitativen Forschung von großer Bedeutung, bereits **vor** der Datenerhebung und Datenanalyse alle theoretischen Konzepte präzise auszuarbeiten und die standardisierten Messverfahren für die einzelnen Variablen detailliert festzulegen (■ [Tab. 8.2](#)).

Die Operationalisierung in der quantitativen Forschung umfasst sowohl die bewusste Auswahl vorhandener als auch die Entwicklung eigener standardisierter Messinstrumente (z. B. Beobachtungspläne, Fragebögen). Die Konzeptbildung erfolgt deduktiv bzw. „top-down“: Aus Theorien werden Hinweise darauf entnommen, welche Konzepte wie zu verstehen und zu messen sind, um mithilfe der entsprechenden Messinstrumente empirische Daten zu gewinnen. Das vorliegende Kapitel zur Operationalisierung ist somit nur für die quantitative sowie die Mixed-Methods-Forschung relevant.

8.2 Deduktive Konzeptspezifikation und Variablen

Quantitative empirische Sozialforschung strebt auf der Basis der Wissenschaftstheorie des Kritischen Rationalismus (► [Abschn. 2.2](#)) eine Analyse der sozialen Wirklichkeit an, bei der **einzelne Merkmale** der Erfahrungswirklichkeit und deren Relationen untereinander genau betrachtet werden. Es wird also nicht der „ganze Mensch“, das „komplette Unternehmen“ oder eine „vollständige Gesellschaft“ untersucht. Denn dies wäre für eine einzelne empirische Studie viel zu komplex und letztlich

nicht leistbar. Indem man theoriegestützt und systematisch jeweils Ausschnitte der Erfahrungswirklichkeit analysiert, soll gemäß kritischem Rationalismus ein wissenschaftlicher Erkenntnisfortschritt möglich sein. Dabei interessieren generell nur Merkmale, die unterschiedliche Ausprägungen annehmen können (**Variablen**). Merkmale, die bei allen Untersuchungsobjekten unveränderlich sind bzw. völlig identisch ausgeprägt sind (**Konstanten**), bieten keine Grundlage für Forschungsaktivitäten. Hinsichtlich der unterschiedlichen Ausprägungen von Variablen unterscheiden wir in der quantitativen Forschung

- **stetige bzw. kontinuierliche Variablen** („continuous variable“) mit unendlich vielen, beliebig fein abgestuften Ausprägungen (z. B. Reaktionszeit, Intelligenz) und
- **diskrete bzw. diskontinuierliche Variablen** („discrete variables“, „discontinuous variables“), die nur wenige verschiedene Werte annehmen können (z. B. Kinderzahl, Automarke, Parteipräferenz).

Variable – Die Variable („variable“) ist ein Merkmal, das verschiedene Ausprägungen annehmen kann. Man unterscheidet **stetige Variablen** („continuous variable“) mit unendlich vielen, fein abgestuften Ausprägungen von **diskreten Variablen** („discrete variable“) mit einer kleinen und begrenzten Zahl an Ausprägungen.

Welche Merkmale bzw. Variablen im Rahmen einer empirischen Studie betrachtet werden, hängt vom Forschungsproblem ab. In dem aus dem Forschungsproblem abgeleiteten Satz an Forschungsfragen bzw. Forschungshypothesen (► Abschn. 5.2) werden jeweils alle relevanten Merkmale genannt. Dabei lassen sich hinsichtlich Messbarkeit und theoretischem Gehalt zwei Typen von Merkmalen bzw. Variablen differenzieren:

1. **Manifeste Merkmale** („manifest variables“, „observable variables“) sind in ihrer Ausprägung unmittelbar beobachtbar. So muss ein Interviewpartner im persönlichen Gespräch nicht erst ausdrücklich danach gefragt werden, ob er im Rollstuhl sitzt oder eine weiße Hautfarbe hat – entsprechende Merkmale sind meistens offensichtlich. Auch Merkmale wie Körpergröße oder Lebensalter, die z. B. im Ausweis registriert sind, lassen sich den manifesten Variablen zuordnen. Bei manifesten Variablen ist es nicht nur sehr einfach, die Ausprägung festzustellen bzw. zu messen, sondern sie sind in der Regel auch theoretisch nicht sehr kompliziert. In einer Studie, in der manifeste Variablen wie „Alter“, „Familienstand“ oder „Wohnort“ erfasst werden, wird man kaum ausführliche theoretische Erläuterungen dazu finden, was mit „Alter“ oder „Wohnort“ gemeint ist, denn die Bedeutung dieser einfachen manifesten Variablen kann in der Regel als bekannt und unstrittig vorausgesetzt werden.

2. **Latente Merkmale** („latent variables“) sind im Gegensatz zu manifesten Variablen nicht direkt beobachtbar und zudem in ihrer theoretischen Bedeutung oft kompliziert und erklärungsbedürftig. Bevor man überhaupt messen kann, welche Ausprägung von „Intelligenz“, „Aggressivität“, „Religiosität“, „Ausländerfeindlichkeit“, „Eifersucht“ oder „Ehrgeiz“ eine Person aufweist, muss erst theoretisch geklärt werden, was im Einzelnen unter diesen Konzepten verstanden werden soll.

Manifestes Merkmal bzw. manifeste Variable – Bei einem **manifesten Merkmal** bzw. einer **manifesten Variable** („manifest variable“, „observable variable“) sind die Ausprägungen direkt beobachtbar bzw. leicht feststellbar und die theoretische Bedeutung gilt als eindeutig und bekannt.

Latentes Merkmal bzw. latente Variable – Bei einem **latenten Merkmal** bzw. einer **latenten Variablen** („latent variable“) bzw. einem **theoretischen Konzept/Konstrukt** („theoretical concept“, „theoretical construct“) sind die Ausprägungen nicht direkt beobachtbar und die theoretische Bedeutung ist erklärungsbedürftig.

Die besonderen Anstrengungen der präzisen Konzeptspezifikation und anschließenden Operationalisierung beziehen sich in der quantitativen Forschung primär auf latente Merkmale bzw. theoretische Konzepte. Zuweilen kann es aber auch sinnvoll und notwendig sein, vermeintlich einfache manifeste Variablen (z. B. „Familienstand“ oder „Geschlecht“) theoretisch genauer zu betrachten und diese theoretischen Überlegungen bei der Operationalisierung zu berücksichtigen (► Abschn. 8.5.1).

8.2.1 Nominaldefinition von theoretischen Konzepten

In einer quantitativen empirischen Studie müssen alle interessierenden theoretischen Konzepte spezifiziert bzw. alle latenten Merkmale präzise definiert werden. Dabei handelt es sich jeweils um sog. Nominaldefinitionen (Kromrey 2000a, S. 163; Schnell, Hill, & Esser, 2008, S. 128). Bei einer **Nominaldefinition** wird der zu definierende Begriff (Definiendum) durch einen oder mehrere bereits bekannte Begriffe (Definiens) festgelegt. Beispiel: Organisationale Gerechtigkeit (Definiendum) meint die Wahrnehmung der Beschäftigten von der Fairness am Arbeitsplatz (Definiens; zur Entwicklung von Definitionen der organisationalen Gerechtigkeit Greenberg 1990).

Voraussetzung für Nominaldefinitionen ist also ein Bestand an bereits eingeführten bekannten Begriffen, die als nicht weiter erklärungsbedürftig gelten und somit zur Präzisierung des zu definierenden Konzeptes verwendbar sind. Nominaldefinitionen sind nichts anderes als begriffliche Umformulierungen. Sie sollen die intersubjektive

Nachvollziehbarkeit der verwendeten Konzepte sicherstellen. Die Präzisierung eines theoretischen Begriffs kann auch als **Explikation** bezeichnet und von der Definition insofern abgegrenzt werden, als die Definition im engeren Sinne auf die Klärung des Sprachgebrauchs abzielt, während die Explikation theoretische Festlegungen vornimmt.

Welche unterschiedlichen Arten von Definitionen neben der Nominaldefinition existieren (z. B. Realdefinition) oder wie Begriffsdefinitionen und Explikationen formalisiert und analysiert werden können, wird im Rahmen der Sprachphilosophie, der Logik sowie der Wissenschaftstheorie detaillierter behandelt. Entsprechende Details sind für die empirische Forschungspraxis jedoch von untergeordneter Bedeutung und werden hier nicht vertieft (zur Wissenschaftstheorie ► [Kap. 2](#)).

Konzeptspezifikation – Unter Konzeptspezifikation („conceptualization“, „concept specification“) versteht man die Nominaldefinition bzw. Explikation eines theoretischen Konzeptes, Konstruktes bzw. latenten Merkmals. Bei komplexen Merkmalen nennt die Definition theoriebasiert in der Regel noch verschiedene Subdimensionen oder Aspekte des Konzeptes.

Man unterscheidet bei theoretischen Konzepten die intentionale Bedeutung (**Begriffsinhalt**) und die extensionale Bedeutung (**Begriffsumfang**) und somit **zwei Varianten von Nominaldefinitionen**:

1. Soll definiert werden, was unter dem Konzept der „Gewalthaltigkeit“ eines Computerspiels zu verstehen ist, so sind im Rahmen einer **intensionalen Nominaldefinition** eine Reihe von Eigenschaften zu nennen, anhand derer man gewalthaltige von weniger oder gar nicht gewalthaltigen Spielen abgrenzen kann. Entsprechende Merkmale könnten z. B. sein: a) dass die Spielerin oder der Spieler im Rahmen des Spiels selbst Gewalt ausüben kann, b) dass man im Zuge der Gewalthandlung im Spiel andere Spieler töten kann und c) dass die Tötungshandlungen im Spiel für den Spielerfolg notwendig sind. Die begründete Auswahl dieser Merkmale entspricht dem Begriffsinhalt. Sofern ein Spiel genau diese Merkmale erfüllt, gilt es als gewalthaltiges Spiel. Bei der Konzeptspezifikation greift man meistens auf intentionale Nominaldefinitionen zurück.
2. In manchen Fällen ist es jedoch schwierig, den Begriffsinhalt vollständig anzugeben. Dann wird im Sinne einer **extensionalen Nominaldefinition** eine Liste von Beispielen angegeben, um das interessierende Konzept zu definieren. So würde man gewalthaltige Computerspiele laut extensionaler Nominaldefinition spezifizieren als Spiele in der Art wie „Counterstrike“, „Battlefield“, oder „Call of Duty“. Ein Spiel, das den genannten Beispielen ähnelt, würde somit als ge-

Studienbeispiel

Nominaldefinition von Mobbing („bullying“)

In einem von der American Psychological Association (APA) herausgegebenen Sammelband zur Gewaltprävention befindet sich auch ein Beitrag zum Mobbing an Schulen. Dabei wird auf der Basis vorliegender Literatur das Mobbing – im Englischen „bullying“ – als eine Spezialform der Aggression eingeordnet mit folgender intensionaler Nominaldefinition (Orpinas & Horne, 2006):

„Bullying is a type of aggression in which the bully (or the aggressor) is stronger or more powerful than the victim, and the aggressive behaviors are committed intentionally, repeatedly, and over time.“ (Orpinas & Horne, 2006, S. 148)

Zur Erläuterung dieser Definition wird ein Beispiel für Nichtmobbing sowie eines für Mobbing beigefügt:

„Therefore, a fight among gang members of equal power would be aggressive but not bullying, whereas pushing, teasing, or threatening a child who is younger, smaller, or somehow weaker would be bullying if the behavior is repeated over time.“ (Orpinas & Horne, 2006, S. 148)

walthaltiges Spiel gelten. Extensionale Definitionen haben den Vorteil, dass sie leicht generierbar und sehr anschaulich sind, sie haben im Vergleich zu intentionalen Definitionen jedoch den Nachteil geringerer Präzision.

Nominaldefinitionen sind Festlegungen, sie können nicht als wahr oder falsch überprüft werden, sondern sind daran zu messen, ob sie für die Forschung **fruchtbar** sind. Dies setzt voraus, dass verwendete Definitionen sich am bisherigen Stand der Forschung und Theorieentwicklung orientieren. Im einfachsten Fall greift man eine in der Literatur häufig verwendete und brauchbar erscheinende Definition heraus und stützt sich auf diese (► [Studienbeispiel „Nominaldefinition von Mobbing“](#)).

Vielfach finden sich in der Literatur auch Studien, die auf eine theoretische Konzeptspezifikation ganz verzichten und unmittelbar die für die relevanten latenten Merkmale verwendeten Messinstrumente nennen (für Studien zur kurzfristigen Wirkung von Computerspielen auf Aggression ohne Konzeptspezifikation von Aggression siehe z. B. Anderson et al., 2010; Ferguson & Rueda, 2010). Dieses Vorgehen ist insofern nachteilig, als ohne ausdrückliche Konzeptspezifikation der Bezug zwischen den mit den genannten Messinstrumenten erhobenen Daten einerseits und ihrer übergreifendem theoretischen und praktischen Bedeutung andererseits nicht ausdrücklich hergestellt wird. Liest man etwa, dass eine Studie Aggression mittels der Lautstärke von Bestrafungstönen misst, die eine Versuchsperson einem vermeintlichen Spielpartner im Labor appliziert, so muss man selbst den Rück-

Studienbeispiel

Bedeutungsanalyse von Aggression

In einem Forschungsüberblick sollen die vorliegenden Befunde zu den wechselseitigen Beziehungen zwischen Aggression, Feindseligkeit und Ärger dargestellt werden (Ramírez & Andreu, 2009). Die Autoren referieren dazu über drei Seiten hinweg verschiedene Definitionen von Aggression. Dabei wird die „Absicht, einem anderen Lebewesen zu schaden“, als **allgemein konsensfähige Definition von Aggression** herausgearbeitet und mit diversen Literaturquellen belegt. Des Weiteren werden diverse Typen von Aggression auf unterschiedlichen Dimensionen unterschieden:

- Art der Absicht bzw. Zielsetzung: **feindselige Aggression**, bei der die unmittelbare Schädigung Selbstzweck ist (z. B. Schläge im Rahmen von Hass), vs. **instrumentelle Aggression**, bei der die unmittelbare Schädigung Mittel für ein übergeordnetes Ziel ist (z. B. Schläge im Rahmen eines Raubüberfalls).
- Mittel der Schädigung: **physische Aggression**, bei der eine andere Person körperlich angegriffen wird, vs. **verbale/soziale/relationale Aggression**, bei der eine andere Person verbal bzw. im Kontext des Sozialverhaltens angegriffen wird (z. B. Beschimpfungen).
- Aktivität des Aggressors: **aktive Aggression**, bei der die Schädigung durch aktives Handeln erzeugt wird, vs. **passive Aggression**, bei der die Schädigung durch Nichthandeln hervorgerufen wird (z. B. nicht grüßen, nicht zuhören).
- Unmittelbarkeit des schädigenden Verhaltens: **direkte Aggression**, bei der die Schädigung im direkten Face-to-Face-Kontakt auftritt, vs. **indirekte Aggression**, bei der die Schädigung nicht im direkten Kontakt erfolgt (z. B. heimlich Eigentum des Opfers zerstören, hinter dem Rücken der betroffenen Person Gerüchte verbreiten).

Die Autoren konzentrieren sich auf die erstgenannte Dimension (Art der Absicht) und führen genauer aus, warum diese Dimension für das Verständnis von Aggression so wichtig ist, was feindselige vs. instrumentelle Aggression laut vorliegenden Definitionen im Einzelnen kennzeichnet und welche synonymen Bezeichnungen (z. B. reaktive vs. proaktive Aggression) dafür in der Literatur verwendet werden.

schluss ziehen, welche Art von Aggression hier eigentlich gemessen wird und inwiefern diese überhaupt alltagsrelevant ist (► Studienbeispiel „Bedeutungsanalyse von Aggression“).

8.2.2 Bedeutungsanalyse von etablierten theoretischen Konzepten

Bei theoretisch umstrittenen Konzepten, für die mehrere unterschiedliche Definitionen vorliegen, muss eine **Bedeutungsanalyse** bzw. **semantische Analyse** der vorliegenden Begriffsverwendungen und Definitionen vorgenommen und eine begründete Auswahl getroffen werden (für eine ausführliche Darstellung siehe Kromrey 2000b,

S. 130ff.). Hierfür greift man auf die existierende Literatur zurück, die im jeweiligen Theorieteil meistens Konzeptualisierungsfragen diskutiert (► Studienbeispiel „Bedeutungsanalyse von Aggression“). Zudem existieren Publikationen, die sich im Sinne einer Bedeutungsanalyse ausschließlich der Darstellung, dem Vergleich und der Integration unterschiedlicher Konzeptualisierungen eines Phänomens in der wissenschaftlichen Literatur widmen (z. B. Konzeptualisierung von „Mobbing am Arbeitsplatz“: Crawshaw 2009; Konzeptualisierung von „Aggression im Straßenverkehr“: Galovski, Malta, & Blanchard, 2006).

Deskriptive und explanative quantitative Studien sind nur sinnvoll, wenn eine entsprechende Forschungs- und Theoriebasis existiert, so dass man auf vorhandene Definitionen zurückgreifen kann. Für die eigene Studie wird man entweder die einschlägigste Definition herausgreifen und übernehmen/anpassen oder zunächst mittels Bedeutungsanalyse alle vorliegenden Definitionsvorschläge analysieren und auf dieser Basis eine Auswahl treffen oder eine Modifikation vornehmen.

8.2.3 Dimensionale Analyse von neuen theoretischen Konzepten

Es sind vor allem explorative Studien, die sich neuen oder bislang wenig untersuchten Gegenständen zuwenden, die es erforderlich machen, eine eigene Konzeptspezifikation zu entwickeln und zu diesem Zweck eine **Dimensionsanalyse** bzw. **dimensionale Analyse** durchzuführen. Während die Bedeutungsanalyse bei den bereits vorliegenden Definitionen und Begriffsverwendungen in der Literatur ansetzt und diese vergleichend beleuchtet, betrachtet die Dimensionsanalyse direkt den **Gegenstand** selbst und versucht anhand von konkreten Beispielen dessen zentrale Merkmale und Aspekte möglichst systematisch herauszuarbeiten.

Diese angestrebte begriffliche Strukturierung erfolgt im Rahmen der dimensionalen Analyse in vier Schritten (Kromrey 2000b, S. 116ff.):

1. **Ideen- und Materialsammlung:** Für eine möglichst breite Sammlung von Ideen und Materialien zum Gegenstand ist auf verschiedene Quellen zurückzugreifen: eigene Alltagserfahrung, Konsultation von Experten, Presseberichte, Fachliteratur zu verwandten Themen, Brainstorming mit Kollegen und vor allem die Betrachtung von Beispielen. Will man z. B. ein Konzept wie „Handykompetenz von Kindern“ einer dimensional Analyse unterziehen, um zu einer genauen Konzeptspezifikation mit allen relevanten Unteraspekten zu kommen, so kann man auf allgemeine Definitionen von „Medienkompetenz“ zurückgreifen,

Studienbeispiel

Konzeptspezifikation von „Globalisierung“ mittels dimensionaler Analyse

Eine politikwissenschaftliche Methodenstudie widmet sich der Konzeptspezifikation von „Globalisierung“, einem Begriff, der bekanntlich unser Zeitalter prägt, dessen genaue Bedeutung jedoch schwer greifbar ist. Auf der Basis vorliegender Theorien und Definitionsvorschläge (**Bedeutungsanalyse**) präsentiert der Autor zunächst eine **allgemeine Nominaldefinition**: „Globalisierung bezeichnet Prozesse der Zunahme sowie der geographischen Ausdehnung grenzüberschreitender anthropogener/sozialer Interaktion.“ (Kessler 2009).

Das Konzept der Globalisierung wird dabei als Variable eines Nationalstaates betrachtet, d. h., verschiedene Länder bzw. Staaten unterscheiden sich im Grad ihrer Globalisierung. Die auf der Basis der Bedeutungs-

analyse herausgearbeitete Nominaldefinition ist für eine Messung noch zu allgemein, so dass mittels einer **dimensionalen Analyse** die **Unterdimensionen des Konstruktes** zu spezifizieren sind. Im Zuge der Dimensionsanalyse wurde hierbei die „grenzüberschreitende soziale Interaktion“ in den Mittelpunkt gerückt, deren Menge und Ausdehnung zunehmen soll. Zur Systematisierung wird ein einfaches Kausalmodell zugrunde gelegt, dementsprechend werden Voraussetzungen, Arten und Folgen grenzüberschreitender Interaktionen differenziert. Die gesammelten Beispiele für Unterdimensionen werden in dieses Schema eingeordnet. Eine solche Konzeptspezifikation durch dimensionale Analyse hilft der weiteren Theoriebildung sowie der Planung empirischer Un-

tersuchungen, weil nun ganz gezielt einzelne Globalisierungsaspekte herausgegriffen und gemessen werden können und man gleichzeitig darstellen kann, welche Globalisierungsaspekte in der jeweiligen Studie nicht untersucht werden (Tab. 8.3).

Für eine konkrete **Messung des Globalisierungsgrades** verschiedener Staaten kommt die Bildung eines **Globalisierungsindex** infrage. Dazu müssen auf der Basis der dimensionalen Analyse Anzahl und Art der relevanten Indikatoren ausgewählt werden. Zudem muss inhaltlich begründet und entschieden werden, welcher Indextyp (z. B. gewichteter additiver Index, ungewichteter multiplikativer Index) zum Einsatz kommen soll (zur Indexbildung ► Abschn. 8.7).

■ **Tabelle 8.3** Ergebnis einer dimensionalen Analyse des Konzepts „Globalisierung“. (Mod. nach Kessler 2009, S. 38)

Voraussetzungen für grenzüberschreitende soziale Interaktionen	→ Arten von grenzüberschreitenden sozialen Interaktionen (über nationalstaatliche Grenzen und wachsende Entfernung hinweg)	→ Folgen von grenzüberschreitenden sozialen Interaktionen
Beispiele:	Beispiele:	Beispiele:
<ul style="list-style-type: none"> – Abbau von Handelsbeschränkungen – Abbau von Kapitalverkehrsbeschränkungen – Liberale politische Systeme – Fremdsprachenkenntnisse – Abbau von Einreise- und Aufenthaltsbeschränkungen – Verfügbarkeit von Transportmitteln, Infrastruktur und Know-how – Verfügbarkeit von Computer-Hard- und Software, Internet, PC-Kenntnissen – Verfügbarkeit von Telefon, Telefax, Fernsehen, Radio 	<ul style="list-style-type: none"> – Zunahme der Handelsströme – Zunahme der Direkt- und Portfolioinvestitionen – Wachsender Austausch von internationalen Organisationen, Politikern, Wissenschaftlern – Zunahme von Touristenströmen, des Personenverkehrs – Zunahme des Gütertransports – Zunahme der übermittelten Informationen in Form von Datenströmen – Zunahme der empfangenen Fernseh- oder Radiosendungen, der Telefongespräche 	<ul style="list-style-type: none"> – Wettbewerbsintensivierung – Entwicklungsimpulse – Höhere externe Verwundbarkeit – Bedeutungszuwachs transnationaler Akteure, internationaler Organisationen – Beschleunigte Ausbreitung von Krankheiten – Ökologische Probleme – Verbreitung extremistischer, krimineller Information – Kulturelle Durchdringung

Presseartikel zu Handygefahren für Kinder sichten oder Eltern und Lehrer dazu befragen, welche nützlichen und schädlichen Umgangsweisen sie bei Kindern mit dem Handy beobachten.

2. **Systematisierung**: Die gesammelten Beispiele und Aspekte werden in eine gegenstandsbezogene Ordnung gebracht: Dabei bietet sich eine Systematisierung nach Ursache-Wirkungs-Relationen, nach zeitlichen Abläufen oder logischen Strukturen (z. B. vom Allgemeinen zum Speziellen) an. Eine Visualisierung in Form von Pfad- oder Ablaufdiagrammen kann dabei helfen, die konzeptuelle Struktur zu verdeutlichen und hinsichtlich möglicher Lücken, Redundanzen

oder Inkonsistenzen zu bearbeiten. Ziel ist eine möglichst erschöpfende Auffächerung der verschiedenen relevanten Aspekte des Gegenstandes. So lässt sich das Konzept Handykompetenz entlang üblicher Dimensionen von Medienkompetenz gliedern (z. B. Medienkunde, Medienkritik, Mediennutzung, Mediengestaltung), möglich ist aber auch eine Strukturierung nach unterschiedlichen Anwendungen und Diensten des Mobiltelefons (z. B. Mobiltelefonie, Handyspiele, Handykamera, Mobile Apps, mobiles Internet) und/oder nach Nutzungskontexten (z. B. Handynutzung zu Hause, in der Schule, unterwegs).

3. **Auswahl der untersuchungsrelevanten Aspekte:** Da die dimensionale Analyse als Hilfsmittel der Konzeptspezifikation bzw. der Gegenstandsstrukturierung eine empirische Studie vorbereiten soll, müssen schließlich auch theoretische und forschungspraktische Erwägungen herangezogen werden: Welche Aspekte lassen sich mit den vorhandenen zeitlichen, finanziellen und personellen Mitteln überhaupt untersuchen? Welche Aspekte sind besonders interessant und vielversprechend hinsichtlich ihres theoretischen oder praktischen Erkenntniswertes? So mag man sich z. B. dafür entscheiden, den Aspekt der Videotelefonie per Handy nicht in die Studie einzubeziehen, weil zum Untersuchungszeitpunkt wenige Kinder diese Option nutzen. Oder man wird insbesondere jene Aspekte der Handykompetenz umfassend untersuchen, deren praktische Förderung – z. B. im Rahmen des Schulunterrichts – als besonders gut möglich erscheint.
4. **Entwicklung eines deskriptiven Begriffsschemas:** Nachdem die in die empirische Untersuchung einzubeziehenden Subdimensionen des Konzepts identifiziert wurden, müssen für diese jeweils schlüssige Nominaldefinitionen formuliert werden. Soll z. B. im Bereich Handywissen der Aspekt „Kenntnis über die Kosten der Mobilkommunikation“ erfasst werden, so muss genau definiert werden, ob damit z. B. gemeint ist, dass Kinder wissen, wie viel ihr Handy als Endgerät kostet, wie teuer Inlands- und Auslandsgespräche mit dem Handy sind oder welche verschiedenen Bezahlmodelle existieren und welchen Tarif sie selbst nutzen.

Neben explorativen empirischen Studien sind es vor allem **Theorie- und Methodenstudien** (► Abschn. 7.3), die komplexe Konzepte einer systematischen dimensional Analyse unterziehen, um deren Messung vorzubereiten (► Studienbeispiel „Konzeptspezifikation von ‚Globalisierung‘ mittels dimensionaler Analyse“).

8.3 Grundlagen zur Durchführung einer Operationalisierung

Auf die Konzeptspezifikation mittels Nominaldefinition, die mithilfe von Bedeutungsanalyse sowie dimensionaler Analyse zu erarbeiten ist, folgt die Operationalisierung, die in eine operationale Definition mündet. Im Folgenden wird die operationale Definition eingeführt. Dabei werden Besonderheiten bei abhängigen und unabhängigen Variablen hervorgehoben und verbreitete Fehlinterpretationen problematisiert.

8.3.1 Operationale Definition von theoretischen Konzepten

Für ein und dasselbe theoretische Konzept existieren meistens verschiedene Möglichkeiten der Operationalisierung. So kann das **theoretische Konstrukt** „Intelligenz“ mit unterschiedlichen **Messinstrumenten** (Intelligenztests) operationalisiert werden, die sich jeweils aus verschiedenen **Indikatoren** (Intelligenztestaufgaben) zusammensetzen, welche einzeln und gemeinsam dazu dienen, die Intelligenzleistung einer Person zu erfassen. Eine **operationale Definition** wie „Intelligenz ist, was der Intelligenztest misst“ bzw. genauer: „Intelligenz ist, was der Hamburg-Wechsler-Intelligenztest misst“ oder „Intelligenz ist, was der Raven-Test misst“ erscheint für sich genommen meistens unbefriedigend. Deswegen sollten operationale Definitionen nicht die Konzeptspezifikation ersetzen. Hat man zuerst im Rahmen der Konzeptspezifikation theoretisch festgelegt, welches Intelligenzkonzept man zugrunde legen möchte und die entsprechende Nominaldefinition angegeben, dann wird man auf dieser Basis gezielt die passende Operationalisierung bzw. den passenden Intelligenztest auswählen. Bei dieser Vorgehensweise ist die verwendete operationale Definition nicht zirkulär, sondern **in einem theoretischen Verständnis von Intelligenz verankert**.

Operationalisierung – Die Operationalisierung („operationalization“) eines theoretischen Konzepts bzw. einer latenten Variable legt fest, anhand welcher beobachtbaren Variablen (**Indikatoren**) die Ausprägung des theoretischen Konzepts bei den Untersuchungsobjekten festgelegt werden soll. Neben der Auswahl der Indikatoren gehört zur Operationalisierung auch die **Festlegung der Messinstrumente**, mittels derer den Ausprägungen der einzelnen Indikatoren jeweils entsprechende numerische Werte zugeordnet und zu einem Gesamtmesswert für das Konstrukt verrechnet werden. Komplexe theoretische Konstrukte werden selten mit einem einzigen Indikator (**Einzeldikator** als Messinstrument) operationalisiert, sondern meist über einen Satz von Indikatoren (d. h. über eine **psychometrische Skala** oder einen **Index**). Mit der Festlegung der Operationalisierung wird für ein theoretisches Konzept (dargelegt über seine Nominaldefinition) eine konkretisierende **operationale Definition** („operational definition“) vorgenommen.

Angenommen, die **Intensität des Mobbing von Schulkindern** wurde im Zuge der Konzeptspezifikation über die Dimensionen a) Anzahl der Aggressoren, b) Dauer des Mobbing und c) Schwere der aggressiven Handlungen definiert. Indikatoren für diese theoretischen Aspekte können nun entsprechend detaillierte Fragen in einem Elternfragebogen oder in einem Schülerinterview sein. Welche Operationalisierungsvariante und damit auch **Datenerhebungsmethode** (► Kap. 10) gewählt wird, hängt u. a. von forschungspraktischen sowie von theoretischen Erwägungen ab. So mag es zwar weniger aufwändig sein,

Eltern kollektiv beim Elternabend einen standardisierten Fragebogen zur Mobbingbetroffenheit ihres Kindes ausfüllen zu lassen als alle Kinder einzeln in kindgerechter Weise zu interviewen. Da jedoch nicht davon auszugehen ist, dass Eltern über alle Mobbingvorfälle ihrer Kinder genau informiert sind, wären vermutlich Indikatorvariablen, die direkt an den Kindern selbst erhoben werden, zu bevorzugen.

Die zur Konzeptspezifikation der **Globalisierung unterschiedlicher Länder** angegebenen theoretischen Dimensionen (z. B. „Touristenströme“; ▶ [Studienbeispiel](#) „Konzeptspezifikation von ‚Globalisierung‘ mittels dimensionaler Analyse“) sind im Zuge der Operationalisierung in konkret messbare Indikatoren zu übersetzen (z. B. Anzahl der Ankünfte internationaler Touristen pro 1000 Einwohner pro Jahr), die z. B. den amtlichen Statistiken zu entnehmen wären. Zuweilen muss aus **forschungspraktischen bzw. forschungswirtschaftlichen Gründen** auf Indikatoren zurückgegriffen werden, die das theoretische Konzept nur teilweise oder nur ungenau abbilden (etwa weil relevante Kennwerte nicht schnell oder kostengünstig genug beschaffbar sind). Abweichungen zwischen theoretischem Konstrukt und den zur Operationalisierung genutzten Indikatoren sind zu begründen und bei der Ergebnisinterpretation zu berücksichtigen (z. B. wenn die interessierenden internationalen „Touristenströme“ über die Zahl der Hotelübernachtungen operationalisiert würden, wobei dann auch nationale Touristen sowie Geschäftsreisende enthalten wären).

Multiple Indikatoren tragen dazu bei, dass die verschiedenen Aspekte eines komplexen theoretischen Konstruktes möglichst vollständig abgebildet werden und somit wirklich das erfasst wird, was gemessen werden soll (Kriterium der Gültigkeit bzw. **Validität** des Messinstrumentes). Zudem können durch den Einsatz multipler Indikatoren auch Messfehler reduziert werden, so dass die Messgenauigkeit bzw. **Reliabilität** des Instruments steigt (zu psychometrischen Gütekriterien im Überblick ▶ [Abschn. 10.4.1](#)). Dementsprechend wird eine groß angelegte bevölkerungsrepräsentative Studie zur Verbreitung von Depressionen mit einem etablierten Depressionsmessinstrument arbeiten, welches das Konstrukt möglichst differenziert mithilfe einer Reihe sorgfältig entwickelter und aufeinander abgestimmter Selbstauskunftsfragen bzw. Indikatoren erfasst. Demgegenüber wird eine Studie, die sich dem Essverhalten widmet, das Konstrukt Depression – wenn es denn am Rande auch erhoben werden soll – möglicherweise nur mit einem **Einzelindikator** (z. B. mit einer einzigen Frage im Fragebogen) erheben, weil alles andere das ohnehin umfangreiche Erhebungsinstrument sonst zu sehr aufblähen, die Untersuchungspersonen demotivieren und somit letztlich den Erfolg der

gesamten Studie einschränken könnte (zur Messung von latenten Merkmalen wie z. B. Depression mittels Einzelindikatoren ▶ [Abschn. 8.5.2](#)).

Wenn anstelle eines Einzelindikators ein aus mehreren Indikatoren bestehendes Messinstrument zur Operationalisierung eines theoretischen Konstruktes genutzt wird, so stellt sich die Frage, welche Indikatoren auf welche Weise zu einem Messinstrument zusammengefasst werden sollen. Die Auswahl, Modifikation oder Neuentwicklung von Messinstrumenten und Indikatoren ist die Kernaufgabe bei der Operationalisierung. Wissenschaftlich unbrauchbar sind **Ad-hoc-Instrumente**, die aus einer mehr oder minder willkürlichen und ungeprüften Menge an Indikatoren bestehen. Gefordert sind stattdessen wissenschaftliche Messinstrumente, deren Aufbau theoretisch begründet und deren Gütekriterien empirisch geprüft sind, so dass man sichergehen kann, dass sie tatsächlich genau das Zielkonstrukt erfassen (Validität) und wenig durch Messfehler verzerrt sind (Reliabilität). Die Forschung rund um die Voraussetzungen und Gütekriterien der Messung psychologischer Merkmale wird auch als **Psychometrie** („psychometrics“) bezeichnet. Eine wichtige Grundlage für die Konstruktion und Bewertung von Messinstrumenten ist die **Testtheorie** (▶ [Abschn. 10.4.4](#)). In Abgrenzung von Ad-hoc-Instrumenten werden überprüfte Messinstrumente auch als **psychometrische Skalen** und **psychometrische Tests** bezeichnet. Typische Gegenstände psychometrischer Messung sind Einstellungen, aber auch Wissen und kognitive Leistungen, Persönlichkeitsmerkmale oder psychologische Störungen, zu denen eine Fülle von Messinstrumenten vorliegen (für eine Auswahl gebräuchlicher Testverfahren ▶ [Abschn. 10.4.3](#)). Auf die Konstruktionsprinzipien von unterschiedlichen psychometrischen Skalen (▶ [Abschn. 8.6](#)) sowie von Indizes (▶ [Abschn. 8.7](#)) wird im Verlauf dieses Kapitels noch genauer eingegangen.

Wichtig ist es im Hinblick auf das **Verhältnis zwischen theoretischem Konstrukt und Indikator**, zwischen zwei Typen von Indikatoren zu unterscheiden: Den reflektiven und den formativen Indikatoren (auch Bühner 2011, S. 37).

■ Beim **reflektiven Messmodell**, das psychometrischen Skalen (▶ [Abschn. 8.6](#)) zugrunde liegt, wird das **theoretische Konstrukt als Ursache** und die **Indikatoren** werden **als Wirkungen** betrachtet. Beispiel: Dadurch, dass Menschen sich im Grad ihrer Schüchternheit unterscheiden (theoretisches Konstrukt als Ursache), beantworten sie Indikatorvariablen bzw. Skalen-Items wie „Mir fällt es schwer, mit Fremden ins Gespräch zu kommen“ oder „Wenn ich mit Fremden spreche, fühle ich mich gehemmt“ in systematischer Weise unterschiedlich (reflektive Indikatoren als Wirkungen).

In den Ausprägungen dieser Indikatoren „reflektiert“ sich also die Ausprägung des latenten Merkmals: Weil eine Person schüchtern ist, stimmt sie entsprechenden Aussagen über schüchternes Verhalten und Erleben tendenziell zu. Dasselbe Ursache-Wirkungs-Prinzip zwischen Konstrukt und Indikatoren gilt für Skalen, mit denen z. B. durch mehrere Skalen-Items (Fragen oder Aufgaben) Konstrukte wie Depression, Intelligenz, Aggressivität oder Religiosität erfasst werden. Die in einer psychometrischen Skala enthaltenen reflektiven Indikatoren sind einander **formal und inhaltlich ähnlich** (z. B. bedeutungsähnliche Selbsteinschätzungen zur Schüchternheit) und relativ hoch miteinander korreliert.

8

- Demgegenüber betrachtet das **formative Messmodell**, welches den Indizes (► Abschn. 8.7) zugrunde liegt, die **Indikatoren als Ursachen** bzw. Determinanten des im Zuge der Konzeptspezifikation definierten theoretischen Konstruktes. Die Ausprägung des **Konstruktes ist eine Wirkung** der Indikatoren. Beispiel: Erst wenn nachweisbar ist, dass ein bestimmtes Land über zunehmende grenzüberschreitende Touristen-, Daten-, Handels-, Finanzströme etc. verfügt, dann wird diesem Land gemäß der Nominaldefinition und Konzeptspezifikation von Globalisierung ein hoher Globalisierungsgrad zugeschrieben. Also nicht weil ein Land globalisiert ist, entstehen dadurch Touristen- und Finanzströme, Wirtschafts- und Wissenschaftsaustausch etc., sondern das Auftreten der inhaltlich ganz verschiedenen grenzüberschreitenden sozialen Interaktionen erzeugt bzw. „formt“ den Globalisierungsgrad eines Landes. Die in einem Index enthaltenen formativen Indikatoren können einander **formal und inhaltlich sehr unähnlich** sein und müssen auch nicht miteinander korrelieren. Ein weiteres Beispiel: Wenn eine Person einen geringen Bildungsstand hat, einen Beruf mit geringem Prestige ausübt und über ein unterdurchschnittliches Einkommen verfügt (drei formative Indikatoren als Determinanten), dann wird ihr definitionsgemäß ein niedriger sozioökonomischer Status zugeschrieben (theoretisches Konstrukt als Wirkung).

Neben Einzelindikatoren werden zur Operationalisierung theoretischer Konstrukte oft **Messinstrumente** verwendet, die aus mehreren Indikatoren bestehen und auf zwei grundlegend verschiedenen Messmodellen basieren.

Reflektives Messmodell – Bei einem reflektiven Messmodell geht man davon aus, dass das zu messende Konstrukt die Ursache für die Merkmalsausprägungen auf den gewählten Indikatoren ist. Als Messinstrument wird eine **psychometrische Skala** genutzt, die aus homogenen bzw. inhaltsähnlichen Fragen, Aussagen oder Aufgaben besteht (sog. **reflektive Indikatoren**, in denen sich das Konstrukt widerspiegelt).

Formatives Messmodell – Bei einem formativen Messmodell geht man davon aus, dass das zu messende Konstrukt die Wirkung oder Folge der Merkmalsausprägungen der Indikatoren ist. Als Messinstrument wird ein **Index** gebildet, in den heterogene Kennwerte eingehen (sog. **formative Indikatoren**, durch die das Konstrukt ursächlich gebildet wird).

Bereits bei der Konzeptspezifikation (► Abschn. 8.2), spätestens im Zuge der operationalen Definition, also bei der Auswahl oder Konstruktion von Indikatoren und Messinstrumenten, sollte man sich darüber klar geworden sein, ob man ein reflektives oder ein formatives Messmodell anzulegen hat, also ob eine psychometrische Skala oder ein Index zu verwenden ist.

8.3.2 Operationalisierung von abhängigen Variablen

In der quantitativen Sozialforschung messen wir Variablen vor allem, um im Kontext explanativer Studien Hypothesen über Variablenzusammenhänge, Gruppenunterschiede oder Veränderungen über die Zeit zu prüfen. Besonders aussagekräftig ist dabei die Prüfung von Kausalhypothesen, wie sie in experimentellen und quasi-experimentellen Designs angestrebt wird (► Abschn. 7.6). In diesen Designs wird mindestens eine unabhängige Variable systematisch variiert (z. B. unterschiedliche Behandlungen, Interventionen, Stimuli), um die Auswirkung auf die abhängige(n) Variable(n) zu prüfen. Dabei versucht man meistens, die Ausprägungen der abhängigen Variablen in **möglichst feinen Abstufungen** zu erfassen. Aber auch in nicht-experimentellen Studien ist eine differenzierte Messung der Variablenausprägungen nützlich. Zur Messung feiner Merkmalsabstufungen sind folgende sechs **Operationalisierungsvarianten** (modifiziert nach Conrad & Maul, 1981, S. 151) besonders geeignet:

1. **Häufigkeit:** Wie oft tritt ein bestimmtes Verhalten auf? (Beispiele: Anzahl der Fehler in einem Diktat, Häufigkeit der Blickkontakte beim Flirt, Häufigkeit von Sprechpausen in einer Vernehmungssituation, Häufigkeit von Ehestreits vor und nach einer Paartherapie)
2. **Reaktionszeit:** Wie viel Zeit vergeht, bis eine Person nach Auftreten eines Stimulus reagiert? (Beispiele: Reaktionslatenz nach Auftreten eines unerwarteten Verkehrshindernisses, Reaktionszeit bis zur Identifikation eines Wortes). Die Reaktionszeitmessung ist in der Regel experimentellen Laborstudien vorbehalten.
3. **Reaktionsdauer:** Wie lange reagiert eine Person auf einen Stimulus bzw. auf eine Intervention? (Beispiele: Lösungszeit für eine Mathematikaufgabe, Verweildauer des Auges auf einem bestimmten Bildausschnitt, Dauer des Nichtrauchens nach einem Anti-Rauch-Training)

4. **Reaktionsstärke:** Wie intensiv reagiert eine Person auf einen Stimulus bzw. auf eine Intervention? (Beispiele: Stärke der Muskelanspannung als Indikator für Aggressivität, geäußerte Stärke von Meinungen auf Ratingskalen, Höhe des Blutdruckanstiegs, Intensität der Zustimmung zu einer Partei vor und nach der Rezeption von Pressebeiträgen über einen Partei-skandal). Die Reaktionsstärke kann im Rahmen von Messwiederholungsdesigns bzw. Längsschnittstudien auch mehrfach erhoben und in ihrem Verlauf betrachtet werden (► Abschn. 7.8).
5. **Reaktionsqualität:** Welche Wertigkeit (Valenz) hat eine Reaktion auf einen Stimulus: Ist sie eher positiv oder negativ bzw. beinhaltet sie Zuwendung oder Abwendung? (Beispiele: Bewertung von Lebensmitteln hinsichtlich Geruch und Geschmack auf Schulnoten-skalen; Einstufung der Sympathie oder Antipathie gegenüber einer Person auf einer Ratingskala; Einstufung der eigenen Stimmungslagen – angespannt, aufmerksam, fröhlich etc. – auf einer psychometrischen Skala.)
6. **Wahlreaktion:** Welche Wahl trifft eine Person angesichts mehrerer Wahlmöglichkeiten? (Beispiele: Bevorzugung eines von zwei Kunstwerken als ästhetischer im Paarvergleichsurteil; Bevorzugung eines von mehreren Reisezielen bei einer Mehrfachwahlaufgabe; Nennung der Lieblingsmarke aus einem Spektrum an Marken bei einem Präferenzurteil.)

Bei den Operationalisierungsvarianten für abhängige Variablen ist der **Zeitpunkt der Messung** zu beachten: Die Messung kann **nach Abschluss der Intervention** (z. B. Stimuluspräsentation) erfolgen, etwa wenn nach dem Betrachten eines Films, nach einem Bewerbungsgespräch oder nach einer Unterrichtsstunde jeweils die Ausprägungen der interessierenden abhängigen Variablen erfasst werden (z. B. mit einer psychometrischen Skala, mit einem psychologischen Testverfahren oder per Expertenurteil). Es besteht aber auch die Möglichkeit, die Messung **prozessbegleitend** durchzuführen (d. h. **während** die Untersuchungsperson auf den Stimulus bzw. die unabhängige Variable reagiert wird z. B. die Reaktionsdauer per Stoppuhr gemessen und die Reaktionsqualität durch Beobachtung von Mimik und Gestik erfasst). Besonders gut geeignet für prozessbegleitende Messungen sind **physiologische Messverfahren** (► Abschn. 10.5), mit denen kontinuierlich während des gesamten Versuchsdurchlaufs z. B. der Blutdruck oder die Blickbewegungen aufgezeichnet werden. Es besteht auch die Möglichkeit, subjektive Bewertungen prozessbegleitend zu erheben. Hierfür werden als Operationalisierungsvarianten das sog. **Real Time Response (RTR) Measurement** bzw. das **Continuous Response Measurement (CRM)** genutzt (zum

Überblick Maier, Maier, Maurer, Reinemann, & Meyer, 2009). Bei diesen Verfahren erhält die Untersuchungsperson als Messinstrument einen **Dreh- oder Schieberegler**, über den sie fortlaufend stufenlos angeben kann, ob und wie stark sie einen präsentierten Reiz gerade positiv oder negativ bewertet (auch andere Reaktionsqualitäten wie interessant vs. langweilig etc. können erfasst werden). Die RTR-Measurement- bzw. CRM-Methode wurde bereits in den 1930er-Jahren entwickelt und wird bis heute u. a. zur prozessbegleitenden Messung von Publikumsreaktionen auf Medienangebote wie Radio- oder TV-Sendungen, Werbespots und Kinofilme genutzt (Maier, Maurer, Reinemann, & Faas, 2006; Reinemann, Maier, Faas, & Maurer, 2005; Schmeisser, Bente, & Isenbart, 2004; Schneider et al., 2011).

Je komplexer ein theoretisches Konstrukt und je wichtiger seine genaue Messung für die Studie ist, umso eher wird man zu seiner Operationalisierung statt auf einen **Einzelindikator** (► Abschn. 8.5) auf ein **Messinstrument mit mehreren Indikatoren** zurückgreifen. Sehr verbreitet zur Operationalisierung abhängiger Variablen in der empirischen Sozialforschung sind die **psychometrische Skala** (► Abschn. 8.6) sowie der **Index** (► Abschn. 8.7). Sie messen Qualitäten sowie mehr oder minder fein abgestuft die Intensitäten von Merkmalsausprägungen.

8.3.3 Operationalisierung von unabhängigen Variablen

Während abhängige Variablen oft stetige Merkmale sind, die mit möglichst vielen Abstufungen gemessen werden, handelt es sich bei unabhängigen Variablen meistens um diskrete Variablen mit wenigen Ausprägungen. Die unabhängige Variable bzw. der Ursachenfaktor fungiert in experimentellen und quasi-experimentellen Studien (► Abschn. 7.6) als **Gruppierungsvariable**, wobei der Vergleich von zwei bis ca. zehn Gruppen typisch ist. Bei nicht-experimentellen Studien werden diese Gruppen durch eine vorgefundene Variable gebildet, die z. B. mittels Beobachtung oder Befragung operationalisiert wird (z. B. Vergleiche zwischen Altersgruppen, Geschlechtern, Nationalitäten auf der Basis entsprechender soziodemografischer Angaben im Fragebogen).

Die Operationalisierung von experimentellen und quasi-experimentellen unabhängigen Variablen läuft oft auf die Produktion von unterschiedlichem **Stimulusmaterial** oder die Konzeption von unterschiedlichen **Behandlungsformen** bzw. Interventionen oder Treatments hinaus. Soll etwa der Grad der Gewalthaltigkeit von Computerspielen als experimentelle unabhängige Variable variiert werden, um den Effekt auf die abhängige Variable

Aggressivität zu erfassen, so müssen mehrere Varianten eines Computerspiels mit unterschiedlicher Gewalthaltigkeit a) gezielt ausgewählt oder b) selbst produziert werden. Dabei kommt es darauf an, dass sich die Spiele möglichst nur in ihrer Gewalthaltigkeit und nicht in anderen Merkmalen unterscheiden. Würde man im Experiment z. B. die Aggressivität nach einem Denkspiel mit der Aggressivität nach dem Spielen eines Ego-Shooters vergleichen und in der Shooter-Gruppe tatsächlich erhöhte Aggressionswerte messen können, so wäre damit nicht belegt, dass es sich um einen Effekt der Gewalthaltigkeit des Spiels handelt. Denn Denkspiel und Shooter unterscheiden sich auch in vielen anderen Aspekten – von der Hintergrundmusik über die Farbgebung bis zur Spieleraktivität – deutlich voneinander. Die verschiedenen Untersuchungsbedingungen einschließlich der Stimulusmaterialien müssen idealerweise so gestaltet sein, dass sie sich exakt nur hinsichtlich der unabhängigen Variable voneinander unterscheiden und **alle anderen Aspekte der Untersuchungsbedingungen gleich** sind.

Zu beachten ist zudem, dass die **Dosierung der unabhängigen Variable** maßgeblich darüber entscheidet, welche **Effektstärke** sich in einer Studie zeigt (z. B. Messung der Aggressivität nach 10 Minuten, nach 3 Stunden oder nach 6 Monaten Nutzungszeit eines gewalthaltigen Medienangebots; ► [Abschn. 14.2](#)).

Eine theoretisch fundierte Operationalisierung der unabhängigen Variablen ist in der Experimentalforschung (► [Abschn. 7.6](#)) eine besondere Herausforderung und gleichzeitig entscheidende Voraussetzung für die Aussagekraft eines Experiments oder Quasi-Experiments.

8.3.4 Fehlinterpretation von Operationalisierungen

Die Auswahl der Indikatoren und die Wahl des Messinstrumentes sind bei **latenten Merkmalen** erklärungsbedürftig, weil eine Brücke zwischen den beobachtbaren Sachverhalten einerseits und der theoretischen Konzeptualisierung andererseits geschlagen werden muss. Bei **manifesten Variablen**, die praktisch unmittelbar als beobachtbare Indikatoren vorliegen, besteht kein besonderes konzeptuelles Überbrückungsproblem. Dennoch müssen strenggenommen auch manifeste Variablen operationalisiert werden. So etwa, wenn zur Erfassung von Alter, Geschlecht oder Wohnort entsprechende Fragen in einem Fragebogen gestellt werden (zur Messung soziodemografischer Variablen ► [Abschn. 8.5.1](#)).

Indem die Operationalisierung angibt, über welche Indikatoren und mit welchem standardisierten Messinstrument (z. B. einem standardisierten Fragebogen oder psychologischen Test) ein theoretisches Konstrukt em-

pirisch zu erfassen ist, wird dieses greifbarer. Dabei ist im Auge zu behalten, dass die beobachteten Merkmale keine voraussetzungslosen Tatsachen, sondern immer Ergebnis eines **theoretischen Konstruktionsprozesses** sind. Der Umstand, dass eine Person auf einem Messinstrument für „Internetsucht“ eine hohe Punktzahl erreicht, bedeutet nicht, dass die Person tatsächlich internetsüchtig „ist“. Es bedeutet, dass ihr Verhalten und Erleben auf der Basis bestimmter theoretischer Vorannahmen mit dem Konzept der Sucht beschrieben und erklärt wird. Internetsucht als reale Tatsache – anstatt als theoretische Konstruktion – aufzufassen, käme einer unzulässigen Verdinglichung bzw. **Reifizierung** („reification“) gleich. Eine andere Theorie könnte dasselbe Verhalten nicht als „Sucht“, sondern als „Zwang“ oder auch als „Gewohnheit“ auffassen, woraus sich dann andere Schlussfolgerungen hinsichtlich Entstehung oder Behandlung ergeben würden. Empirische Forschung, die gemessene Variablen als Tatsachen auffasst, mündet in einen **naiven Empirismus bzw. Positivismus**.

Deswegen ist die **theoretische Konstruiertheit aller wissenschaftlichen Messungen** bei der Diskussion von empirischen Forschungsprozessen und ihren Ergebnissen stets zu berücksichtigen. Dies wird im quantitativen Paradigma der empirischen Sozialforschung im Rahmen der Wissenschaftstheorie des Kritischen Rationalismus ausdrücklich betont (► [Abschn. 2.2.3](#)). Theoretische Konzepte zu operationalisieren läuft somit keineswegs auf ein datengläubiges „empiristisches“ oder „positivistisches“ Vorgehen hinaus, vielmehr verlangt eine seriöse wissenschaftliche Operationalisierung transparente und fundierte theoretische Argumente sowohl bei der Auswahl und Konstruktion von Indikatoren und Messinstrumenten als auch bei der Interpretation der so gewonnenen quantitativen Daten.

8.4 Messung und die vier Skalenniveaus

Wurden im Zuge der Operationalisierung für ein latentes Merkmal die manifesten Indikatorvariablen ausgewählt und die Art des Messinstruments und damit auch der Datenerhebungsmethode festgelegt (z. B. Beobachtungsschema, standardisierter Fragebogen, psychologischer Test), so steht noch die Messung im engeren Sinne an, d. h., die aussagekräftige Zuordnung von numerischen Messwerten zu den beobachteten Ausprägungen der Untersuchungseinheiten auf den einzelnen Indikatorvariablen. In Abhängigkeit von der Art der Messung unterscheiden wir vier **verschiedene Messniveaus bzw. Skalenniveaus**. Je höher das Skalenniveau der Messung, umso informationshaltiger sind die erzeugten Messwerte und umso vielfältiger die Möglichkeiten der statistischen