
Konzeptspezifikation und Messung

Rinaldo Kühne

Abstract Die quantitative Untersuchung sozialer Phänomene setzt voraus, dass die relevanten theoretischen Konzepte quantifiziert bzw. gemessen werden können. Der vorliegende Beitrag widmet sich diesem Prozess der Verknüpfung von Theorie und Empirie. Als Erstes wird der idealtypische Ablauf quantitativer Studien skizziert. Daraufhin wird erörtert, wie durch die Konzeptspezifikation und Operationalisierung Indikatoren für theoretische Konzepte generiert werden. Um zu erklären, wie Ausprägungen von Indikatoren Zahlenwerte zugewiesen werden, werden die Grundzüge der Messtheorie dargelegt und verschiedene Verfahren zur Konstruktion von Messinstrumenten dargestellt. Geschlossen wird mit einer Darstellung der zentralen Gütekriterien von Messungen und einem Ausblick auf aktuelle methodische Entwicklungen.

Schlagwörter Quantitativer Forschungsprozess, Konzeptspezifikation, Operationalisierung, Messen, Skalierungsverfahren, Skalenniveau, klassische Testtheorie, Reliabilität, Validität

1 Der quantitative Forschungsprozess

Die empirische Sozialforschung kennt mehrere typische Arbeitsschritte, die bei der Durchführung von Studien vollzogen werden. Übersichtswerke zu Methoden der empirischen Sozialforschung nennen in der Regel sechs Schritte: die Auswahl des Forschungsproblems, die Theoriearbeit, die Konzeptspezifikation und Operationalisierung, die Konzeption der Untersuchung, die Datenerfassung und -auswertung sowie die Ergebnisdarstellung und -interpretation (Atteslander 2000: 21–70; Diekmann 1999: 162–173; Friedrichs 1990: 50–59; Schnell, Hill & Esser 1999: 7–14). Analoge Ablaufmodelle finden sich in kommunikationswissenschaftlichen Methodenabhandlungen (Brosius, Koschel & Haas 2009: 42–46) – insbesondere auch in Übersichten zur Methode der Inhaltsanalyse (Früh 2007: 77–141; Merten 1995: 314–332; Rössler 2005: 34–49).

Am Anfang des Forschungsprozesses stehen die Wahl eines Forschungsproblems bzw. Themas und die Formulierung einer Forschungsfrage, die beantwortet werden soll. Der Anlass für eine Studie kann ein gesellschaftliches Problem (z. B. eine Wissenslücke zwischen Personen unterschiedlicher sozioökonomischer Schichten) oder ein Problem der Theoriebildung sein (z. B. die mangelnde Erklärungskraft eines theoretischen Modells). Zudem kann eine Auftraggeberin oder ein Auftraggeber die Untersuchung eines Problems veranlassen (Atteslander 2000: 63–65; Friedrichs 1990: 50–52; Schnell et al. 1999: 7). Geklärt wird also zunächst, welches Problem erforscht werden soll und weshalb dieses Problem relevant ist. In einem zweiten Schritt findet die theoretische Auseinandersetzung mit dem Problem statt. In einer ausführlichen Literaturrecherche werden empirische Befunde und theoretische Erklärungsansätze zusammengetragen. Die bestehenden Wissensbestände werden hinsichtlich ihrer Nützlichkeit für die Beantwortung der Forschungsfrage evaluiert. Man kann auf bereits ausgearbeitete Theorien zurückgreifen oder bei mangelnder Eignung bisheriger Ansätze neue Erklärungen erarbeiten (Brosius et al. 2009: 44; Schnell et al. 1999: 9–10). Aus der Theoriearbeit resultieren Hypothesen zur Forschungsfrage, die anhand von Daten überprüft werden sollen. An die Theoriearbeit schließt die Konzeptspezifikation und Operationalisierung an. Hier werden zunächst die zentralen Konzepte bzw. Variablen, über welche Hypothesen formuliert wurden, definiert. Anschließend wird auf Basis der Definitionen die Operationalisierung der theoretischen Konzepte vorgenommen. Bei der Operationalisierung werden theoretische Konzepte, die nicht direkt beobachtbar sind, „messbar“ gemacht, indem ihnen beobachtbare Indikatoren zugewiesen werden (Bortz & Döring 2006: 62–65; Kromrey 1998: 165–195; Weaver 2008). In einem engen Zusammenhang mit der Operationalisierung steht die weitere Konzeption der Untersuchung. Die Forscherin oder der Forscher müssen entscheiden, welche Methode (oder welche Methoden) für die Datenerhebung verwendet wird, wie die Datenerhebung im Detail ausgestaltet wird und welche Untersuchungseinheiten berücksichtigt werden (Brosius et al. 2009: 44–46; Diekmann 1999: 168–170). Anschließend werden die Daten erhoben und ausgewertet. Zuletzt werden die Ergebnisse dargestellt und zur Beurteilung der Hypothesen bzw. zur Beantwortung der Forschungsfrage herangezogen. An dieser Stelle kann der Forschungsprozess wieder auf der ersten Stufe beginnen: Anhand der Ergebnisse werden Rückschlüsse auf das soziale bzw. theoretische Problem gezogen, was neue Fragestellungen aufwerfen und weitere Forschung stimulieren kann.

Sowohl qualitative als auch quantitative Studien können diesem idealtypischen Ablauf folgen. Allerdings bestehen auch bedeutende forschungspraktische Unterschiede zwischen den beiden Herangehensweisen, die aus unterschiedlichen Wissenschaftsverständnissen und Erkenntnisinteressen resultieren. Quantitative Ansätze streben nach der Erklärung von sozialen Phänomenen bzw. dem Erkennen von Mustern und Gesetzmäßigkeiten. Qualitative Ansätze versuchen hingegen soziale Probleme zu verstehen bzw. die subjektive Perspektive und die Beweggründe von Individuen nachzuvollzie-

hen.¹ Der quantitative Ansatz zeichnet sich in der Forschungspraxis durch die Quantifizierung bzw. Messung von Realitätsausschnitten aus und resultiert in der statistischen Auswertung der aus diesem Vorgang resultierenden Daten (Bortz & Döring 2006: 296–302; Brosius et al. 2009: 19–21). Entsprechend besitzen die Konzeptspezifikation und Operationalisierung *vor* der Datenerhebung einen besonderen Stellenwert: Eine präzise Definition und Operationalisierung der relevanten Konzepte ist Voraussetzung für deren Quantifizierung. Zudem spielt in quantitativen Ansätzen die *ex ante* Formulierung von Hypothesen eine wichtige Rolle, da meist die Erklärung eines Realitätsausschnitts angestrebt wird (Friedrichs 1990: 50, 62–73; Kromrey 1998: 33–58; Opp 1999: 45–101). Über die Formulierung von Hypothesen und deren quantitative Auswertung kann eruiert werden, ob bestimmte theoretische Erklärungen plausibel sind.

Der qualitative Ansatz arbeitet mit Verbalisierungen von subjektiven Erfahrungen (also mit nichtnumerischem Material), die interpretativ ausgewertet werden, um das soziale Phänomen zu verstehen (Bortz & Döring 2006: 296–302; Flick, von Kardorff & Steinke 2005; Scholl 2008). Von der *ex ante* Formulierung von Hypothesen und der abschließenden Operationalisierung von Konzepten wird häufig Abstand genommen. Vielmehr sollen Operationalisierungen im Forschungsprozess durch die Interaktion mit den Betroffenen angepasst werden. Ebenso werden Hypothesen nicht im Vorhinein festgeschrieben, um dem Forschungsproblem mit einer größtmöglichen Offenheit zu begegnen. Hypothesen können aber das Resultat eines interaktiven Forschungsprozesses sein (Flick et al. 2005; Meinefeld 2005; Scholl 2008). Der idealtypische quantitative Forschungsprozess fokussiert also die deduktive Überprüfung von Theorien, die mittels Quantifizierung der relevanten Variablen und Variablenzusammenhänge vollzogen wird. Idealtypische qualitative Studien basieren auf einem induktiven Vorgehen, in welchem mittels interpretativer Verfahren ein Verständnis eines sozialen Problems angestrebt wird. Dieser Beitrag fokussiert den quantitativen Forschungsprozess.

2 Konzeptspezifikation und Operationalisierung

Sozialwissenschaftliche Forschung operiert oft mit Begriffen und Konzepten, die Bestandteil des alltäglichen, nichtwissenschaftlichen Sprachgebrauchs sind. Diese alltagssprachlichen Begriffe sind oft unpräzise und es existiert kein einheitliches Begriffsverständnis (Friedrichs 1990: 73–74; Weaver 2008). Zudem sind sozialwissenschaftliche Begriffe oft abstrakt und beziehen sich nicht auf unmittelbar in der Realität beobachtbare Objekte und Ereignisse (Kromrey 1998: 145). Um eine sinnvolle Quantifizierung bzw. Messung von Begriffen und Konzepten zu ermöglichen – und letzten Endes Hypo-

1 Eine ausführliche Diskussion der unterschiedlichen Wissenschaftsverständnisse und Erkenntnisinteressen quantitativer und qualitativer Ansätze ist bei Opp (1999: 45–101) zu finden.

thesen mit quantitativen Methoden zu testen – müssen diese Begriffe definiert werden, wobei die Definitionen bestimmten Anforderungen genügen müssen.

Eine Definition kann ganz allgemein als „Festlegung des Sprachgebrauchs eines Begriffs“ (Friedrichs 1990: 74) bezeichnet werden. Ein Begriff besteht aus einem Zeichen, Designata und semantischen Regeln (Friedrichs 1990: 74; Kromrey 1998: 145–149; Opp 1999: 102–104). Ein Zeichen kann ein Wort sein wie ‚Themenrelevanz‘ oder ‚Politikverdrossenheit‘. Designata bezeichnen reale Sachverhalte oder Wörter, deren Bedeutung bekannt ist. Über semantische Regeln werden Zeichen und Designata miteinander verbunden bzw. eine Bedeutungszuweisung vorgenommen. Die Definition eines Zeichens geschieht dann auf folgende Weise: Diejenigen Designata, die einem zu definierenden Zeichen zugeordnet werden, werden mit Zeichen beschrieben, deren Bedeutung als bekannt vorausgesetzt wird (Opp 1999: 102–104).

Die Anforderungen an Begriffsdefinitionen sind *Präzision*, *Eindeutigkeit* und *Adäquatheit* (Clark & Watson 1995; Friedrichs 1990: 74; Opp 1999: 127–132; Westermann 2000: 106–110). Bei unpräzisen Begriffsdefinitionen ist es nicht klar, ob der Begriff auf ein Objekt oder Ereignis angewendet werden kann. Das heißt, auf Basis eines unpräzisen Begriffs ist es nicht möglich zu entscheiden, ob ein bestimmtes Objekt oder Ereignis zur durch den Begriff designierten Menge von Objekten und Ereignissen gehört. Ist ein Begriff präzise, dann ist eine klare Zuordnung von Objekten und Ereignissen möglich. Ein Begriff ist eindeutig (oder auch konsistent; Friedrichs 1990: 74), wenn verschiedene Forscherinnen und Forscher den Begriff nicht in unterschiedlicher Bedeutung verwenden (Opp 1999: 129). Das heißt, auf Basis eines eindeutigen Begriffs gelangen alle Forscherinnen und Forscher zum gleichen Schluss, ob spezifische Ereignisse oder Objekte zur designierten Menge von Ereignissen oder Objekten gehören. Eine Definition sollte darüber hinaus adäquat sein, also nicht zu eng oder zu weit (Westermann 2000: 106). Bei einer zu engen Definition existieren kaum Fälle, die der durch den Begriff designierten Mengen zugeordnet werden können. Bei einer zu weiten Definition ist die designierte Menge hingegen zu groß. Eine adäquate Definition deckt hingegen gerade die relevanten Objekte und Ereignisse ab.

Verschiedene Möglichkeiten bestehen, um Begriffe zu definieren. Eine *Realdefinition* ist die Beschreibung des Wesens oder der Natur eines Begriffs (Friedrichs 1990: 75–76; Opp 1999: 109–110). Das heißt, es werden Aussagen über die Eigenschaften eines Gegenstands gemacht, die bezüglich des Gegenstandes als wesentlich erachtet werden (Kromrey 1998: 160–163). Allerdings werden Realdefinitionen in der sozialwissenschaftlichen Forschung in der Regel mit der Begründung abgelehnt, dass Wesensbestimmungen nicht möglich sind (Brosius et al. 2009: 36; Kromrey 1998: 162; Opp 1999: 109–110). Allenfalls können Realdefinitionen als Behauptungen über die Beschaffenheit eines Phänomens verstanden werden, wobei sie dann denselben Status wie Hypothesen haben und einer empirischen Überprüfung zugeführt werden müssen (Kromrey 1998: 161).

In der Regel operieren die Sozialwissenschaften mit *Nominaldefinitionen*. Hier wird der zu definierende Begriff mit einem oder mehreren bereits definierten Begriffen

gleichgesetzt. Eine Nominaldefinition hat also zwei Komponenten: einen oder mehrere Begriffe, deren Bedeutung vorausgesetzt wird (Definens), und den Begriff, der als synonym mit dem bekannten Begriff bzw. den bekannten Begriffen verstanden wird (Definiendum) (Bortz & Döring 2006: 60–61; Kromrey 1998: 149–160; Opp 1999: 104–107). Die Bedeutung des Definiendums ergibt sich also vollständig aus dem Definens. Wirth (1999: 55) schlägt zum Beispiel folgende Definition von Glaubwürdigkeit vor: „Glaubwürdigkeit kann als prinzipielle Bereitschaft verstanden werden, Botschaften eines bestimmten Objektes als zutreffend zu akzeptieren und bis zu einem gewissen Grad in das eigene Meinungs- und Einstellungsspektrum zu übernehmen.“ Die Bedeutung des Begriffs Glaubwürdigkeit ergibt sich hier vollständig aus der Kombination der einzelnen Bestandteile des Definens (Bereitschaft, Meinungs- und Einstellungsspektrum, etc.). Die Definition verdeutlicht auch, dass sozialwissenschaftliche Begriffe oft eine hohe Komplexität aufweisen. Einer Nominaldefinition muss daher eine ausführliche semantische Analyse² der relevanten Begriffe vorausgehen, um unterschiedliche Bedeutungsfacetten zu eruieren und definitorisch abzudecken (Kromrey 1998: 130–145). Zu beachten ist allerdings, dass Nominaldefinitionen lediglich Konventionen für die Verwendung von Begriffen sind, keinen empirischen Informationsgehalt haben und folglich nicht wahr oder falsch sein können (Friedrichs 1990: 76; Kromrey 1998: 149–150). Nominaldefinitionen können verschiedene Fehler aufweisen (Opp 1999: 107–108; Westermann 2000: 106–108; Schnell et al. 1999: 48–51). Zu vermeiden sind definitorische Zirkel, in welchen zu definierende Ausdrücke zu ihrer eigenen Definition werden, unpräzise und mehrdeutige Ausdrücke im Definens sowie Definitionen durch Beispiele (da eine solche Definition nicht sagt, wie mit anderen Sachverhalten als den erwähnten zu verfahren ist).

Nominal definierte Begriffe der Sozialwissenschaften weisen oft einen hohen Abstraktionsgrad auf. Um die Begriffe messbar zu machen, ist es deshalb notwendig, die Begriffe in beobachtbare Ereignisse zu übersetzen bzw. diese operational zu definieren. Die *operationale Definition* stellt demnach einen Arbeitsschritt dar, der an die nominale Begriffsdefinition anschließt.³ Hierbei werden einem Begriff präzise und beobachtbare Designate, die auch Indikatoren genannt werden, zugeordnet (Atteslander 2000: 50–53; Opp 1999: 118–127; Weaver 2008). Besozzi und Zehnpfennig sprechen in diesem Zusammenhang von der Notwendigkeit, „dass man latente (d.h., nicht direkt beobachtbare, theoretische) Begriffe mit Hilfe bestimmter Transformationsregeln auf die Beobachtungsebene abbildet.“ (1976: 14). Operationalisierungen sind somit nichts anderes als Anweisungen für die Messung von Begriffen und Konzepten, die sich auf direkt beobachtbare Sachverhalte bzw. Indikatoren beziehen (Schnell et al. 1999: 10). In Abhängig-

2 Eine ausführliche Diskussion der einzelnen Schritte einer semantischen Analyse ist bei Kromrey (1998: 130–141) zu finden.

3 Im Folgenden werden die Ausdrücke ‚Operationalisierung‘ und ‚operationale Definition‘ synonym verwendet. Alternativ wird die operationale Definition auch als genuine Vorgehensweise für die Definition von Begriffen verstanden. Diese Vorgehensweise ist aber umstritten (Bortz & Döring 2006: 62–64; Opp 1999: 120–123).

keit von der spezifischen Datenerhebungsmethode (Befragung, Inhaltsanalyse oder Beobachtung) sind unterschiedliche Indikatoren denkbar. Die bedeutendsten Formen von Indikatoren in der Kommunikationswissenschaft sind Items in Fragebögen sowie inhaltsanalytische Kategorien in Codebüchern, wobei die klassische Beobachtung, physiologische Messungen und Blickaufzeichnungen (vgl. die Beiträge von Gehrau & Schulze, Fahr & Hofer sowie Blake im vorliegenden Band) an Bedeutung gewinnen. Das zentrale Problem der Operationalisierung – auch Korrespondenzproblem genannt (Schnell et al. 1999: 71–78) – besteht nun darin, wie einem theoretischen Konzept überhaupt Indikatoren zugewiesen werden können (Opp 1999: 118–127; Kromrey 1998: 88–91; Steyer & Eid 2001: 2).

Besozzi und Zehnpfennig (1976) unterscheiden drei Vorgehensweisen, um dem Korrespondenzproblem zu begegnen, die jeweils mit spezifischen Vorstellungen über die Relation zwischen Konzept und Indikatoren einhergehen. Die *operationalistische* Lösung geht davon aus, dass Indikatoren Elemente der Definition eines Konzepts sind. Ein Konzept wird also erst durch die Messanweisung definiert. Die Problematik der Zuordnung von Indikatoren fällt folglich weg, da die Indikatoren erst das Konzept konstituieren. Die *typologisch-induktive* Lösung geht von Konzepten als Dispositionen von Indikatoren aus. Theoretische Konzepte sind demnach latente Variablen, welche die Kovariation der beobachtbaren Indikatoren erklären. Die Zuordnung von Indikatoren geschieht in einem mehrstufigen Verfahren, das deduktive und induktive Elemente aufweist. So wird anstelle einer präzisen Nominaldefinition eine vorläufige Definition verwendet, welche die Bestimmung von Subdimensionen des Konzepts erlaubt. Anschließend wird eine Menge von Indikatoren mittels strukturentdeckender Verfahren (Analyse latenter Klassen, explorative Faktorenanalyse) analysiert und Indikatoren für die einzelnen Dimensionen selektiert.

Sowohl die operationalistische als auch die typologisch-induktive Lösung wurden verschiedentlich kritisiert. So ist die operationalistische Vorgehensweise zwar geeignet, wenn die Bedeutung eines Konzepts a priori gegeben ist, allerdings dürfte dies bei sozialwissenschaftlichen Konzepten selten der Fall sein. Zudem vernachlässigen beide Vorgehensweisen den Theoriebezug, was zur Spezifikation theoretisch irrelevanter Konzepte führen kann (Besozzi & Zehnpfennig 1976; Schnell et al. 1999: 125–127). Der verbreitetste Ansatz ist deshalb die *kausal-analytische* Lösung. Diese geht davon aus, dass Indikatoren als empirische Wirkungen von abstrakten Konzepten verstanden werden können (Besozzi & Zehnpfennig 1976), und folgt so einem vergleichbaren Verständnis wie der typologisch-induktive Ansatz. Allerdings betont der kausal-analytische Ansatz die Notwendigkeit, Indikatoren für ein Konzept deduktiv abzuleiten. Indikatoren sollen also theoriegeleitet formuliert werden, bevor empirische Analysen einsetzen.

Die kausal-analytische Perspektive basiert auf der Annahme, dass für die Messung eines Konzepts ein Universum von Indikatoren vorliegt (Besozzi & Zehnpfennig 1976; Kromrey 1998: 180–181; Schnell et al. 1999: 125–130) und dass einzelne Indikatoren unterschiedliche Facetten des Konzepts abbilden. Aus der Annahme eines Indikatorenuni-

versums folgt, dass mehrere Indikatoren für die Erfassung eines Konzepts verwendet werden sollen – da sich einzelne Indikatoren ja auf bestimmte Facetten des Konzepts beziehen –, um das Konzept breit abzubilden und indikatorspezifischen Messungen vorzubeugen (Besozzi & Zehnpfennig 1976; Friedrichs 1990: 165). Zudem können mit Messungen, die auf mehreren Indikatoren basieren, zufällige Messfehler ausgeglichen werden (Kromrey 1998: 173–174; Schnell et al. 1999: 129–130). Eine weitere Annahme ist, dass das Indikatorenuniversum homogen ist. Aus dieser Annahme folgt, dass die Teilmenge der Indikatoren, die für die Erfassung eines Konzepts verwendet wird, austauschbar ist (Besozzi & Zehnpfennig 1976; Schnell et al. 1999: 128).⁴ Das heißt, für jede homogene Dimension eines Konzepts sollten unendlich viele und austauschbare Indikatoren existieren.

Die erörterten konzeptionellen Überlegungen, die der kausal-analytischen Lösung zugrunde liegen, stellen die Basis für die forschungspraktische Formulierung bzw. Selektion von Indikatoren dar. Verschiedene Empfehlungen wurden für diesen Schritt ausgesprochen (Clark & Watson 1995; Haynes, Richard & Kubany 1995; Opp 1999: 123–127). Zudem existieren Anleitungen für die Indikatorenbildung bei spezifischen Erhebungsmethoden wie der Inhaltsanalyse (Früh 2007; Rössler 2005) oder der Befragung (Holm 1982; vgl. auch den Beitrag von Hartmann & Reinecke im vorliegenden Band). Grundsätzlich sollten Indikatoren so gewählt werden, dass sie alle Bedeutungsaspekte des Konzepts abdecken (Clark & Watson 1995). Für alle theoretisch identifizierten Dimensionen des Konzepts sollen mehrere Indikatoren selektiert werden. Zudem soll die Anzahl an Indikatoren pro Dimension ausreichend groß sein. Die genaue Anzahl Indikatoren pro Begriffsaspekt sollte sich an der Relevanz der jeweiligen Subdimension für das Gesamtkonzept orientieren (Haynes et al. 1995). Ob einzelne Indikatoren tatsächlich das zugrundeliegende Konzept abbilden, muss durch eine ausführliche Auseinandersetzung und Diskussion der Indikatoren vor dem Hintergrund der gewählten Nominaldefinition entschieden werden (Haynes et al. 1995; Opp 1999: 123–127; Pett, Lackey & Sullivan 2003: 13–50). Letztlich handelt es sich hierbei um Fragen der Validität von Messungen, für deren Überprüfung auch empirische Verfahren existieren, welche in Abschnitt 5 diskutiert werden (z. B. Moosbrugger & Kelava 2007).

4 Kausal-analytisch spezifizierte Konzepte werden auch reflektive Konzepte genannt. Hiervon zu unterscheiden sind formative Konzepte, die durch die Indikatoren verursacht werden. Bei formativen Konzepten ist der willkürliche Austausch von Indikatoren nicht möglich, da die Indikatoren das Konzept konstituieren (Ingenhoff & Sommer 2010).

3 Grundlagen des Messens

Aus der Konzeptspezifikation und Operationalisierung resultieren Indikatoren, welche die theoretischen Konzepte repräsentieren. Diese Indikatoren können unterschiedliche Ausprägungen annehmen und stellen somit (manifeste) Variablen dar (Atteslander 2000: 53; Schnell et al. 1999: 125). Beim Messen werden den einzelnen Ausprägungen der Variablen systematisch bzw. nach bestimmten Regeln Zahlenwerte zugewiesen (Friedrich 1973; Früh 2007: 28–29; Orth 1974: 18–19). Die Zuweisung erfolgt nach bestimmten Regeln, um zu gewährleisten, dass das resultierende numerische Relativ das empirische Relativ strukturerhaltend bzw. *homomorph* abbildet. Das heißt, die empirischen Relationen zwischen Objekten bezüglich relevanter Merkmalsausprägungen müssen sich in den vergebenen Zahlenwerten widerspiegeln. Zum Beispiel sollte bei der Messung der täglichen Fernsehnutzungsdauer Individuen mit einer längeren Nutzungsdauer ein höherer Zahlenwert zugewiesen werden als Individuen mit einer kurzen Nutzungsdauer, um die empirische Mehr-Weniger-Relation zu erhalten. Eine Skala ist eine homomorphe Abbildung eines empirischen Relativs in ein numerisches Relativ. Die numerischen Werte stellen die Skalen- oder Messwerte dar (Bortz 2005: 17; Orth 1974: 22; Schnell et al. 1999: 132–133). Die Herausforderung zu beweisen, dass für ein empirisches Relativ eine homomorphe Abbildung unter bestimmten Bedingungen existiert, wird Repräsentationsproblem genannt (für die folgenden Ausführungen vgl. Bortz 2005: 16–23; Diekmann 1999: 244–249; Orth 1974: 22). Die Lösung des Problems ist mittels Repräsentationstheoremen möglich, die aussagen, dass (mindestens) eine homomorphe Abbildung bzw. Skala existiert, sofern ein empirisches Relativ bestimmte Bedingungen erfüllt. In der Regel gibt es für ein Repräsentationstheorem mehrere unterschiedliche Skalen. Das heißt, wenn ein empirisches Relativ bestimmte Eigenschaften besitzt, existieren unterschiedliche Möglichkeiten, dieses homomorph abzubilden. Eine Menge von Skalen zu einem Repräsentationstheorem wird charakterisiert durch die zulässigen Transformationen, durch welche die Skalen ineinander überführt werden können. Diejenigen Transformationen sind dabei zulässig, welche die Eigenschaft der Skalen, das Repräsentationstheorem zu erfüllen, nicht tangieren. In Abhängigkeit der zulässigen Transformationen werden unterschiedliche Skalenniveaus unterschieden. Die zulässigen Transformationen sind wiederum von den Eigenschaften des empirischen Relativs abhängig (Bortz 2005: 18–23; Orth 1974: 24–25; Steyer & Eid 2001). Je weniger Transformationen zugelassen sind, desto eindeutiger ist eine Skala. Ausführliche Abhandlungen zu den Skalenniveaus finden sich bei Bortz (2005: 18–23), Diekmann (1999: 249–260), Orth (1974: 25–28) und Steyer und Eid (2001). Im Folgenden werden die unterschiedlichen Skalenarten beschrieben. Es sei angemerkt, dass die einzelnen Skalentypen jeweils eine Klasse von Skalen repräsentieren und innerhalb jeder Klasse verschiedene Skalierungsverfahren existieren (vgl. Abschnitt 4).

Nominalskalen sind die am wenigsten eindeutigen Skalen und setzen ein empirisches Relativ voraus, für welches die Äquivalenzrelation gilt. Das heißt, dass im empirischen

Relativ die Gleichheit und Ungleichheit von Objekten festgestellt werden kann.⁵ Eine Nominalskala klassifiziert Objekte anhand ihrer Gleichheit und Ungleichheit bezüglich einer Merkmalsausprägung: Objekten mit gleicher Merkmalsausprägung werden gleiche Zahlenwerte zugewiesen und Objekten mit unterschiedlichen Ausprägungen unterschiedliche Zahlenwerte. Die spezifischen Zahlenwerte, die zugewiesen werden, sind dabei nicht relevant. Somit sind alle eindeutigen Transformationen der Skala zugelassen, bei welchen jedem Ausgangswert ein spezifischer neuer Wert zugewiesen wird. Beispiele für Nominalskalen sind die Variablen ‚Geschlecht‘, ‚Religionszugehörigkeit‘ oder ‚Muttersprache‘.

Ordinalskalen setzen voraus, dass für das empirische Relativ eine schwache Ordnungsrelation gilt. Für jedes Paar von Objekten muss feststellbar sein, bei welchem Objekt das Merkmal stärker ausgeprägt ist oder ob die Objekte äquivalent sind, und Transitivität muss gegeben sein: Wenn Objekt a eine stärkere Ausprägung besitzt als Objekt b und Objekt b eine stärkere Ausprägung als Objekt c, dann muss auch gelten, dass Objekt a eine stärkere Ausprägung besitzt als Objekt c. Eine Ordinalskala weist einem Objekt a, dessen Merkmalsausprägung mindestens so groß ist wie diejenige eines Objekts b, einen Zahlenwert zu, der mindestens so groß ist wie der Zahlenwert von Objekt b. Eine Ordinalskala ermöglicht es also, alle Objekte in eine Rangordnung zu bringen. Erlaubt sind alle monotonen Transformationen, welche die Rangordnung der Skalenwerte nicht tangieren. Eine Rangordnung der am häufigsten genutzten Medienangebote stellt ein Beispiel für eine Ordinalskala dar.

Intervallskalen setzen ein empirisches Relativ voraus, dass eine algebraische Differenzstruktur aufweist (Bortz 2005: 19–21) und somit für jedes Paar von Objekten die Größe des Unterschieds bezüglich der Merkmalsausprägung feststellbar ist. Die Zuweisung von Zahlen funktioniert so, „dass die Rangordnung der Zahlendifferenzen zwischen je zwei Objekten der Rangordnung der Merkmalsunterschiede zwischen zwei Objekten entspricht“ (Bortz 2005: 21). Gleich großen Merkmalsunterschieden entsprechen also gleich große Differenzen der zugewiesenen Zahlenwerte. Daraus folgt, dass Objekte bezüglich eines Merkmals nicht nur in eine Rangordnung gebracht werden können, sondern dass auch die exakten Abstände der Merkmalsausprägungen angegeben werden können. Für Intervallskalen sind alle linearen Transformationen $y = b \cdot x + a$ (mit $b > 0$) zulässig. Der Nullpunkt einer Intervallskala ist demnach willkürlich gewählt und ist nicht durch Eigenschaften des empirischen Relativs determiniert. Beispiele für Intervallskalen sind die Celsius- und Fahrenheitskala für Temperaturen oder der Intelligenzquotient.

Verhältnis- oder Ratioskalen sind die eindeutigsten Skalen. Ratioskalen setzen ein empirisches Relativ mit einer extensiven Messstruktur voraus (vgl. Bortz 2005: 22): Über die Gleichheit von Differenzen muss auch die Gleichheit von Verhältnissen feststellbar

5 Für eine formale Abhandlung der erforderlichen Eigenschaften bzw. Axiome, die für empirische Relative gelten müssen, vgl. Bortz (2005), Orth (1974) und Kromrey (1998).

sein. Zum Beispiel sollten Aussagen möglich sein wie: „Die Objekte a und b haben zusammen dasselbe Gewicht wie Objekt c.“ Eine Ratioskala „ordnet den Objekten eines empirischen Relativs Zahlen zu, die so geartet sind, dass das Verhältnis zwischen je zwei Zahlen dem Verhältnis der Merkmalsausprägungen der jeweiligen Objekte entspricht“ (Bortz 2005: 22). Folglich können Objekte bezüglich einer Merkmalsausprägung in eine Reihenfolge gebracht werden und Abstände zwischen Merkmalsausprägungen sowie Verhältnisse zwischen Merkmalsausprägungen können exakt angegeben werden. Für Ratioskalen sind alle Ähnlichkeitstransformationen $y = b \cdot x$ (mit $b > 0$) zulässig. Das heißt, die Einheit der Messung kann verändert werden (z. B. Meter statt Zentimeter), aber der Nullpunkt der Skala ist fixiert. Beispiele für Ratioskalen sind physikalische Messungen wie Länge, Gewicht oder Zeit. Intervall- und Ratioskalen werden auch unter der Bezeichnung metrische Skalenniveaus zusammengefasst.

Die sozialwissenschaftliche Forschung operiert meist mit nominal bis intervallskalierten Variablen. Wie erörtert wurde, werden die Skalenarten zunehmend genauer, weshalb eine geringere Anzahl von Transformationen zulässig ist, welche die Aussagen der Skala nicht verändern. Möglichst hohe Skalenniveaus werden beim Messen angestrebt, da diese eine größere Zahl an Rechenoperationen und statistischen Auswertungen zulassen (Diekmann 1999: 249–260; Friedrichs 1990: 97–100; Kromrey 1998: 231–237).

4 Skalierungsverfahren und Indexbildung

Die beschriebenen Skalenniveaus stellen Skalenklassen dar, denen verschiedene Skalierungsverfahren und spezifische Messinstrumente zuzuordnen sind. Der Begriff Skalierungsverfahren wird hierbei in Anlehnung an Schnell et al. (1999: 173) als Methode zur Konstruktion spezifischer Messinstrumente verstanden (z. B. die Likert- oder die Guttman-Skalierung; vgl. auch Hartmann & Reinecke im vorliegenden Band). Insofern können Anweisungen zur Konstruktion von Kategoriensystemen als inhaltsanalytisches Skalierungsverfahren verstanden werden. Ein Messinstrument stellt hingegen eine Sammlung von Indikatoren inklusive der dazugehörigen Messanweisungen für die Erfassung von theoretischen Konzepten dar. Ein Messinstrument ist demnach z. B. eine Sammlung von Testfragen oder Items in einem Fragebogen inklusive Ausfüllanweisungen oder ein inhaltsanalytisches Kategoriensystem mit Anweisungen für die Codierenden und Codierer. Wie in der Literatur üblich, werden im Folgenden sowohl einzelne Items als auch Sammlungen von Items bzw. Messinstrumente als Skalen bezeichnet.

Der Skalierungsprozess bei Inhaltsanalysen ist weniger standardisiert als etwa bei der Befragung (vgl. Früh 2007). Fokussiert wird dabei die Frage, wie Kategoriensysteme zur Erfassung der relevanten Konzepte erstellt werden können, wobei deduktive und induktive Ansätze zu unterscheiden sind. Ausführliche Darstellungen des Prozesses der Kategorienbildung finden sich bei Früh (2007), Merten (1995) und Rössler (2005). Bei der Befragung stehen bei der Skalierung von theoretischen Konzepten zwei Fragen im

Handbuch standardisierte Erhebungsverfahren in der
Kommunikationswissenschaft

Möhring, W.; Schlütz, D. (Hrsg.)

2013, VIII, 389 S. 15 Abb., Hardcover

ISBN: 978-3-531-18775-4