
8.1 Grundzüge und Herkunft

Die Inhaltsanalyse ist diejenige Methode, die in der Kommunikationswissenschaft am weitesten verbreitet ist. Die Kommunikationswissenschaft ist sozusagen die Wissenschaft, die Inhaltsanalysen kultiviert und auch weiterentwickelt hat. In den anderen Fächern, wie etwa der Psychologie, spielt sie eine geringere Rolle. Wie bei anderen Begriffen, mit denen sich diese Einführung befasst, findet man auch zum Begriff „Inhaltsanalyse“ Synonyme. Zunächst ist in der deutschsprachigen Literatur die englische Fassung geläufig: *Content Analysis*. Aber auch die direkten Übersetzungen *Kontentanalyse* und *Medienresonanzanalyse* finden sich zuweilen.

Der Erste Weltkrieg war der erste Krieg, der unter massivem Einsatz von Propaganda und Gegenpropaganda betrieben wurde. Die Zersetzung des Gegners durch Propaganda und Falschmeldungen, durch alle möglichen Arten von Kommunikationsmitteln, war im Ersten Weltkrieg zum ersten Mal eine Waffe bei einer kriegerischen Auseinandersetzung. Im Zweiten Weltkrieg wurde dieses Mittel noch perfektioniert. Zum Beispiel warfen die Kriegsparteien Flugblätter im Feindesland ab, in denen auf die desolate Situation und auf die marode Führerschaft hingewiesen wurde. Im Zusammenhang mit den vermuteten Propagandamöglichkeiten der Massenmedien wurde dann in den USA die Inhaltsanalyse systematischer entwickelt. Dabei ging es in dieser Phase vor allem um die Frage, was eigentlich die Inhalte dieser Botschaften sind und wie diese Botschaften von Menschen verstanden und verarbeitet werden. Es ging also um ein ganz profanes kriegerisches Ziel, nämlich die Propaganda des eigenen Landes zu optimieren und die Propaganda des Gegners in ihrer Effektivität einzuschränken. In diesem Kontext, etwa in den 1940er-Jahren, entstand die Inhaltsanalyse bzw. deren Vorläufer. Wichtige Namen aus dem Bereich der Kommunikationsforschung sind in dem Zusammenhang Lasswell, Lazarsfeld und Berelson, einige der Gründerväter des Fachs.¹

1 Dies ist übrigens nicht der einzige Fall, in dem der Zweite Weltkrieg für die Kommunikationsforschung eine entscheidende Rolle gespielt hat. Carl Hovland hat nach dem Krieg in den sogenannten „Yale-Studien“ Experimente zur Wirkung von persuasiven Botschaften durchgeführt, um festzustel-

Neben der politischen Kommunikationsforschung, die auf die Propagandafor- schung zurückgeht und auch heute noch das wichtigste Forschungsgebiet ist, in dem diese Methode eingesetzt wird, werden Inhaltsanalysen in der Kommunikationswis- senschaft außerdem in der Gewaltforschung, der Forschung über Minderheiten so- wie den Public Relations angewandt. Auf einzelne Fragestellungen wird im Verlauf dieses Abschnitts noch näher eingegangen. Doch zunächst zu der Frage, was Inhalts- analyse eigentlich bedeutet.

8.2 Definitionen

Die erste wichtige Bestimmung dieser Methode geht über einen Vergleich: Der Un- tersuchungsgegenstand der Inhaltsanalyse sind nicht Personen wie bei der Befragung, sondern es sind in der Regel Medienprodukte. Dies können Texte aus Tageszeitun- gen sein, aber ebenso werden Rundfunksendungen, Filme, Werbespots oder Propa- gandamaterial, also sämtliche Formen von textlichen oder visuellen Botschaften, die massenmedial verbreitet werden, analysiert.

Bernard Berelson war es, der 1952 ein grundlegendes Buch vorlegte, in dem die Methode erstmals systematisch und theoretisch fundiert dargestellt wurde: „Content Analysis in Communications Research“ ist der Titel und denen, die sich einmal dezi- diert mit dieser Methode auseinandersetzen wollen, sei das Buch auch heute noch als Lektüre empfohlen. Um der Inhaltsanalyse auf die Spur zu kommen und um die Entwicklungslinien bis heute zu verstehen, lohnt sich ein Blick auf die verschiedenen Definitionen, wie sie in einschlägigen Lehrbüchern angeboten werden. Bei näherer Betrachtung wird man schnell erkennen, dass sich hinter diesen kurzen Abhandlun- gen recht unterschiedliche wissenschaftliche Auffassungen dieser Methode verber- gen. Zunächst die ursprüngliche Definition von Berelson:

- **„Content analysis is a research technique for the objective, systematic, and quanti- tative description of the manifest content of communication.“ (Berelson 1952, S. 18)**

Die beiden weit verbreiteten Lehrbücher über Inhaltsanalyse im deutschen Sprach- raum (Früh und Merten) haben diese Definition der Inhaltsanalyse etwas unter- schiedlich interpretiert und eigene Definitionen vorgelegt:

- **„Die Inhaltsanalyse ist eine empirische Methode zur systematischen, intersubjek- tiv nachvollziehbaren Beschreibung inhaltlicher und formaler Merkmale von Mit-**

len, wie die Wirkkraft einer (politischen) Botschaft optimiert werden kann. Hovland und sein Team (Hovland et al. 1953) untersuchten dabei nicht das Propagandamaterial gegen den Feind, sondern in- teressierten sich für die Propaganda innerhalb der eigenen Reihen: Was könnte man z. B. unterneh- men, um die eigenen Leute stärker zu mobilisieren?

teilungen, meist mit dem Ziel einer darauf gestützten interpretativen Inferenz auf mitteilungsexterne Sachverhalte.“ (Früh 2011, S. 27)

- „Die Inhaltsanalyse ist eine Methode zur Erhebung sozialer Wirklichkeit, bei der von Merkmalen eines manifesten Textes auf Merkmale eines nicht-manifesten Kontextes geschlossen wird.“ (Merten 1995, S. 59)

8.2.1 Manifeste oder latente Bedeutung von Texten?

Alle drei Definitionen bestimmen die Inhaltsanalyse als eine empirische Methode, mit deren Hilfe sich etwas *beschreiben* lässt: „Inhalte von Texten“, „Merkmale von Mitteilungen“, „soziale Wirklichkeit“. Der Scheideweg, an dem sich die drei Definitionen trennen, wird mit den Begriffen *manifest* und *nicht-manifest*² bestimmt. Während Früh in der Folge seiner Definition explizit die Verwendung dieses Begriffes ablehnt, da er „in der Vergangenheit oft mehr Verwirrung gestiftet als zur Klärung beigetragen“ habe (Früh 2011, S. 27), stellt Merten klar, dass die Methode (d. h. systematisches, planvolles Vorgehen) der Inhaltsanalyse dazu dienen muss, auf der Grundlage vorliegender, also manifesten, Texte soziale Wirklichkeit, d. h. nicht-manifeste Kontexte, zu erfassen bzw. zu verstehen. Sobald man also über den analysierten Text hinausgeht, führen die dabei notwendigen Inferenzen (Was hat der Kommunikator gemeint? Welche Wirkungen hat die Botschaft wohl gehabt?) zu einer Interpretation, die nicht-manifest ist, weil sie sich nicht mehr direkt aus dem Text erschließt.

Berelson erhob in seiner Definition die zentrale Forderung, dass nur die manifesten Inhalte von Texten in Kategorien messbar sind, man also sicherstellen müsse, dass Kommunikator, Leser und Forscher ein und desselben Textes tatsächlich dasselbe Verständnis von diesem Text hätten. Er spricht in diesem Zusammenhang von einem „common meeting-ground“ (Berelson 1952, S. 19) an dem sich Urheber, Leser und Forscher quasi „treffen“ müssen. Damit ist gemeint, dass diese Beteiligten dem Inhalt eine identische Bedeutung zuschreiben. Genau diese Forderung ist aus heutiger Sicht problematisch. Kann man wirklich davon ausgehen, dass Informationen derart allgemein verarbeitet werden, dass alle alles gleich verstehen? Oder haben nicht Alter, Bildung, Kulturkreis oder individuelle Lebenskonzeption zur Folge, dass eben nicht alle alles gleich verstehen? Dies gilt umso mehr, wenn man eine historische Dimension anlegt. Können wir heute noch Texte aus den 1950er-Jahren so verstehen, wie die Menschen sie damals auch verstanden haben? Manifeste (offenkundige) bzw. latente (der Möglichkeit nach vorhandene) Bedeutungen von Begriffen sind vermutlich nicht so einfach zu bestimmen, wie es Berelson versuchte. Man kann

2 Das begriffliche Gegenstück zu „manifest“ wird üblicherweise mit „latent“ bezeichnet. Warum Merten und Teipen (1991) den Begriff „nicht-manifest“ verwenden, wird im Verlauf ihrer Darstellung nicht weiter erläutert.

Textinhalte nur schwer danach unterteilen, welche Inhalte latent oder manifest im Sinne des „common meeting-ground“ sind. Selbst ganz einfache Aussagen können in einem bestimmten Kontext eine latente Bedeutung haben, die sich nur aus dieser konkreten Situation heraus ergibt. Dies gilt ganz eklatant für den Sprachstil von Subkulturen, die vermeintlich ganz eindeutigen Begriffen neue Bedeutungen beimessen. Weiß man wirklich, was gemeint ist, wenn man das Wort „Zecke“ liest? Wahrscheinlich gibt es Botschaften, die offenkundiger zu sein scheinen als andere. Vielleicht ist der Inhalt einer Börsennachricht tatsächlich manifeste als der einer Kulturkritik. Aber auch da muss man schon wissen, dass es sich um diesen Kontext handelt. Eine Dichotomie latent – manifest, ein „entweder – oder“ gibt es wohl nicht: Vielmehr scheint die semantische und syntaktische Eindeutigkeit von Texten auf einem Kontinuum zwischen Offenkundigkeit und Verborgtheit zu liegen, worauf auch Berelson schon hinweist.

Die kann an dem Beispiel verdeutlicht werden, das wir in den ersten Auflagen dieses Buches angeführt hatten. In einem bestimmten Kommunikationskontext, Anfang 1999, „wusste“ vermutlich ein großer Teil der deutschen Bevölkerung etwas mit Begriffen wie Nussecken und Guildo Horn anzufangen. Aber gibt es tatsächlich einen „common meeting-ground“? Anhänger einer klaren Unterscheidung zwischen latenten und manifesten Inhalten hätten ihre Schwierigkeiten gehabt: Nicht jeder kannte den „Meister“ und seine Vorliebe für Nussecken. Anders ausgedrückt: Konnotationen, d. h. gedankliche Verknüpfungen zwischen einem manifesten Text (Nussecke) und einer Abstraktion (Lieblingsspeise) konnte und wollte die erste Definition von Inhaltsanalyse nicht erfassen. Gleichwohl wird an diesem Beispiel deutlich, dass sich der Sinn oder die Bedeutung von Begriffen häufig erst zur Gänze über das erschließt, was man zwischen den Zeilen liest. Wäre man 1965 mit Nussecken und Guildo Horn konfrontiert worden, hätten diese Begriffe in einem völlig anderen Bedeutungskontext gestanden: „Guildo Horn? – kenne ich nicht.“ Das Gleiche gilt aber wahrscheinlich inzwischen wieder. Im Jahr 2015 dürfte ein Großteil der 20-Jährigen mit dieser Figur nicht mehr viel anfangen können. Wenn man so will, ist der „common meeting-ground“ wieder „verschwunden“. Anders ausgedrückt: Möglicherweise sind Texte ohne ihren Zeitbezug und ohne die Messung dessen, wie die Leser diese Texte verstehen bzw. verstanden haben, inhaltsanalytisch nicht zu erfassen. Der große Vorteil, den diese Methode bietet – nämlich, raum- und zeitunabhängig Texte aller Art zu erfassen und zu analysieren – ginge mit dieser Sichtweise verloren. Deshalb ist es bis heute ein Anliegen, theoretisch wie praktisch Möglichkeiten zu finden, die Texte über ihren Aktualitätsbezug hinaus messbar zu machen.

8.2.2 Quantitativ ...

Inhaltsanalysen sind natürlich nicht die einzige Methode, mit der man wissenschaftlich Texte bearbeitet. Und neben der Kommunikationswissenschaft beschäftigen sich auch andere Disziplinen mit der Analyse von Mitteilungen, wie etwa die Literatur- und die Sprachwissenschaften oder die Psychologie – jede der Disziplinen mit ganz unterschiedlichen Intentionen. Deshalb sind auch die methodischen Herangehensweisen verschieden. Beispielsweise ist die Hermeneutik, die man vielleicht als methodisches Gegenstück zur quantitativen Inhaltsanalyse betrachten kann, eine Methode, um die Situation des Entstehens, die Motivation und Intentionen des Verfassers zu analysieren. Die Hermeneutik ist also eine Lehre zur Interpretation von Kommunikationsinhalten. Die Anwendungen reichen von der rein werkimmanenten Interpretation eines Gedichts bis hin zur Interpretation von Texten in ihrer Funktion innerhalb der sozialen Realität.

Die quantitative Inhaltsanalyse versucht im Gegensatz dazu nicht, einen singulären Text zu interpretieren, sondern große Textmengen.

- **Eine Inhaltsanalyse, wie sie in aller Regel in der Kommunikationsforschung zum Einsatz kommt, hat das Anliegen, formale und inhaltliche Merkmale großer Textmengen zu erfassen.**

Die Zerlegung von Texten in Kategorien heißt natürlich zunächst, dass Bedeutungen, semantische Differenzierungen und Singularitäten, das Verständnis von Texten im Sinne sozialer Beziehungen und Ähnliches – eben alles, was sich zwischen den Zeilen abspielt – außer Acht gelassen werden. Darin gleicht die Inhaltsanalyse der Befragung. Nicht die ganze Komplexität eines Textes/einer Person wird erfasst, sondern nur wenige ausgewählte Merkmale derselben werden reduktiv analysiert. Typische Fragestellungen in der Kommunikationswissenschaft sind:

- Wie häufig verwendet der Sprecher/die Sprecherin der Tagesschau männliche und weibliche Endungen?
- Wie häufig und in welchen thematischen Kontexten wurde das Thema AIDS in den letzten 25 Jahren in den deutschen Tageszeitungen besprochen?
- Verwenden Teilnehmer von Chatrooms eher Verben oder Substantive?
- Welche Kommunikatoren kommen in bestimmten Zeitungsmeldungen vor?
- Welche Persuasions-Techniken werden in Werbespots eingesetzt?
- Wie wurden die Kanzlerkandidaten bei den letzten Bundestagswahlen dargestellt?

Solche Fragestellungen sind ganz typisch für einfache quantitative Inhaltsanalysen. Im Ergebnis stellt man dann fest, dass zum Beispiel männliche Endungen wesentlich häufiger gebraucht werden und dass sich daran seit 30 Jahren nichts geändert hat. Oder man stellt fest, dass das Wort AIDS in den 1990er-Jahren in deutschen

Zeitungen sehr häufig im Kontext mit Homosexualität und Drogenkonsum verwendet wurde. Große Textmengen werden auf ganz bestimmte Merkmale hin untersucht, häufig wird nur gezählt, wie oft bestimmte Wörter beispielsweise in Zeitungstexten vorkommen. Die reine Häufigkeit eines Merkmals, eine Prozentzahl oder ein Koeffizient sind für sich gesehen noch nicht besonders ergiebig. Ein Zahlenwert erhält erst dann eine Relevanz, wenn man einen Vergleichsmaßstab hat. Wenn man etwa feststellt, dass 1970 in 30 % aller Werbespots Hausfrauen vorkamen, während das 1990 nur noch auf zehn Prozent der Spots zutraf und man damit beispielsweise Rückschlüsse auf die Darstellung von Geschlechterrollen in der Werbung ziehen kann, dann werden die Zahlenwerte interessant. Dies gilt natürlich nicht nur für den Einsatz der Inhaltsanalyse, sondern trifft auch auf Befragungen zu. Auch deren Ergebnisse können im Vergleich sinnvoller interpretiert werden. Wenn z. B. Vielseher von Krimis, Krimiserien und Gerichtsshowes den Wert derer, die beruflich mit der Durchsetzung von Recht und Gesetz befasst sind (Polizisten, Richter etc.), auf zehn Prozent schätzen, kann man daraus kaum Schlüsse ableiten. Interpretierbar wird dieser Wert erst, wenn man ihn mit der Einschätzung der Wenigseher (fünf Prozent) und den Realdaten (nicht einmal ein Prozent) abgleicht. Bei der Inhaltsanalyse können solche Vergleichsmaßstäbe durch unterschiedliche Zeitpunkte, unterschiedliche Medien oder unterschiedliche Kulturen hergestellt werden.

8.2.3 ... und intersubjektiv nachvollziehbar

Mit systematisch und objektiv, bzw. intersubjektiv nachvollziehbar, ist das gemeint, was jeder wissenschaftlichen Methode zu eigen ist und bereits im ersten Kapitel besprochen wurde: Derartige Inhaltsanalysen sind nur dann wissenschaftlich korrekt durchgeführt, wenn das Ergebnis unabhängig vom Forscher ist und heute wie übermorgen nachvollzogen werden kann. Das betrifft den gesamten Forschungsablauf von der theoretischen Fundierung über die Hypothesenbildung, die Stichprobenziehung, die Konzeption des Messinstrumentes bis hin zur statistischen Auswertung. Es betrifft aber auch, und das ist ein ganz wesentliches Merkmal des Ablaufes einer Inhaltsanalyse, das Verständnis, das zwischen den Mitarbeitern – sie heißen hier Codierer –, die die ausgewählten Texte lesen (respektive messen), hergestellt werden muss. Intersubjektive Nachvollziehbarkeit bedeutet ganz praktisch, dass jeder, der einen Text vorgelesen bekommt, dieselbe Messung vornimmt.³

Es muss sichergestellt sein, dass die Codierer, die natürlich zuvor solide geschult werden müssen, im Prinzip austauschbar sind, dass sie alle zu jeder Zeit jeden Text so wie der Kollege „verstehen“, also bestimmte Begriffe gleich häufig finden und für

3 Das Thema wird unter dem Stichwort Reliabilität weiter unten behandelt, vgl. aber auch Kap. 3.

denselben Begriff dieselbe Codenummer vergeben.⁴ Unter der Prämisse der intersubjektiven Nachvollziehbarkeit erschließt sich dann die Bedeutung von „manifest“ durchaus neu. Während Berelson noch davon ausging, dass die Begriffe a priori einen manifesten bzw. latenten Charakter haben, werden die Begriffe nach der modernen Auffassung erst durch die Bestimmung ihres Bedeutungskerns *manifest gemacht*. Ob der Prozess gelungen ist, erschließt sich dann indirekt über die Reliabilität des inhaltsanalytischen Messinstruments.

8.3 Inhaltsanalyse als Methode zur Erfassung sozialer Realität

Die Analyse von Botschaften kann verwendet werden, um auf den Kontext der Berichterstattung, die Motive und Einstellungen der Kommunikatoren oder auf die mögliche Wirkung bei Rezipienten der Botschaften zu schließen.

8.3.1 Rückschlüsse auf den Kontext

Die typischen Ergebnisse einer Befragung befassen sich mit Einstellungen, Meinungen und dem Verhalten von Personen, z. B. mit der Zustimmung oder Ablehnung der zivilen Nutzung der Atomenergie oder dem Wahlverhalten, kurz, allen aktuellen Sachverhalten, die im weitesten Sinn die Meinung oder das Verhalten der Bürger widerspiegeln. Bei der Inhaltsanalyse erhält man ganz ähnliche Aussagen: Man beschreibt zum Beispiel die *Tendenz der Berichterstattung* hinsichtlich des Pro und Contra zur doppelten Staatsbürgerschaft, ebenfalls in Prozentpunkten. Wie schon im letzten Absatz ausgeführt, bleibt es jedoch selten bei der reinen Frequenzmessung bzw. Häufigkeitsauszählung. Das eigentlich Spannende ist der Schluss auf den Kontext.

Ein Beispiel soll dies verdeutlichen. Mit einer Inhaltsanalyse soll erfasst werden, wie viele Gewaltakte im TV pro Tag gezeigt werden. Wiederum sind die nackten Zahlen für sich genommen uninteressant (z. B. 70 Morde pro Tag; vgl. Groebel und Gleich 1993). Erst der Schluss auf den aktuellen Kontext, nämlich die medienpolitische Debatte über die Qualität der privaten und öffentlich-rechtlichen Sender, verleiht den Zahlen Relevanz. Dasselbe gilt für Studien zur Konvergenzdebatte, d. h. zu der Frage nach einer inhaltlichen Angleichung des öffentlich-rechtlichen Programms an die Privaten. Rückschlüsse von erhobenen Textmengen können natürlich auch auf historische Kontexte gezogen werden, wie etwa die Propagandaforschung über das Dritte Reich, oder, um einen aktuellen Bezug zu nehmen, die Berichterstattung der Medien über Atomkraft. Ohne den gesellschaftspolitischen Bezug, z. B. ohne die

4 Man erinnere sich an die Überführung eines empirischen in ein numerisches Relativ im zweiten Kapitel.

Berücksichtigung des atomaren GAU von 2011 in Fukushima, sind die gewonnenen Zahlen, die man dann in einem Graphen als Fieberkurve ausdrücken kann, ganz und gar unverständlich. Man muss von den Ereignissen des Jahres 2011 wissen, um den „Knick“ in der Tendenz der Berichterstattung angemessen interpretieren zu können. Der Rückschluss von einem bestimmten Datenmaterial bzw. dessen Ergebnis auf einen die Forschungsfrage betreffenden Kontext berücksichtigt folgende Aspekte:

- historischer bzw. aktueller Kontext, d.h. das gesellschaftliche Verständnis von Sachverhalten der jeweiligen Zeit
- kultureller Kontext, einmal im Laufe der Zeit innerhalb einer Gesellschaft, wie etwa Enttabuisierungen, einmal interkulturell, etwa im Vergleich Deutschland – Japan
- sozioökonomischer Kontext, d.h. die unterschiedliche Informationsverarbeitung von Menschen innerhalb einer Gesellschaft aufgrund ihres sozialen und/oder ökonomischen Status

Die Rezeption von Texten kann also nicht objektiv sein, sondern nur intersubjektiv nachvollziehbar, und dies heißt, sie ist gebunden an einen bestimmten raum-zeitlichen Kontext. An dieser Stelle ist noch einmal ein Rückbezug auf die Definition von Inhaltsanalyse notwendig: Während Berelson diesen Gesichtspunkt nicht in Erwägung zieht, macht die Definition von Merten diesen Aspekt deutlich: Man schließt von Merkmalen eines manifesten Textes auf Merkmale eines nicht-manifesten Kontextes. Das methodische Instrumentarium der Inhaltsanalyse muss also gewährleisten, dass Texte so messbar gemacht werden, dass valide erfasst wird, auf welchen Kontext sich die Inhalte beziehen, welche soziale Wirklichkeit sie widerspiegeln.

8.3.2 Rückschlüsse auf den Kommunikator

Die zweite Interpretationsmöglichkeit von Inhaltsanalysen ist der Rückschluss von Texten auf Motive, Ursachen und Gründe der Berichterstattung durch den Kommunikator. Im Fall der Kommunikationswissenschaft sind das natürlich hauptsächlich journalistische Urheber, aber es ist nicht ausschließlich darauf beschränkt. Man schließt auf die Motivation, die einen Journalisten zu dem jeweiligen Beitrag veranlasst hat oder versucht, anhand des Textes auf seine Einstellungen zu schließen. Dieser Gesichtspunkt ist immer dann relevant, wenn es darum geht, eine bestimmte Art der Berichterstattung in den Medien zu bewerten.

Angenommen, man will etwas über die Berichterstattung über das „Garching Atomkraftwerk“, den größten deutschen atomaren Forschungsreaktor, wissen. Man würde vielleicht die Berichterstattung der Süddeutschen Zeitung (SZ) und die des Münchner Merkur, einer regionalen Tageszeitung, zu diesem Thema untersuchen und die Häufigkeit sowie die Tendenz der Berichterstattung erheben. Angenommen, man käme

zu dem auf den ersten Blick etwas überraschenden Ergebnis, dass die SZ im Trend äußerst negativ über dieses Atomei berichtet, während der Merkur eher ausgewogen schreibt, die SZ wenige Wissenschaftler zu Wort kommen lässt, dafür aber viele Politiker, der Merkur dagegen erstaunlich viele Wissenschaftler. Jetzt wäre erst einmal eine Kontextfolgerung fällig: Die SZ gilt als überregional verbreitete Qualitätszeitung, beim Merkur handelt es sich um eine lokale Abonnementzeitung und man würde deshalb eigentlich erwarten, dass die Süddeutsche über die besseren, ausgewogeneren Quellen verfügt. Irgendetwas scheint nicht mit unseren Erwartungen übereinzustimmen. Man würde sich dann das verwendete Untersuchungsmaterial, also die Artikel, vornehmen, überprüfen, wer die Artikel geschrieben hat und feststellen, dass alle negativen Artikel in der SZ von einer einzigen Person stammen. Die wenigen positiven Artikel sind entweder Agenturmeldungen oder kommen von anderen Urhebern. In dem Fall könnte man schlussfolgern: Offenbar schreibt die persönliche Einstellung eines einzelnen Journalisten mit, prägt die Berichterstattung eklatant. Und man würde recherchieren, warum das wohl so ist und vielleicht darauf kommen, dass dieser Journalist selbst in Garching wohnt und in seiner Freizeit bei einem Bürgerkomitee gegen das Atomei engagiert ist. In dem Fall stellt man also eine Verbindung zwischen den Motiven, den Einstellungen, den Wünschen von Journalisten und der Berichterstattung her.

Schwieriger wird der Rückschluss von Texten auf die Einstellungen von Journalisten bzw. Kommunikatoren, wenn es sich um ältere Texte handelt und der Urheber nicht mehr lebt. Es gibt zum Beispiel Inhaltsanalysen zur Berichterstattung von Qualitätszeitungen in der Weimarer Republik. Diese Untersuchungen versuchen zu ermitteln, warum die Presse so wenig gegen die zunehmende Radikalisierung in der Politik unternommen hat. Auch in diesem Fall schließt man von der Berichterstattung – also von dem, was man inhaltsanalytisch herausgefunden hat – auf die Motive der Journalisten. Man würde sich vielleicht mit Biografien, Tagebüchern und ähnlichem Sekundärmaterial behelfen und auf diese Weise eine Außen-Validierung vornehmen. Rückschlüsse auf den Kommunikator wollen also Aussagen über dessen Einstellungen und Motive machen, befassen sich aber auch mit dessen sozialer und künstlerischer Herkunft, seinem Stil oder der Verständlichkeit von Texten. Insgesamt muss man aber festhalten, dass der Rückschluss auf den Kommunikator nicht allein auf Basis seiner Texte erfolgen kann, sondern nur mit weitergehenden Recherchen abgesichert ist. Genau genommen bewegt man sich gerade bei Rückschlüssen auf Einstellungen und Motive von Kommunikatoren auf einem spekulativen Feld.

8.3.3 Rückschlüsse auf den Rezipienten

Die dritte Interpretationsrichtung nimmt von inhaltsanalytischen Ergebnissen Rückschlüsse auf Rezipienten vor. Dieser Bereich der Inhaltsanalyse ist wissenschaftlich umstritten, aber trotzdem sehr beliebt. Bereits die Propagandaanalyse im frühen

zwanzigsten Jahrhundert ging mit einer aus heutiger Sicht problematischen Prämisse an die Untersuchung von massenmedial verbreiteten Texten heran. Man hoffte, dass Medieninhalte die Einstellung und das Verhalten von Rezipienten unmittelbar verändern: Kaum wären fünf Millionen Flugblätter abgeworfen und gelesen, schon würden alle Deutschen ihren König davonjagen und kapitulieren. Diese Wirkungsvorstellung ist in der Öffentlichkeit, im kollektiven Bewusstsein, fest verankert und bis heute anzutreffen: „Werbung korrumpiert die Leute“ und „Gewalt im Fernsehen macht gewalttätig“ sind die prominenten Beispiele. Man unterstellt dabei eine direkte Beziehung zwischen der Tendenz oder Häufigkeit von bestimmten Inhalten und dem, was Leute aus diesen Inhalten machen.

Auf der Grundlage einer Inhaltsanalyse auf die Wirkung bei den Rezipienten zu schließen, beruht auf einer Wirkungsvorstellung, wie sie das sogenannte *Stimulus-Response-Modell* vertritt (zur Relevanz dieses Modells vgl. Brosius und Esser 1998): Ein bestimmter Stimulus führt immer und bei allen Menschen zu der gleichen Reaktion. Dass Medienwirkung stattdessen ein komplexer Prozess ist, der von vielschichtigen Randbedingungen abhängt, ist heute unstrittig. Zwischen Medieninhalt und messbarem Verhalten agieren aktive Rezipienten, die bestimmte Inhalte aussuchen und andere überhaupt nicht beachten, die ihre festen politischen und religiösen Überzeugungen haben, kurz, die Medien sehr selektiv nutzen. Wirkungen sind daher von vielen Randbedingungen auf Seiten der medialen Darstellung (Gewalt ist nicht gleich Gewalt) und der Rezipienten (Zuschauer ist nicht gleich Zuschauer) abhängig. Dem trägt man methodisch Rechnung, indem nicht einfach von Medieninhalten auf Meinungen der Rezipienten geschlossen wird, sondern Rezipienten zu ihren Meinungen befragt oder in einem wissenschaftlichen Experiment untersucht werden. Folglich verlässt man sich nicht allein auf die Ergebnisse der Inhaltsanalyse, sondern zieht weitere Methoden, z. B. die Befragung, hinzu. Parallelen zum Fall einer Inhaltsanalyse, die Rückschlüsse auf den Kommunikator ziehen möchte, sind erkennbar: Das Hinzuziehen einer zweiten Methode ist notwendig, um den Zusammenhang zwischen den beiden Untersuchungsebenen herzustellen.

8.4 Anwendungsgebiete und typische Fragestellungen

Nach diesen generellen Überlegungen soll die Anwendung der Methode anhand typischer Fragestellungen bzw. Praxisbeispiele konkretisiert werden.

8.4.1 Inhaltsanalysen auf dem Feld der politischen Kommunikation

Auf dem Feld der politischen Kommunikation werden Inhaltsanalysen häufig eingesetzt für:

- die Erfassung von Themenspektren von verschiedenen Medien
- Wahlwerbeanalysen
- Propagandaforschung
- die Erfassung der zeitlichen Entwicklung der Berichterstattung über Themen
- die Feststellung von Tendenzen der Berichterstattung

Die politische Kommunikation ist sicherlich das bedeutendste Feld der Inhaltsanalyse. Wie Medien über Politik berichten, spielt für eine Gesellschaft, in der wir fast alle politischen Sachverhalte und Personen nur über Medien vermittelt bekommen, eine ganz zentrale Rolle. Politische Einseitigkeit, Bevorzugung bestimmter Kandidaten und Parteien sowie die Herausstellung bestimmter Themen und Standpunkte sind zentrale Diskussionspunkte der Gesellschaft. Hier ist die Kommunikationswissenschaft immer wieder gefragt. Die Kernfrage des Fachs, ob und wenn ja wie, unter welchen Bedingungen, Massenmedien bei den Rezipienten wirken, spielt hier implizit eine Rolle. An einem Beispiel sei dies verdeutlicht:

Man erinnere sich an den Truppeneinmarsch der USA in Somalia, der genau zur amerikanischen Hauptsendezeit stattfand. Es ist kein Geheimnis, dass dies kein Zufall war, sondern zwischen Militär und CNN bis ins Kleinste abgestimmt. Die Kameras waren die Vorhut und konnten die einmarschierenden Soldaten optimal filmen. Natürlich führt kein Land Krieg wegen der Medien, aber in diesem Fall handelt es sich um ein Ereignis, das ohne die Medien sicher zu einem anderen Zeitpunkt und in einer anderen Art und Weise stattgefunden hätte: ein typisches *mediatisiertes* Ereignis. Sie unterscheiden sich von sogenannten *genuine* Ereignissen wie etwa Naturkatastrophen dadurch, dass sie aufgrund der Präsenz der Medien anders stattfinden, in der Regel so, dass damit eine intendierte Wirkung auf die Rezipienten ins Kalkül gezogen wird. Das Ereignis findet also bei Anwesenheit der Medien anders statt als ohne Medien. Schließlich kennt man noch *Pseudoereignisse*. Diese finden *nur* statt, weil es die Medien gibt bzw. weil die Medien dabei sind.

Klassisches Beispiel hierfür sind Pressekonferenzen, die es ohne die Medien nicht gäbe. Mittels einer Inhaltsanalyse lässt sich nun der Frage nachspüren, ob sich der Anteil der verschiedenen Ereignistypen am gesamten Nachrichtenaufkommen im

Zeitverlauf verändert hat. Damit könnte man die gesellschaftspolitisch brisante Frage beantworten, ob die Politik in den letzten Jahren einen stärkeren Einfluss auf die Berichterstattung nimmt. Ein Anstieg des Anteils der Pseudoereignisse wäre ein Indikator für diese These. Um diesen Sachverhalt zu prüfen, würde man etwa folgende Bausteine für die Untersuchung festlegen:

- Zeitraum
- Auswahl der zu untersuchenden Medien
- Festlegung der Untersuchungseinheiten (z. B. welche Rubriken in Tageszeitungen, welche Sendungen in Fernsehen und Radio)
- Festlegung der Stichprobe
- Definition dessen, was Pseudo-, mediatisierte und genuine Ereignisse sind

Nach der Entwicklung des entsprechenden inhaltsanalytischen Messinstruments, des Codebuchs, kann das Codieren beginnen. Als Ergebnis würde man für die einzelnen Jahre die verschiedenen Anteile der drei Ereignistypen erhalten und könnte beispielsweise zeitliche Entwicklungen darstellen. Unter Berücksichtigung aller Randbedingungen lassen sich nun Rückschlüsse eines manifesten Inhaltes (der Anteil der Pseudoereignisse) auf einen nicht-manifesten Kontext (Einfluss des politischen Systems auf das Mediensystem) treffen.

8.4.2 Inhaltsanalysen in der Gewaltforschung

In der Kommunikationswissenschaft wird Gewaltforschung häufig aus der theoretischen Perspektive der *Kultivierungshypothese* betrachtet. Auch hier spielen Inhaltsanalysen eine Rolle. Gerbner selbst hat jährlich sogenannte „Violence-Profiles“ erstellt, in denen detailliert die Anzahl der Gewaltakte pro Sendung und Stunde, die Merkmale der Täter und Opfer etc. ausgezählt wurden. Die erwähnte Studie von Groebel und Gleich hat nach dem eben beschriebenen Schema gezählt, wie häufig im Fernsehen Menschen umgebracht werden. Die Zahl der Morde an sich besagt nicht viel; erst durch entsprechende Kontexte wird ihnen Aussagekraft verliehen, indem man beispielsweise nach Sendern und Tageszeiten differenziert, Vergleiche mit früheren Studien oder anderen Ländern anstellt. In diesem Zusammenhang werden häufig Methodenkombinationen gewählt, um festzustellen, ob und in welchem Maße die Medieninhalte das Weltbild der Rezipienten prägen. Mittels Inhaltsanalyse stellt man beispielsweise fest, wie viel Prozent der Protagonisten in Fernsehserien mit Verbrechensbekämpfung beschäftigt sind. Anschließend ermittelt man mit einer Befragung, wie viel Prozent der Bevölkerung nach Meinung von Viel- und Wenigsehern mit Gewaltbekämpfung beschäftigt sind. Divergierende Prozentzahlen könnten dann ein Indiz für die Kultivierungshypothese sein. Hier muss noch einmal betont werden, dass eine Inhaltsanalyse allein nichts über die Wirkung eines Mediums aussagt, son-

dern nur die Inhalte des Mediums erhebt. Wenn man also feststellt, dass 70 Morde im deutschen Fernsehen pro Tag verübt werden, heißt das noch nicht, dass Gewalt im Fernsehen schädlich ist. Rückschlüsse auf nicht-manifeste Kontexte sind immer heikel und nur unter der Bedingung möglich, dass man zuvor diesen Kontext genau untersucht, definiert und kontrolliert hat.

8.4.3 Inhaltsanalysen in der Minderheitenforschung

Mittels Inhaltsanalyse wird hier festgestellt, wie häufig und in welchen Rollen verschiedene Minderheiten in den Medien dargestellt werden. Die Ergebnisse belegen immer wieder, auch im internationalen Vergleich, dass die dargestellte Realität in den Medien von der tatsächlichen Realität abweicht. Beispielsweise sind im Fernsehen die Figuren in Serien meist zwischen 20 und 40 Jahre alt. Ältere Menschen und kleine Kinder werden kaum gezeigt und wenn überhaupt, dann nur in bestimmten, klischeehaften Rollenkontexten. Das heißt, man findet per Inhaltsanalyse Verteilungen, die man zunächst einmal mit denen aus dem wirklichen Leben vergleicht. Ein weiterer Befund aus dieser Forschungsrichtung zeigt, dass Frauen zumeist jünger als Männer sind, im Schnitt etwa zehn Jahre. Die Beispiele zeigen, dass man nicht nur Texte im engeren Sinne inhaltsanalytisch untersuchen kann, sondern auch bildliche Darstellungen. Ein interessanter Unterschied in dem Zusammenhang beruht auf dem sogenannten *Face-ism-Index*, den man für Inhaltsanalysen entwickelt hat. Hierbei wird gemessen, wie viel Prozent der gesamten sichtbaren Darstellung eines Körpers der Kopf einnimmt. Je kleiner der Index, desto stärker ist die entsprechende Person körperbetont dargestellt. Auch hier sind die Befunde über Grenzen, Zeiträume und Kulturen hinweg sehr ähnlich: Frauen werden körperbetonter dargestellt als Männer. Das heißt, Männer werden entsprechend „kopflastiger“ abgebildet. In derartigen Studien werden dann nicht nur Zeitschriften und Nachrichtensendungen der Gegenwart untersucht, sondern auch Gemälde des 15. Jahrhunderts und auch damals war es so, dass Männer häufiger kopfbetont dargestellt wurden als Frauen.⁵

5 Wie bei der Darstellung der Gewaltforschung schon angedeutet, erhält die Minderheitenforschung ihre Brisanz aus der Kontrastierung der Medieninhalte mit statistischem Datenmaterial über die Realität (vgl. hierzu Haas & Scheufele 2012). Ob es allerdings die Funktion von Soaps und Sitcoms ist, Realität abzubilden, erscheint fraglich. Ein vergleichender Rückschluss auf soziale Realität ist also nicht in jedem Fall angemessen.

8.5 Die Vorteile der Inhaltsanalyse gegenüber anderen Methoden

8.5.1 Darstellung vergangener Kommunikationsprozesse

Die erste Stärke der Inhaltsanalyse gegenüber Befragung oder experimentellem Design ist die Tatsache, dass man Aussagen über Medieninhalte und Kommunikationsprozesse der Vergangenheit machen kann. Der Forscher ist unabhängig von der physischen Anwesenheit und dem Erinnerungsvermögen von Personen; Tote kann man nicht befragen oder sie als Versuchspersonen in einem Labor einsetzen. Und auch wer Menschen nach ihrem Verhalten oder ihren Einstellungen von vor 20 Jahren fragt, sollte keine allzu validen Antworten erwarten. Vergangene Kommunikationsprozesse sind dagegen für sich schon sehr erhellend und gewinnen noch an Aussagekraft, wenn sie als vergleichendes Material zur Analyse von gegenwärtigen Kommunikationsprozessen hinzugezogen werden.

Ein Beispiel: Man will die gegenwärtige Berichterstattung über Kernenergie, vielleicht auch eine mögliche Veränderung durch die Katastrophe in Fukushima untersuchen. Gleichzeitig interessiert man sich für die öffentliche Meinung zu dem Thema. In beiden Fällen kann die Analyse der vergangenen Berichterstattung zur Einordnung der Ergebnisse dienen. Die Ergebnisse der Analyse der gegenwärtigen Berichterstattung können vor dem Hintergrund der Berichterstattung der 1970er- und 1980er-Jahre besser interpretiert werden. Man kann etwa den Anteil der positiven und negativen Aussagen zur Kernenergie im Zeitverlauf betrachten. Um die Veränderungen der Berichterstattung vor und nach Fukushima besser beurteilen zu können, ist es hilfreich, die Veränderungen der Berichterstattung im Umfeld der Störfälle von Harrisburg und Tschernobyl zu untersuchen.

Etwas vorsichtiger sollte man die damalige Berichterstattung dagegen im Hinblick auf die Frage nach der öffentlichen Meinung zur Kernenergie interpretieren. Es ist offensichtlich, dass man die Bürger nicht mehr nach ihrer Einstellung zur Kernenergie in den 1970er- und 1980er-Jahren befragen kann. Daran kann sich (abgesehen vielleicht von extremen Gegnern und Befürwortern) niemand mehr zuverlässig erinnern. Eine Analyse der damaligen Medieninhalte kann zumindest Hinweise auf das gesamtgesellschaftliche Stimmungsbild liefern. Dabei ist allerdings zu berücksichtigen, dass die in den Medien veröffentlichte Meinung nicht zwangsläufig mit der Mehrheitsmeinung innerhalb der Bevölkerung übereinstimmen muss. Was in der Darstellung vergangener Kommunikationsprozesse schließlich nicht sichtbar wird, sind Wirkungsprozesse in der Vergangenheit. Man kann zwar untersuchen, ob und wie sich die Berichterstattung seit den 1970er-Jahren verändert hat. Und man kann beispielsweise auch Leserbriefe zum Thema als Indikator für die Meinungsverteilung in der damaligen Gesellschaft auswerten. Man sollte auf Basis dieser Datenquellen allerdings keinen Wirkungszusammenhang zwischen Medieninhalten und der Einstellung der Bevölkerung in dieser Zeit konstruieren.

8.5.2 Papier ist geduldig – Befragte sind es nicht

Der zweite Vorteil der Inhaltsanalyse gegenüber anderen empirischen Methoden ist die Tatsache, dass der Forscher nicht auf die Kooperation von Befragten oder Versuchspersonen angewiesen ist. Niemand verweigert ein Interview oder bricht es ab, kein Interviewer muss sich an der Tür beschimpfen lassen. Zeitungsartikel und Videobänder laufen nicht weg. Man kann sie zeitunabhängig analysieren und man kann sie immer wieder untersuchen. Sie verändern sich in der Regel nicht. Für einen Masterkandidaten spielt es keine Rolle, ob er seine Zeitungsartikel morgens, mittags oder in der Nacht codiert, Papier ist geduldig und immer erreichbar. Abgesehen von der Verfügbarkeit ist es natürlich ein Unterschied, ob man Menschen morgens beim Frühstück oder nach der Arbeit interviewt. Sie könnten je nach Stimmung und Tagesform unterschiedliche Antworten geben. Das heißt, sie reagieren in irgendeiner Form mit dem Untersuchungsinstrument.

- **Die Inhaltsanalyse dagegen gilt allgemein als ein nicht-reaktives Verfahren,⁶ d. h., der Untersuchungsgegenstand verändert sich nicht – egal, wann und wie oft man ihn untersucht.**

Daher sind Inhaltsanalysen prinzipiell auch beliebig reproduzierbar und damit auch modifizierbar. Man kann eine Inhaltsanalyse, die irgendjemand gemacht hat, in genau der gleichen Form wiederholen, wenn man das Ergebnis anzweifelt. Der gleiche Gegenstand kann wiederholt untersucht werden, ohne dass er sich verändert. Man kann zwar auch die gleiche Person wieder befragen, aber sie wird nicht mehr dieselbe sein wie noch vor zwei Tagen.

Der Reproduzierbarkeit sind natürlich Grenzen gesetzt. Und diese Grenzen liegen in der Reliabilität, der Zuverlässigkeit des Messinstruments. Eine Untersuchung völlig zu reproduzieren, setzt voraus, dass das Messinstrument, in dem Fall das Codebuch der Inhaltsanalyse, vollständig reliabel ist. Das bedeutet praktisch, dass neue Codierer/Codiererinnen genau zu dem gleichen Ergebnis kommen müssen wie die, die bereits länger mitarbeiten. Und da stößt man schnell an Grenzen, wenn zwei gleiche Untersuchungen zeitlich weit auseinanderliegen. **Die Reliabilität des Messinstruments (des Codebuchs) hat vermutlich ein Verfallsdatum.** Anders als ein Zentimetermaß misst es nach fünfzig Jahren nicht mehr dasselbe, weil sich der soziale, kulturelle und politische Kontext verändert hat, in dem die Codierer leben. Unter diesem Aspekt erscheint die Inhaltsanalyse dann als reaktives Verfahren, weil die Codierer, nicht der Untersuchungsgegenstand selbst, auf das Messinstrument reagieren.

Um das Verständnis der Methode zu erleichtern, haben wir abschließend noch einmal die wesentlichen Elemente der beiden Methoden zusammenfassend einander gegenübergestellt (Tab. 8.1).

6 Eine ausführliche Diskussion des Konzepts Reaktivität findet sich bei Merten und Großmann (1996).

Tab. 8.1 Elemente von Befragung und Inhaltsanalyse

Kriterien	Befragung	Inhaltsanalyse
Grundgesamtheit	Bevölkerung, Teile davon	Berichterstattung der Massenmedien, Teile davon
Merkmalsträger	Personen	Texte, Spielfilme, Reden Artikel, Teile davon
Datenerhebung	Interview	Codierung
Erhebungsinstrument	Fragen	Kategorien
Reaktivität	ja	nein
Realitätsbezug	Gegenwart	Gegenwart, Vergangenheit

9.1 Kategorien als Erhebungsinstrument der Inhaltsanalyse

In diesem Abschnitt wird genauer auf das Erhebungsinstrument, das Kategoriensystem bzw. das Codebuch, eingegangen. Ebenso wie in den Abschnitten zur Befragung die Fragen im Mittelpunkt der Darstellung standen, ist für das Verständnis dieser Methode die Kategorienbildung essentiell. Was bei der Befragung der Fragebogen ist, ist für die Methode der Inhaltsanalyse das Kategorienschema. Man sollte sich hierbei noch einmal Kapitel 2 vor Augen führen, in dem Messen und Zählen sowie die verschiedenen Skalierungsmöglichkeiten erläutert wurden. Denn das Prinzip, nach dem Messinstrumente entwickelt werden, ist immer dasselbe, nur die Form verändert sich je nach Methode.

9.1.1 Inhaltliche Kategorien

Kategorien sind vergleichbar mit den Programmfragen einer Befragung. Die Kategorien sind zunächst einmal die exakte Definition dessen, was erhoben bzw. gemessen werden soll. Wie man für eine Befragung die Programmfragen in Testfragen – also solche, die dann tatsächlich im Fragebogen stehen – ausformuliert, werden die Kategorien anhand der Indikatoren, mit denen man seine Fragestellung entfaltet hat, gebildet. Je nach Differenziertheit der Fragestellung werden sie in Unterkategorien aufgeteilt. Am Ende der Kategorienbildung steht dann die Festlegung der Merkmalsausprägungen an. Das lässt sich wieder mit der Fragebogenkonstruktion vergleichen: Je nach Untersuchungstiefe entwickelt man unterschiedlich differenzierte Kategorienausprägungen. Zusätzlich wird häufig die Ausprägung „Sonstiges“ vergeben.

Ein Beispiel: Es geht um die Frage, ob sich die Anforderungen an den Beruf des Journalisten in den letzten dreißig Jahren gewandelt haben. Um dies herauszufinden, sollen Stellenanzeigen inhaltsanalytisch untersucht werden, weil angenommen wird, dass der Stellenmarkt einen guten Indikator für Veränderungen (Ausdifferenzierung oder Professionalisierung) von Berufsbildern abgibt. Eine ganz zentrale Dimension sind natürlich die Anforderungen, die in diesen Stellenanzeigen an Bewerber gestellt

Tab. 9.1 Beispiel für die Darstellung der beiden Kategorien Studienabschluss sowie Berufserfahrung in bestimmten Medien

Studienabschluss (Uni, FH, TU, Journalistenschule)
0 = nicht gefordert
1 = Studium gewünscht, Abschluss nicht gefordert (z. B. „einige Semester Kommunikationswissenschaft“)
2 = Studienabschluss gewünscht
3 = Studienabschluss gefordert bzw. unabdingbar
Berufserfahrung in bestimmten Medien
0 = nicht gefordert
1 = allgemein Medien („journalistischer Bereich“)
2 = Print
3 = Hörfunk
4 = Fernsehen
5 = Verlag
6 = Presseagentur
7 = Sonstiges

werden. Die *Hauptkategorie* wird also „Geforderte Eigenschaften des Bewerbers“ genannt. Man wird dann bei der Entwicklung der *Unterkategorien* empirisch vorgehen und alle denkbaren Anforderungen auflisten. Dabei ist an formelle Anforderungen wie „Schulbildung“ oder „Berufserfahrung“ genauso zu denken wie an die persönlichen Eigenschaften eines Bewerbers, also „Kontaktfreudigkeit“ oder „Zuverlässigkeit“. Im nächsten Schritt müssen die Merkmalsausprägungen der Kategorien (Skalierung!) festgelegt werden. Beispiele für solche Kategorien einschließlich der Merkmalsausprägungen finden sich in Tab. 9.1.

Wenn im Text einer Anzeige ein Journalist mit Studienabschluss und Erfahrung im Verlagswesen gesucht wird, müsste in der ersten Kategorie eine 2, in der zweiten Kategorie eine 5 vergeben werden. Neben den konkreten Ausprägungen der Kategorie „Studienabschluss“ sind zusätzlich Vermerke (hier in Klammern) angebracht. Dies ist die operationale Definition des Begriffs „Studienabschluss“. Hier handelt es sich natürlich um ein simples Beispiel, aber es demonstriert das Prinzip. Wenn die Kategorie nicht eindeutig für sich spricht, muss angegeben werden, was mit dieser Kategorie genau gemeint ist. Die Begriffsbestimmung ist ja der zentrale Punkt einer Untersuchung, an dem auf Basis einer theoriegeleiteten Hypothese eindeutig bestimmt wird, was hier untersucht wird. So werden Transparenz und intersubjektive Nachvollziehbarkeit hergestellt.

9.1.2 Formale Kategorien

- **Formale Kategorien beschreiben die formalen Merkmale der jeweiligen Untersuchungseinheit.**

Das können ganze Artikel oder auch nur die Überschriften von Artikeln sein. In unserem Beispiel bildet jede Stellenanzeige eine Untersuchungseinheit, die exakt identifiziert werden muss: Wann ist sie erschienen? In welchem Medium? Welche Größe hat sie? Diese formalen Variablen dienen nicht nur der Wiedererkennung, sie sind besonders dann wichtig, wenn formale Gesichtspunkte bei der Forschungsfrage eine Rolle spielen. Wenn also, wie in diesem Fall, eine Trendstudie vorgenommen wird, ist das Erscheinungsjahr natürlich wesentlich. Auch der Umfang kann entscheidend sein, wenn man etwa herausfinden möchte, wie intensiv über ein bestimmtes Thema berichtet wird. Im Falle der vorliegenden Studie sehen die formalen Kategorien etwa so aus, wie Tab. 9.2 zeigt.

Tab. 9.2 Beispiele für formale Kategorien

Medium (Erscheinungsmedium der Anzeige)
1 = journalist
2 = Die Feder
3 = Die Zeit
4 = FAZ
5 = SZ
Tag (sofern vorhanden zweistellig eintragen)
Monat (zweistellig)
Jahr (zweistellig)
Seite (dreistellig)
Lfdnr (laufende Nummer der Anzeige pro Seite)
Text (Textanfang im Klartext)
Rubrik (Anzeigenrubrik nur bei SZ, FAZ, ZEIT)
0 = keine Rubriken vorhanden
1 = Anzeige in Medienrubrik
2 = Anzeige in Rubrik „Dienstleistung“
3 = Anzeige in anderer spezifischer Rubrik
4 = Anzeige in Rubrik „Sonstige“
Illustration
0 = keine Illustration (nur Fließtext)
1 = nur Firmenlogo
2 = nur Bild/Grafik
3 = Bild, Grafik und Firmenlogo
4 = Sonstiges

Anhand dieser Kategorien sind alle Anzeigen eindeutig zu identifizieren und zu codieren. Bei Inhaltsanalysen zur Medienberichterstattung werden üblicherweise noch Angaben zur Größe des Textes und zur Platzierung gemacht. Es gibt Codierungen des Umfanges eines Beitrags, die relativ grob sind. Die Ausprägungen könnten zum Beispiel „1 Seite“, „½ Seite“, „¼ Seite“, „1/8 Seite“ heißen. Bekanntermaßen wählt man die Skalierung nach der Wichtigkeit des Merkmals für die Untersuchung. Eine grobe Skalierung deutet also an, dass der Umfang keine besondere Rolle spielt. Anders ist es, wenn der Umfang in Quadratzentimetern oder bei elektronischen Medien in Sekunden, also intervallskaliert, gemessen wird.

Die Platzierung eines Berichtes, sei es in der Zeitung oder in Nachrichtensendungen, sagt im weitesten Sinn etwas über die diesem Beitrag zugeschriebene Relevanz aus. Dinge, die auf der ersten Seite stehen, hat die Redaktion für wichtiger gehalten als eine Meldung im Blattinneren. Rezipienten lesen Meldungen auf der Titelseite auch häufiger als Meldungen im Inneren der Zeitung. Derartige Gesichtspunkte sind zum Beispiel für Agenda-Setting-Studien von Belang.

Die formalen Kategorien stehen – ebenso wie die soziodemografischen Daten in Befragungen – nicht im eigentlichen Zentrum der Untersuchung, liefern jedoch wichtige Zusatzinformationen. Mit ihnen können schon einige wichtige Aussagen getroffen werden, im Fall der Journalistenstudie beispielsweise darüber, ob das Anzeigenaufkommen im Zeitverlauf zu- oder abgenommen hat.

- Wie hat sich das Anzeigenaufkommen in den untersuchten Medien im Vergleich entwickelt?
- Ist vielleicht ein Trend hin zu den überregionalen Tageszeitungen zu entdecken oder sind es die Fachblätter, in denen sich der Stellenmarkt materialisiert?
- Werden heute Anzeigen aufwendiger gestaltet als früher, legen also die Firmen mehr Wert auf ihre Selbstdarstellung?

Konkret wird so vorgegangen, dass unter der Kategorie „Medium“ die entsprechende Ausprägung codiert wird. Stammt eine Anzeige aus der FAZ, vergibt man beispielsweise eine 4. Analog verfährt man mit den folgenden Merkmalsausprägungen. Man notiert den Tag bis hin zum „Textanfang“, einer quasi offenen Antwortkategorie, in der die ersten Worte der jeweiligen Anzeige vermerkt werden.

1 Zur Erinnerung: Daten, die intervallskaliert erhoben werden, eröffnen mehr Möglichkeiten der Datenanalyse. Während nominale Skalen „nur“ Häufigkeitsauszählungen zulassen, kann man mit intervallskalierten Daten Durchschnittswerte berechnen.

9.2 Codebogen

Auf dem sogenannten Codebogen protokolliert der Codierer die Ergebnisse seiner Messung. Der Codebogen enthält in Spalten angelegt die „Plätze“ für die Vergabe der entsprechenden Codes. Die Zeilen repräsentieren den einzelnen Fall, d. h. die Untersuchungseinheit, in unserem Beispiel also die einzelne Stellenausschreibung. Man nehme noch einmal das Beispiel der Journalistenstudie. Gegeben sei eine Anzeige vom 03. Februar 2012 aus der SZ. Die Kategorie „Medium“ wurde zuvor einstellig vergeben, so dass man in dem dafür vorgesehenen Kästchen eine 5 vergibt. Die nächste Kategorie muss zweistellig angelegt werden, denn „Tag“ kann zweistellig werden; man vergibt eine 03. „Monat“ und „Jahr“ sind ebenfalls zweistellig angelegt. Am Ende liegt eine Zahlenreihe vor, die sich entsprechend der angegebenen Informationen so liest: 5030212... Hier wird deutlich, was es praktisch heißt, empirische Fakten in ein numerisches Relativ zu überführen (vgl. Kap. 2). Nach der Vercodung hat sich die konkrete Stellenausschreibung in einen „Fall“ gewandelt. Man spricht dann in der Phase der Datenanalyse auch von Fällen. Aus dem Text mit seinen vielen Besonderheiten ist eine lange, maschinenlesbare Zahlenreihe geworden, die nur noch die interessierenden Merkmale enthält.

9.3 Codebuch

Woher wissen die Codierer, was sie zu tun haben? So wie alle Instruktionen für die Befragung im Fragebogen stehen, brauchen Codierer genaue Handlungsanleitungen, wie sie mit den zu analysierenden Medieninhalten umgehen sollen. Diese sind im Codebuch vermerkt, ohne dieses sind die Codierer orientierungslos. In den Codieranweisungen wird jede Kategorie im Detail beschrieben. Ziel ist eine größtmögliche Reliabilität und Validität der Kategorien: Alle involvierten Codierer sollen im Idealfall jeden beliebigen Text gleich verstehen, also dieselben Codierungen vornehmen. Als Beispiel hierfür dient wieder eine Aussage aus einer fiktiven Stellenanzeige aus der Journalistenstudie: „Wir suchen einen Wirtschaftsjournalisten mit möglichst weitreichenden Kenntnissen der gängigen Datenanalyseprogramme.“ Handelt es sich hier um eine Anforderung, die der Bewerber erfüllen muss oder um einen Wunsch? Diese Differenzierung ist wichtig für die Ausgangsfrage, ob und inwiefern sich das Berufsbild des Journalisten verändert hat. In der Codieranweisung liest sich das dann folgendermaßen:

- „Wunsch versus Forderung“: Forderungen sind erkennbar am Fehlen einer Einschränkung (z. B. „Wir suchen Dipl.-Journalisten mit mind. dreijähriger Berufserfahrung“) oder am Terminus „Voraussetzung“. Wünsche sind meist an Einschränkungen („möglichst“, „eventuell“, „vielleicht“) oder konkreten Wunschformen („wünschenswert wäre“, „von Vorteil wäre“) geknüpft.

Codieranweisungen können neben operationalen Definitionen wie der oben genannten auch allgemeine Hinweise enthalten – zum Beispiel zur Frage, was als Fall bzw. Analyseeinheit zu betrachten sei. Hier wieder eine entsprechende Codieranweisung:

- Analyseeinheit ist die angebotene Stelle. Eine Stellenanzeige kann mehrere Angebote enthalten, siehe Anweisung unter „laufender Nummer“. Codiert werden alle Stellenangebote, in denen Journalisten im Sinne der folgenden Definition gesucht werden: (es folgt eine ausführliche Definition des Begriffs Journalist).

Auch die Durchführung der Codierung kann unter Umständen, je nach Komplexität derselben, genau festgelegt werden: Es kann beispielsweise vorkommen, dass man eine Untersuchungseinheit mehrmals lesen oder als Video anschauen muss, bis man alle Codierungen durchgeführt hat. Das trifft vor allen Dingen dann zu, wenn man mit Fernseh- oder Radiomaterial arbeitet, bei dem bis zu fünf Durchgänge notwendig werden, um alle Kategorien zu erfassen. Es gibt Codieranweisungen, die vom Codierer verbindlich fordern, den Beitrag mindestens dreimal jeweils nach verschiedenen Kriterien anzusehen und zugleich festlegen, bei welchem Durchgang welche Kategorien zu vercoden sind. Das könnte in etwa so lauten:

- Erster Durchgang: Der Beitrag wird vollständig angesehen. Dabei wird die Laufzeit des Beitrags verschlüsselt. Zweiter Durchgang: Alle Themen und ihre Urheber, die im Beitrag vorkommen, werden verschlüsselt. Dritter Durchgang: Für jeden Urheber werden Stilformen verschlüsselt.

Auch bei Zeitungen kann man solche Anweisungen geben. Jeder Beitrag wird zunächst einmal kursorisch überflogen. Das ist notwendig, denn obwohl jede inhaltliche Information in einen Zahlenwert verwandelt wird, muss der Codierer erst einmal verstehen, um was es sich hier handelt. Im zweiten Durchgang erhält der Codierer dann die Anweisung, detaillierte Codierungen vorzunehmen. Damit ist das Erhebungsinstrument einer Inhaltsanalyse vollständig. Das Codebuch enthält somit:

- allgemeine Hinweise und Hintergrundinformationen,
- Haupt- und Unterkategorien,
- operationale Definitionen,
- Codieranweisungen und
- Codebogen.

9.4 Formale Anforderungen an Kategorien:
Vollständigkeit und Trennschärfe

9.4.1 Vollständigkeit von Kategorien

An einem Beispiel soll gezeigt werden, welche prinzipiellen Anforderungen an die Bildung von Kategorien gestellt werden müssen. Nehmen wir an, wir hätten in der Studie über journalistische Stellenanzeigen folgende Kategorie entwickelt (Tab. 9.3).

Bei einer Probecodierung stellt sich heraus, dass die Ausprägung „Sonstiges“ sehr häufig vergeben wurde. Das ist nicht gut, denn hier werden Informationen zusammengeworfen, die später nicht mehr auseinanderzuidividieren sind. Man will ja etwas über die (gestiegenen) Anforderungen an Journalisten herausfinden. Wenn im Ergebnis herauskommt, dass 15 % der Firmen keine Anforderung stellen, 25 % Kommunikationswissenschaft, 30 % Diplom-Journalistik und 30 % „Sonstiges“ fordern, ist dieses Ergebnis nicht detailliert genug, weil man von fast einem Drittel nur weiß, dass etwas gefordert wurde, aber nicht, was. In so einem Fall muss man tatsächlich im vorliegenden Untersuchungsmaterial nachprüfen, was sich unter „Sonstiges“ verbirgt. Es gibt verschiedene Möglichkeiten: Zum einen ist es wahrscheinlich, dass eine wichtige Ausprägung nicht berücksichtigt wurde. Es gibt zum Beispiel nicht-akademische Studienrichtungen wie den Kommunikationswirt, der praxisorientiert ausgebildet wird. Stellt man dann fest, dass sich vielleicht noch einige weitere typische Studiengänge (Germanistik, Politik) in den „Sonstigen“ verbergen, die man zusätzlich als Ausprägungen mitaufnehmen kann, hat man diese Kategorie auf ein erträgliches Maß reduziert. Nach der Probecodierung wird man also die genannten Ausprägungen „Kommunikationswirt“ etc. mit aufnehmen, so dass die Kategorie „Studienrichtung“ jetzt sechs Ausprägungen besitzt. Offenbar war diese Kategorie vor der Probecodierung nicht vollständig.

- **Vollständigkeit ist eine zentrale Forderung an ein Kategorienschema, denn nur unter dieser Voraussetzung kann die zentrale Forschungsfrage auch erschöpfend beantwortet werden.**

Wenn bestimmte Aspekte nicht in die Kategorien aufgenommen werden, ist eine Untersuchung unvollständig und im schlimmsten Fall unbrauchbar. Ob ein Begriff, überführt in eine inhaltsanalytische Kategorie (oder eine Reihe von Testfragen in

Tab. 9.3 Beispiel für die Kategorie Journalistische Studienrichtung I

Journalistische Studienrichtung
0 = nicht gefordert
1 = Kommunikationswissenschaft
2 = Diplom-Journalistik
3 = Sonstiges

einer Befragung) vollständig erfasst ist, lässt sich anhand älterer Studien zum selben Thema, zugrunde gelegter Theorien und nicht zuletzt des gesunden Menschenverstandes und der eigenen Erfahrung entscheiden. Insgesamt muss ein vollständiges Kategorienschema mit allen Unterkategorien und ihren Ausprägungen das theoretische Konstrukt der Untersuchung *umfassend und exklusiv* widerspiegeln. Erst unter dieser Voraussetzung misst das Messinstrument das, was es messen soll, ist also valide.

9.4.2 Trennschärfe der Kategorien

Die Inhaltsanalyse zerlegt Begriffe in wohldefinierte Bedeutungsräume, die in Form von Daten vorliegen und weiterverarbeitet werden. Die Sezierung dieser Textmengen wird vorgenommen, um über die singulären Texte hinweg auf verallgemeinerbare Zusammenhänge zu stoßen. Dazu ist es nicht nur notwendig, alle Begriffe der Forschungsfrage gemäß vollständig zu erfassen und abzubilden, sondern auch, sie sauber voneinander zu trennen. Trennschärfe ist eine Anforderung sowohl innerhalb einer Kategorie als auch zwischen Kategorien. Hier wieder das bekannte Beispiel (Tab. 9.4):

Trennschärfe wäre dann *nicht* gegeben, wenn eine weitere Ausprägung „Geisteswissenschaft“ lauten würde, weil Germanistik-Absolventen sowohl dort als auch unter „Germanistik“ verschlüsselt werden könnten.

- **Kategorien sind trennscharf, wenn sich die einzelnen Ausprägungen wechselseitig ausschließen und wenn alle Ausprägungen sich auf das gleiche Merkmal beziehen.**

Der Codierer könnte nicht entscheiden, welche Ausprägung er wählen soll: Im Prinzip müsste er zwei Codes vergeben. Der Informationsgehalt und auch die Reliabilität einer nicht trennscharfen Kategorie sinken, weil Codierer uneinheitlich codieren werden. Man kann Trennschärfe auch unter dem Aspekt der Eindimensionalität innerhalb einer Kategorie betrachten. Im vorliegenden Beispiel wären zwei Dimensionen angesprochen, nämlich konkrete Fächer und allgemeine Studienrichtungen. Eine Kategorie darf nur eine Antwortdimension oder -ebene zulassen.² Es ist ein wenig wie bei dem Kinderspiel: Welcher Ausdruck passt nicht in diese Reihe: Huhn, Fasan, Ente, Geflügel? Ein noch deutlicheres Beispiel: Eine Codierung von Akteuren nach einer Kategorie mit den beiden Ausprägungen „männlich“ und „katholisch“ würde die Forderungen nach Vollständigkeit und Trennschärfe einer Kategorie in besonderem Maße verletzen. Das allgemeine Ziel einer Inhaltsanalyse bleibt, Aussagen über große Textmengen zu machen. Dies erfordert eine exakte Identifizierung und Sortierung jedes Falls. Dies ist nur gegeben, wenn Trennschärfe und Vollständigkeit vorliegen.

2 Das gilt in gleicher Weise natürlich auch für Antwortvorgaben im Fragebogen. Auch diese müssen vollständig und trennscharf sein.

Tab. 9.4 Beispiel für die Kategorie Journalistische Studienrichtung II

Journalistische Studienrichtung
0 = nicht gefordert
1 = Kommunikationswissenschaft
2 = Diplom-Journalistik
3 = Kommunikationswirt (FH)
4 = Germanistik
5 = Politikwissenschaft
6 = Sonstiges

9.5 Feststellung der Validität und Reliabilität des Kategorienschemas

Misst das Messinstrument, was es messen soll, und misst es das auch zuverlässig? Diese generellen Fragen zur Gültigkeit und Zuverlässigkeit finden hier ihre konkrete Anwendung.

Zur *Validitätsprüfung* ist in diesem Abschnitt schon gesagt worden, dass die Forderung nach Vollständigkeit ein zentrales Kriterium für die Güte der gesamten Untersuchung ist. Ob nun das Kategorienschema vollständig, also valide ist, lässt sich anhand früherer Forschungsdesigns, Plausibilitäten, theoretischer Überlegungen und anderer Außenkriterien prüfen. Je häufiger zum Beispiel ein Gegenstand schon untersucht wurde, desto sicherer kann sich der Forscher sein, alle relevanten Aspekte berücksichtigt zu haben. Es ist dann noch seine eigene Aufgabe, aus dem vorhandenen Material die für seine konkrete Forschungsfrage relevanten Definitionen, sprich Kategorien, explizit zu machen.

An dieser Stelle der Konzeption berühren sich Validität und Reliabilität des Messinstrumentes. Generell kann man sagen, dass eine hohe Validität, die jeden noch so kleinen Aspekt des theoretischen Konstruktes zu umfassen versucht, zu Lasten einer hohen Reliabilität geht. Je detaillierter die Verschlüsselung wird, zum Beispiel die Verschlüsselung von Bewertungen politischer Nachrichten, desto größer ist die Fehlerquote bei der Codierung. Umgekehrt gilt natürlich, dass „harte“ Indikatoren wie etwa formale Kategorien eine nahezu hundertprozentige Reliabilität fordern. Wenn ein Codierer nicht in der Lage ist, das Medium, in dem eine Stellenanzeige erscheint, mit der richtigen Codeziffer zu versehen, ist das keine fehlende Reliabilität des Codebuchs, sondern ein Problem des Codierers.

Der Codierer ist die vernunftbegabte Zählmaschine des Forschers. Er entscheidet, um welche Ausprägung eines Merkmals es sich bei einem konkreten Begriff, Satz oder Textbaustein handelt. Das ist keinem in die Wiege gelegt und deshalb bedarf es einer genauen Schulung für jede einzelne Studie. Meist werden sich mehrere Codierer die Arbeit teilen, so dass es nicht nur notwendig ist, das Codebuch gut zu beherrschen, sondern auch die bei allen Codierern gleichbleibende genaue und eindeutige Verschlüsselung zu gewährleisten. Wenn dies der Fall ist, gilt das Kategorienschema als reliabel.

- Eine Reliabilitätsprüfung, die sowohl während der Pretestphase als auch während der Codierung selbst vorgenommen wird, misst, wie groß die Übereinstimmung zwischen verschiedenen Codierern ist und wie zuverlässig gleichbleibend ein Codierer sein Material verschlüsselt.

Die Ergebnisse, ausgedrückt in einem Koeffizienten, geben Auskunft über die Zuverlässigkeit des Messinstruments. Die sogenannte *Intracoderreliabilität* misst man, indem ein Codierer einen Teil seines verschlüsselten Materials in gebührendem zeitlichem Abstand zum zweiten Mal codiert. Man geht davon aus, dass sich niemand an einzelne Verschlüsselungen erinnern kann, beim zweiten Mal also nicht aus der Erinnerung codiert wird. Die Prüfung der *Intercoderreliabilität* misst die Übereinstimmung zwischen zwei oder mehr Codierern. Die einfachste Art der Messung ist die Feststellung eines Quotienten: Man nimmt die Anzahl der übereinstimmenden Codierungen von zwei Codierern und teilt diese durch die Anzahl aller Codierungen. Wenn also zwei Codierer jeweils 100 Untersuchungseinheiten codiert haben, von denen 50 übereinstimmen, erhält man einen Quotienten von 0,50. Der Reliabilitätskoeffizient schwankt demnach zwischen 0 und 1. Bei der Interpretation des Koeffizienten ist zu berücksichtigen, um welche Kategorie es sich handelt. Bei der Verschlüsselung des Datums wird man sich mit einem Koeffizienten von 0,8 nicht zufrieden geben, während dieser Wert für die Verschlüsselung inhaltlicher Kategorien annehmbar ist. Der kombinierte Reliabilitätskoeffizient aller Kategorien sagt etwas über die Güte des Codebuches aus, nämlich darüber, wie gut es gelungen ist, die Codierer zu übereinstimmenden Codierungen anzuleiten. Mit der Konzeption des Kategorienschemas, das die Schulung der Codierer und die Überprüfung der Gültigkeit und Zuverlässigkeit des Messinstrumentes umfasst, liegt das Kernstück einer jeden Inhaltsanalyse vor. Zwei Fragen, die unmittelbar mit der Konzeption des Forschungsvorhabens zusammenhängen, sind die nach der richtigen Stichprobe und der Wahl der Untersuchungs- bzw. Analyseeinheiten.

9.6 Grundgesamtheit und Stichprobenziehung

Die Frage nach der optimalen Stichprobe richtet sich nach dem Forschungsvorhaben und den zur Verfügung stehenden Mitteln. Eng umgrenzte Fragestellungen können möglicherweise sogar mit einer Vollerhebung untersucht werden, beispielsweise die Berichterstattung über einen Staatsbesuch, der ja nur einige Tage dauert. Je breiter die Fragestellung, je länger der Untersuchungszeitraum ist und je mehr Medien in die Analyse eingeschlossen werden, desto stärker wird die Notwendigkeit einer Stichprobenziehung sein. Wer etwas über die Wirtschaftsberichterstattung der letzten dreißig Jahre sagen will, wird keine Vollerhebung machen, sondern muss eine wohlbegründete Stichprobe ziehen. Die Prinzipien der Stichprobenziehung bleiben allerdings gleich, die Stichprobe soll ein verkleinertes Abbild der Grundgesamtheit

sein, wobei die Grundgesamtheit hier durch die jeweils vollständige Berichterstattung definiert ist.

Bei der Inhaltsanalyse ist es, anders als bei der Befragung, auch prinzipiell möglich, größere Mengen von Zeitungen, Fernsehsendungen oder Bildmaterial zu untersuchen – vorausgesetzt, das Material ist irgendwann einmal archiviert worden. Man könnte theoretisch die gesamte Berichterstattung einer Tageszeitung von 1947 bis ins Jahr 2015 untersuchen, weil das Material physisch als Grundgesamtheit in einem Archiv vorhanden ist. Die Bestimmung einer Grundgesamtheit ist also bei der Inhaltsanalyse nicht nur praktisch, sondern auch theoretisch einfacher als bei der Methode der Befragung.³

- **Die Grundgesamtheit einer Inhaltsanalyse bestimmt sich, aus der Forschungsfrage abgeleitet, nach zwei Kriterien: dem zu untersuchenden Zeitraum und dem zu untersuchenden Medium.**

Während das erste Kriterium noch recht leicht zu bestimmen ist, fällt die Auswahl der Medien schon schwerer. Dies liegt hauptsächlich daran, dass unklar ist, welche Relevanz einzelne Publikationen für die Fragestellung haben, und die Menge der möglichen Medien viel zu groß ist. Wenn beispielsweise die AIDS-Berichterstattung untersucht werden soll, dann ist offensichtlich, dass man die aktuellen Nachrichtenmedien (Zeitung, Newssites, Radio- und Fernsehnachrichten) untersuchen sollte. Aber sind auch Frauen- oder Programmzeitschriften relevant? Auch muss man sicherlich aus den verfügbaren Medien wenige auswählen, also auch hier eine Stichprobe ziehen. Aber diese Auswahl gilt es zu begründen. Warum wurde ausgerechnet diese und keine andere Auswahl getroffen? Generell gibt es eine konzeptionelle Leitlinie, der im weitesten Sinn die *Hypothese der Koorientierung* zugrunde liegt: Journalisten orientieren sich an den Inhalten der Beiträge ihrer Kollegen und Journalisten einer Lokalzeitung orientieren sich insbesondere an Kollegen von überregionalen Blättern. Diejenigen Medien, die überregional verbreitet sind und einen gewissen Qualitätsstandard haben, werden von Journalisten anderer, eher lokal ausgerichteter Medien rezipiert. Somit erscheint es dann plausibel, dass sich der Tenor der Berichterstattung aus den großen überregionalen Tageszeitungen in den kleineren wiederfindet. Man kann argumentieren, dass für die Untersuchung überregionaler Themen deshalb die Auswahl der entsprechenden überregionalen Tages- bzw. Qualitätszeitungen ausreichend ist, man also nicht jedes regionale Blatt in eine Stichprobe aufnehmen muss. Solche Probleme müssen bedacht und gelöst werden. Wie immer man sich in so einem Fall entscheidet, die Begründung muss logisch und nachvollziehbar sein.

Ebenso wie bei Befragungen Individuen nicht in ihrer Gesamtheit, sondern nur in Bezug auf bestimmte Merkmale untersucht werden, betrachtet man nicht die Ge-

3 Zur Erinnerung: Bevölkerungsbefragungen sind unter diesem Aspekt theoretisch nicht sauber, weil die Grundgesamtheit weder physisch noch symbolisch dauerhaft vorliegt.

samtheit der Berichterstattung eines Mediums, beispielweise einer Tageszeitung, sondern nur den für die Untersuchung relevanten Teil, etwa Artikel des politischen Teils, den Stellenmarkt oder die Anzeigen. Die Grundgesamtheit ist demnach ebenfalls eine wohldefinierte Menge von N Elementen.

- **Bei Inhaltsanalysen werden einfache Zufallsauswahl, systematische Zufallsauswahl oder Klumpenstichproben angewendet. Üblich sind mehrstufige Verfahren.**

Eine reine Zufallsauswahl bedeutet, dass man aus allen Elementen der Grundgesamtheit (zum Beispiel aus dem politischen Teil eines Mediums von 1950 bis 2010) per Zufall eine Stichprobe zieht. Wenn die Merkmalsträger Zeitungsartikel wären, müsste man also zunächst eine Liste aller relevanten Artikel anlegen und dann daraus eine Zufallsstichprobe ziehen. Dies ist vor allem dann, wenn die Zeitungen gebunden oder auf Mikrofiche vorliegen, sehr unpraktisch. Ein weiterer Grund spricht in vielen Fällen gegen eine reine Zufallsauswahl. Wenn man zeitbezogene Analysen macht, also den Zeitverlauf untersuchen will, führen reine Zufallsstichproben sachlogisch dazu, dass beispielsweise zwei Zeitungsausgaben hintereinander gezogen werden und gleichzeitig an anderen Stellen größere zeitliche Lücken entstehen. Deshalb verwendet man bei zeitbezogenen Fragestellungen in aller Regel keine reine Zufallsauswahl, sondern bedient sich einer systematischen Zufallsauswahl, die in den meisten Fällen auch noch in mehreren Schritten erfolgt.

Nach einem festgelegten Schlüssel wird zunächst entweder jede n -te Zeitungsausgabe gezogen oder es werden alle Zeitungsausgaben einer bestimmten Teilperiode des Jahres (z. B. alle Septemberausgaben) ausgewählt. Wenn der Auswahlmodus und die Auswahlfrequenz feststehen, gelten diese natürlich für alle zu untersuchenden Medien in gleicher Weise. Innerhalb der untersuchten Ausgaben kann man dann noch einmal Stichproben von Beiträgen ziehen (jeden zweiten Beitrag) oder alle in Frage kommenden Beiträge verschlüsseln.

Einfache systematische Auswahlverfahren offenbaren allerdings bei bestimmten Fragestellungen Schwächen und müssen mit Vorsicht vorgenommen werden. Gerade wenn man Themenstrukturen im Zeitverlauf darstellen möchte, kann eine solche Auswahl zu verzerrten Ergebnissen führen. Wenn man zum Beispiel nur die Samstage einer überregionalen Tageszeitung berücksichtigt, würde diese Stichprobe die durchschnittliche Themenstruktur des Blattes nicht widerspiegeln, weil unter Umständen an Samstagen andere Themen im Vordergrund stehen als an einem Wochentag. Wenn man durch systematische Zufallsauswahl nur die Frühjahrs- und Herbstausgaben einer Frauenzeitschrift der letzten dreißig Jahre auswählt, wären Schlankheitskuren und Weihnachtsdekorationen die wichtigsten Themen. Man würde also nicht die tatsächliche Themenverteilung erheben und hätte vor allem starke saisonale Schwankungen in der Themenwahl nicht abgebildet. In der Praxis sind deshalb bezüglich des Zeitverlaufs geschichtete Stichproben üblich.

- **Mit einer Schichtung in Bezug auf den Zeitverlauf stellt man sicher, dass keine saisonal bedingten Faktoren Einfluss nehmen.**

Will man also Themenkarrieren bzw. Trends in der Berichterstattung über einen bestimmten Zeitraum innerhalb eines Mediums erheben, bildet man sogenannte *künstliche Wochen*. Man teilt die Grundgesamtheit, also zum Beispiel die Berichterstattung des politischen Teils einer überregionalen Tageszeitung (z. B. Süddeutsche Zeitung) der letzten zehn Jahre, zunächst in Wochenabschnitte auf. Somit erhält man in diesem Beispiel 10×52 Wochen. Dann wird per Zufallsauswahl für jede Woche ein Wochentag gezogen. Für die erste Woche wird zum Beispiel der Mittwoch gezogen. Für die zweite Woche steht der Mittwoch nicht mehr zur Verfügung (Zufallsauswahl ohne Zurücklegen); man zieht vielleicht den Montag. Auf diese Weise kommt für die ersten sechs Wochen (die Süddeutsche Zeitung erscheint lediglich an sechs Tagen pro Woche) eine *künstliche Woche* zusammen. So verfährt man, bis allen Wochen ein Tag zugeordnet ist. Die Berichterstattung der gezogenen Tage wird dann komplett untersucht, für alle eingeschlossenen Medien natürlich dieselben Tage.

Je nach Fragestellung müssen Stichproben so gewählt werden, dass ein Vorher und ein Nachher miteinander verglichen werden können. Immer wenn Veränderungen von Themen, Blattlinien oder Ähnliches interessant sind, braucht man als Referenzdaten diejenigen aus der Zeit vor einem Schlüsselereignis (wie Naturkatastrophen, terroristischen Anschlägen, Brandanschlägen, Epidemien u. Ä.) und die Daten aus einem gewissen Zeitraum nach dem Schlüsselereignis. Auch hier wird also durch die zentrale Forschungsfrage der Untersuchungszeitraum festgelegt. Im Folgenden können dann wieder künstliche Wochen gezogen werden.

9.7 Analyseeinheiten

Die Merkmalsträger in der Inhaltsanalyse nennt man Analyseeinheiten. Für diese Analyseeinheiten werden Merkmale erhoben. Eine häufig verwendete Analyseeinheit ist bei Printmedien der Artikel, also Meldungen, Berichte, Dokumentationen und ähnliche Formen der Berichterstattung. Der Artikel ist eine in sich geschlossene Einheit, die in der Regel von einem oder manchmal auch mehreren Urhebern zusammen verantwortet wird und in der ein bestimmtes Thema abgehandelt wird. Bei Radio- und Fernsehmaterial fällt die Definition dessen, was eine Analyseeinheit ist, schon etwas schwerer. In den Printmedien bietet schon allein die Anschauung, nämlich die räumliche Aufteilung einer Seite, die Grundlage der Definition. Will man dagegen Fernsehnachrichten untersuchen, gibt es verschiedene Möglichkeiten. Man kann zunächst einen thematisch abgegrenzten Sachverhalt als Trennkriterium einführen oder Nachrichten in Bezug auf ihre Stilformen voneinander abgrenzen. Ein Beispiel: Der Nachrichtensprecher beginnt mit einer Meldung, dass im Kosovo wieder gekämpft

wird. Zum selben Thema folgt dann eine sogenannte „Nachricht im Film“. Von einem anderen Redakteur läuft ein Film über die Kämpfe im Kosovo ab, danach ist der Tagesschausprecher wieder zu sehen. Er schaltet nach Bonn zu einem Politiker, der ebenfalls im Bild zu sehen ist, und befragt diesen Politiker zum Kosovo-Konflikt. Danach leitet er zu anderen Meldungen über. Im Fall einer thematischen Abgrenzung der Analyseeinheit gehören Anmoderation, Filmbericht und Interview zusammen. Unterteilt man die Analyseeinheiten nach Stilform, hat man drei Analyseeinheiten: die Anmoderation Kosovo, den Filmbericht Kosovo und das Interview Kosovo, die allerdings untereinander durch das gemeinsame Thema zusammengehören. Die thematische Abgrenzung stellt die Analyseeinheit höherer Ebene dar, die einzelnen Stilformen innerhalb dieses Themas sind Analyseeinheiten der zweiten Ebene. In diesem Fall spricht man von einer hierarchischen Anordnung der Analyseeinheiten.

Die Festlegung der Analyseeinheit ist wichtig, weil die Ergebnisse alle in Bezug auf diese Analyseeinheit zu sehen sind. Dies ist vergleichbar mit der Ergebnisdarstellung von Befragungen. Auch dort legt man eine Analyseeinheit fest, in der Regel Personen (Merkmalsträger!). Alle Aussagen über Häufigkeiten, Zusammenhänge oder Durchschnittswerte beruhen dann auf dieser Analyseeinheit. Bei hierarchisch aufgebauten Analyseeinheiten kann die Auswertung auf der oberen (hier: thematische Beiträge) oder auf der unteren Ebene (hier: Stilformen) erfolgen. Auch bei Zeitungen gibt es solche hierarchischen Anordnungen, wenn man beispielsweise innerhalb von Artikeln einzelne Aussagen als Merkmalsträger untersucht.

- **Die erste Ebene der Analyse ist meistens ein thematisch abgegrenzter Beitrag, die zweite Analyseeinheit können die Aussagen (Zeitung) oder Stilformen (Fernsehen) innerhalb des thematischen Kontextes sein.**

Man könnte sogar noch eine weitere Analyseeinheit einführen, nämlich das einzelne Wort innerhalb von Aussagen. Insbesondere bei computerunterstützten Inhaltsanalysen werden solche wort- oder phrasenbezogenen Zählungen durchgeführt. Die Häufigkeit des Wortes wird dann als Indikator für bestimmte Trends angesehen.⁴ Es gibt beispielsweise Inhaltsanalysen, die die Häufigkeit von Worten wie „Rezession“ in der Wirtschaftsberichterstattung in einem definierten Zeitraum untersuchen. Man erhält dann Verlaufskurven, die man mit der tatsächlichen Wirtschaftsentwicklung vergleichen kann. Dadurch kann man nachweisen, ob eine negative Wirtschaftsberichterstattung einen Einfluss auf die tatsächliche Wirtschaftslage (Inflationsrate, Staatsverschuldung etc.) ausübt oder ob die Bevölkerungsmeinung über die Wirtschaft sich an der tatsächlichen Lage oder an der Berichterstattung orientiert.

Das Wort ist also die einfachste Analyseeinheit. Worte bilden zusammen Sätze, Sätze formen sich zu Aussagen, und Aussagen lassen sich in Absätze zusammenfassen, die wiederum Artikel bilden. Jeder Baustein für sich kann eine Analyseeinheit

4 Die Vor- und Nachteile dieser Vorgehensweise werden ausführlich im nächsten Kapitel behandelt.

sein. Häufig laufen Inhaltsanalysen auf zwei hierarchisch gegliederten Ebenen ab. Zunächst misst man bestimmte Dinge auf Artekelebene, wie etwa Themenstrukturen. Innerhalb des Artikels misst man auf Aussage- oder Aussageträgerebene, um damit Hinweise auf die Kommunikatoren zu erhalten.

Bei Inhaltsanalysen zu Fernsehinhalten definieren sich die Analyseeinheiten analog. In der Regel werden dort unterschieden: Kameraeinstellungen, die sich zu Stilformen addieren, die sich wiederum zu thematischen Beiträgen zusammenfassen lassen. Die Einstellungen lassen sich durch die Art der Kameraführung relativ gut feststellen: Ein Schnitt, ein Zoom oder ein Schwenk beendet eine Einstellung und die nächste Einstellung beginnt. Stilformen, also Anmoderation, Nachricht im Film oder Interview, lassen sich auch relativ einfach identifizieren. Zusammengenommen bilden sie dann die thematische Analyseeinheit, also auch hier ein hierarchisches Schema. Komplizierter wird es, wenn Unterhaltungssendungen untersucht werden. Natürlich kann man die einzelne Sendung als ganz große Analyseeinheit definieren. Innerhalb dieser Sendungen fehlen dann aber häufig klare thematische und stilistische Abgrenzungen. Bei Talkshows könnte man die Frage oder Antwort eines Gastes als Analyseeinheit nehmen, bei *WETTEN, DASS ...?* eine Wette, bei Soap Operas einen Handlungsstrang.

9.8 Ablauf einer Inhaltsanalyse

Anhand eines Beispiels sollen noch einmal alle Phasen des Ablaufs einer Inhaltsanalyse vergegenwärtigt werden. Die einzelnen Schritte entsprechen wie bei der Befragung dem generellen Schema zum Ablauf des empirischen Forschungsprozesses.

9.8.1 Entdeckungszusammenhang: Ein Phänomen aus der sozialen Wirklichkeit wird in eine wissenschaftliche Fragestellung überführt

Der Auftrag bzw. das allgemeine Forschungsinteresse sei, die Berichterstattung über den Bundestagswahlkampf 1998 zu untersuchen. „Wie haben die Medien über die Kanzlerkandidaten und über den Wahlkampf selbst berichtet?“ wäre die generelle Fragestellung, die sich aus dem konkreten Phänomen der sozialen Wirklichkeit herleitet. Man übersetzt die allgemeine Problemstellung in eine wissenschaftliche Fragestellung, die in diesem Fall mittels einer Inhaltsanalyse empirisch überprüft werden soll. Es werden mehrere Hypothesen entwickelt, eine davon könnte lauten: „Es gibt eine zunehmende *Amerikanisierung* des Wahlkampfes bzw. der Berichterstattung über den Wahlkampf.“ Damit wäre ein Schwerpunkt gesetzt – nicht die Berichterstattung in ihrer gesamten Komplexität, sondern nur ein Ausschnitt davon wird untersucht.⁵

5 Näheres zu diesem Thema findet man bei Brettschneider (2000).

9.8.2 Begründungszusammenhang: Definition der Begriffe, Operationalisierung des theoretischen Konstruktes, Konzeption des Codebuchs, Codierung, Auswertung

An dieser Stelle taucht der erste Definitionsbedarf auf: Was versteht man unter Amerikanisierung? Welche Definitionen wurden bisher in der Literatur angeboten und welche Begriffsbestimmung soll im vorliegenden Fall zugrunde liegen? Das theoretische Konstrukt wird über die Zerlegung in verschiedene Teilkonstrukte expliziert. Diese könnten sein:

- Konzentration der Berichterstattung auf die Person des Kandidaten
- Vernachlässigung von Themen
- Verwendung emotionalisierender Bilder
- Stil eines „horse race journalism“⁶

Über diese Teilkonstrukte wiederum findet sich einiges in der Fachliteratur. Die einschlägigen Untersuchungen sowie deren Ergebnisse und Argumentationslinien werden geprüft und mit der angemessenen Tiefe für die Fundierung der eigenen Untersuchung dargestellt. Man muss sicher nicht jede Literaturstelle kennen und zitieren. Es geht bei der Darstellung des Forschungsstandes und der damit verbundenen theoretischen Grundlegung der eigenen Arbeit vielmehr darum, einhellige oder kontroverse Standpunkte zum Forschungsgebiet so zu nutzen, dass sie sachdienliche Hinweise für das Konzept der anstehenden Studie liefern. Das Studium der Literatur ist deshalb keine Pflichtarbeit, die irgendwie zur Abfassung einer Studie dazugehört, sondern schon in dieser recht frühen Phase der Konzeption die Grundlage und ein professionelles Korrektiv für die eigene Argumentation und den Aufbau der Untersuchung.

Die zentrale Aufgabenstellung bei einer Inhaltsanalyse ist die theorie- und empiriegeleitete Kategorienbildung, also die Entwicklung des Codebuchs. Theoriegeleitet meint unter anderem den Umstand, dass aus bereits vorhandener Literatur entnommen wird, welche Kategorien bisher zu diesem Thema entwickelt wurden. Häufig erfasst man aber das Spezifische eines Untersuchungsthemas besser, wenn man zusätzlich die eigenen Ideen nutzt, also sich das zu untersuchende Material zunächst einmal ansieht und so empiriegeleitet weitere Kategorien gewinnt.⁷ So könnte das Ergebnis einer Inhaltsanalyse sein, dass im Zusammenhang mit der Bundestagswahl 1998 im-

6 Dieser Begriff beruht auf einer Feststellung, die in den USA gemacht wurde. Im Wahlkampf konzentriert sich die Berichterstattung weniger auf die Themen des Wahlkampfes als vielmehr auf den Wettkampf zwischen zwei oder mehreren Kandidaten. Berichtet wird, wer um wie viele Längen „vorne“ liegt, ganz wie bei einem Pferderennen.

7 Die Entwicklung eines Kategorienschemas, die Auswahl der Medien, die Ziehung der Stichprobe und die Rekrutierung der Codierer laufen im Prinzip parallel ab. Was in der abstrakten Form des Schemas wie ein zeitliches Hintereinander wirkt, ist in der Praxis eher ein gleichzeitig ablaufender

mer auf 16 Jahre Kanzlerschaft von Helmut Kohl hingewiesen wird. Also wird man mit einer zusätzlichen Kategorie diesen Bereich abdecken.

- **Der Prozess der Kategorienbildung läuft sowohl deduktiv (theoriegeleitet aus der Literatur) als auch induktiv (empiriegeleitet aus eigener Anschauung) ab. Nur dadurch ist gewährleistet, dass man einen Gegenstandsbereich vollständig erfassen kann.**

Durch das Kategorienschema werden also Indikatoren gebildet und die theoretischen Konstrukte *operationalisiert*. Unter Hinzuziehung theoretischer und empirischer Überlegungen werden zugleich die Vollständigkeit und Trennschärfe der Kategorien hinsichtlich des theoretischen Konstruktes immer wieder geprüft. Das Ergebnis dieses Schrittes ist ein erster Codebuchentwurf. Wie bei einer Befragung folgt nun eine Phase, in der diejenigen, die später das Material codieren sollen, zum ersten Mal mit dem Messinstrument vertraut gemacht werden, um zu prüfen, ob das entwickelte Instrument überhaupt handhabbar ist. Man diskutiert in Codiererbesprechungen, ob die Kategorien wirklich trennscharf und die Codieranweisungen verständlich aufbereitet sind. Und dann werden die ersten Probecodierungen vorgenommen. Der Pretest erfolgt dann mit einer Teilmenge des tatsächlich relevanten Materials, wobei die Codierer während des gesamten Pretests in der Regel etwa 10 % des nachher tatsächlich zu codierenden Materials probecodieren.

Beim nächsten Codierertreffen werden die Ergebnisse der ersten Codierung diskutiert. Voneinander abweichende Codierungen geben meistens schnell Hinweise auf die größten Schwächen des Codebuchs. Kategorien werden dann vielleicht mit neuen Ausprägungen versehen, Codieranweisungen mit noch mehr Beispielen bestückt und deutlicher gemacht. Erneut werden die Codierer gebeten, Material zu codieren, das dann wiederum beim nächsten Treffen besprochen wird. Die Pretestphase ist also ein iterativer Prozess, an dessen Ende das fertige Codebuch steht. Der Pretest durchläuft dabei wiederholt drei Phasen: Codierung, Kontrolle und Anpassung des Codebuchs. Dies geschieht so lange, bis eine befriedigende Inter- und Intracoderreliabilität festgestellt werden kann, d. h. die Codierer fast immer die gleichen Codierungen vornehmen. Die Dauer dieses Prozesses hängt von der Kompliziertheit des Codebuchs und der Erfahrung der Codierer ab.

Nun folgt die eigentliche Codierung, bei der die Verteilung des Materials an die Codierer eine wichtige Rolle spielt. Es sei daran erinnert, dass wir mittels einer Stichprobe das zu codierende Material festgelegt hatten. Wichtig ist, dass das Untersuchungsmaterial *einer* Zeitung oder aus *einem* Zeitraum nicht von *einem* Codierer verschlüsselt werden darf, sondern dass das zu codierende Material möglichst mit einem Zufallsalgorithmus auf die Codierer verteilt wird. Warum dies sinnvoll ist, ist un-

schwer zu verstehen. Menschen lernen und reifen, so auch Codierer. Sie entwickeln (ungewollt) eine besondere Art und Weise, die Analyseeinheiten zu verschlüsseln. Auch wenn eine gute Reliabilität während der Pretestphase gemessen wurde, ist nicht zu verhindern, dass sich durch diese Reifungsprozesse bei einzelnen Codierern Fehler einschleichen, so dass durch individuelle Interpretationen des Materials keine reliable Anwendung der Kategorien erfolgt. Deshalb werden bei größeren Forschungsvorhaben auch während der Codierung Reliabilitätstests durchgeführt. Wenn *ein* Codierer das Material *eines* Mediums, der zweite das zweite Medium verschlüsselt, können systematische Fehler auftreten, die am Ende die ganze Untersuchung unbrauchbar machen. Variiert man dagegen das Untersuchungsmaterial zufällig oder systematisch, werden auch Codierfehler zufällig auftreten und vermutlich innerhalb der Toleranzen der Irrtumswahrscheinlichkeit liegen. Sowohl die Zeiträume als auch die Medien müssen also so auf die Codierer verteilt werden, dass sich keine Häufungen bei bestimmten Medien oder bestimmten Zeiträumen ergeben.

Der letzte Schritt in der Phase des Begründungszusammenhangs ist die Auswertung, d. h. die Dateneingabe (sofern sie nicht im Zuge der Codierung am PC bereits erfolgte), die Datenbereinigung und die komplette Analyse plus Ergebnisdarstellung. Hier wird die zentrale Fragestellung mit den Datenanalyseverfahren beantwortet bzw. der Rückbezug vom zahlenmäßigen Ergebnis auf die soziale Realität vorgenommen. Kam es nun zu einer Amerikanisierung des Wahlkampfes? Eine Antwort könnte darin bestehen, dass sich in der Berichterstattung zum Bundestagswahlkampf 1998 23 % aller Aussagen auf den Wahlkampf selbst („horse race“) bezogen haben. Die Zahl sagt an sich noch nichts über Amerikanisierung, gewinnt aber dann an Wert, wenn man diese fiktiven 23 % mit Ergebnissen früherer Studien vergleicht (oder selbst einen Vergleich zum Beispiel zwischen der Berichterstattung von 1994 und 1998 gemacht hat). Die signifikante Zunahme des Zahlenwertes würde dann die These der Amerikanisierung in Bezug auf den Indikator „horse race journalism“ stützen. Man hat also ganz im Sinn des Zieles einer Inhaltsanalyse systematisch und quantitativ die Inhalte von massenmedial verbreiteten Texten erfasst und somit von einem manifesten Text auf einen nicht-manifesten Kontext geschlossen.

9.8.3 Verwertungszusammenhang: Der praktische und theoretische Gehalt der Studie wird genutzt

Die Studie erhält nur dann eine Relevanz, wenn die Ergebnisse an die Öffentlichkeit gelangen. Üblicherweise versucht der Forscher, einen Aufsatz in den einschlägigen Fachpublikationen unterzubringen, ein Buch zu schreiben und vielleicht seine Erkenntnisse auf einer Tagung zur Diskussion zu stellen. Ziel der Forschung ist es ja, mit jeder Studie zur Entwicklung der theoretischen Grundlagen des jeweiligen Fachs beizutragen. Auf der Seite der Öffentlichkeit soll die Studie einen bestimmten gesellschaftlichen Prozess (zum Beispiel die „Entpolitisierung der Politik“) kritisch beglei-

ten. Sie liefert im Idealfall das Material für eine breite Diskussion über dieses Phänomen. Dass die Argumente der Wissenschaft dann durchaus instrumentell eingesetzt werden, steht wieder auf einem anderen Blatt.