

## Kapitel 5

### Wie man bei qualitativen Analysen vorgeht

Dieses Kapitel versucht auf wenigen Seiten eine Übersicht über typische Phasen bei der Analyse qualitativer Daten zu vermitteln. Dabei werden zum besseren Verständnis einzelne Schritte als kennzeichnend für diese Phasen herausgehoben, auch wenn der analytische Prozess in der Realität immer zyklisch abläuft, man also zeitweilig, zumindest "im Hinterkopf", mit Tätigkeiten befasst ist, die andere Phasen charakterisieren. Weiter versucht dieses Kapitel zu skizzieren, wie man die Möglichkeiten nutzen kann, die mit AQUAD verfügbar werden. Einzelheiten des praktischen Vorgehens werden dann in den späteren Kapiteln genauer erläutert. Mehr als eine erste Übersicht über das methodische Vorgehen bei qualitativen Analysen und Möglichkeiten des Computereinsatzes dabei kann dieses Manual hier nicht leisten. Für ausführlichere Darstellungen des Ansatzes der qualitativen Analyse werden interessierten Lesern/innen folgende Bände empfohlen:

- Allgemeine Beiträge zum Einsatz des Computers in der qualitativen Forschung findet man beispielsweise bei Tesch (1990), Huber (1992), Kelle (1995), Fielding und Lee (1998) oder Lissmann (2001).
- Miles und Huberman (2. Ausgabe 1995) geben eine ausführliche Einführung in die interpretative Analyse qualitativer Daten. Sie stützen sich dabei auf zahlreiche Beispiele. Weitere konkrete Beispiele einer Vielfalt qualitativer Studien aus der erziehungswissenschaftlichen Forschung finden sich in Bos und Tarnai (1998) sowie Schratz (1993), aus der psychologischen Forschung bei Kiegelmann (2001, 2002, 2003).
- Eine Einführung speziell in Techniken der theoriekonstruierenden qualitativen Analyse geben Strauss und Corbin (1990).

Diese Auflistung ist keineswegs erschöpfend, sondern enthält nur ausgewählte Literatur, die zur Einführung besonders geeignet erscheint und in meinen Seminaren erprobt wurde. Alle Bände eröffnen guten Zugang zu weiterführenden Arbeiten. Die folgende Übersicht ist eine stark verkürzte Bearbeitung meines eigenen Beitrags in Huber (1992); für Details und weitere Beispiele sei auf diesen Band verwiesen.

## 5.1 Wie läuft eine qualitative Analyse ab?

Als typische Phasen kann man bei der Analyse qualitativer Daten die Reduktion der ursprünglichen Daten, die Rekonstruktion inhaltlicher Zusammenhänge und den Vergleich der Befunde mit dem Ziel der Generalisierung unterscheiden. Insbesondere in psychologischen Studien interessiert häufig der induktive Schluß von Daten *einer* Person auf Regelmäßigkeiten ihres Erlebens und Verhaltens als generalisierter Befund. Ob sich über mehrere Personen bzw. Beobachtungssituationen hinweg Gemeinsamkeiten feststellen lassen, interessiert zwar auch, aber erst in einem späteren Untersuchungsstadium.

- Die erste Arbeitsphase besteht in der *Reduktion* der Datenmenge und gleichzeitig der darin enthaltenen Vielfalt des Ausdrucks., d.h. der sprachlichen Formulierungen in Texten und zusätzlich der nonverbalen Ausdrucksformen in Videoaufzeichnungen. Dazu werden mehr oder weniger umfangreiche Datensegmente (Text-, Tonband- oder Videoabschnitte bzw. Bildausschnitte) ermittelt, denen jeweils definierbare Bedeutung zugeschrieben wird. Daraufhin wird ein Code als Abkürzung oder Benennung des Bedeutungsgehalts an das Datensegment angehängt. Im folgenden werden diese Codes als Repräsentanten der Datensegmente oder der *Bedeutungseinheiten* der Dateien benutzt. Grundsätzlich handelt es sich hier um einen Kategorisierungsprozess, bei dem die genauen Kategorien meist erst während der Interpretation auftauchen. Sie können aber auch aus einem bereits bestehenden Kategoriensystem übernommen werden. Das hängt ganz von der epistemologischen Orientierung des Forschers/der Forscherin ab.
- In der zweiten Phase rekonstruiert man aus diesen Bedeutungseinheiten die subjektiven Bedeutungssysteme der Datenproduzenten, d.h. der Interviewpartner, Tagebuchschreiber, Beobachter usw. Bei der *Rekonstruktion* sucht man nach regelhaften, für die Datenproduzenten und/oder ihre Situation kennzeichnenden Verknüpfungen von Bedeutungseinheiten innerhalb der Dateien.
- In der dritten Phase schließlich werden durch *Vergleich* der auf diese Art und Weise erhaltenen Einsichten in individuelle Bedeutungssysteme Invarianzen oder generelle Zusammenhänge erschlossen (vgl. Ragin, 1987).



Wichtig ist der Hinweis, dass diese Phasen nicht streng abgrenzbar und linear aufeinander folgen, sondern sich in der Regel überlagern und zyklisch verlaufen (vgl. Shelly & Sibert, 1992).

Bereits während der Reduktion denkt man beispielsweise an implizite Theorien der Datenproduzenten, man beginnt nach der Sichtung einiger weniger Dateien ständig zu vergleichen; dabei tritt dann vielleicht bei Person/Datei B ein Aspekt deutlich hervor, den man bei Person/Datei A übersehen hatte - also durchläuft man dort mindestens partiell nochmals die Reduktionsphase usw. In jeder dieser Phasen ist immer auch wichtig, die Gültigkeit der Kategorisierungen deduktiv wieder abzusichern, indem auf Besonderheiten der vorausgehenden Analysephase rückgeschlossen und nach entsprechenden Belegen gesucht wird.

## 5.2 Wie reduziert man qualitative Daten?

Das Prinzip der Reduktion ist scheinbar einfach, doch erweist sich die Anwendung als sehr arbeitsintensiv, zeitraubend und fehleranfällig: Umfangreiches verbales Material muss auf die darin enthaltenen Bedeutungseinheiten reduziert werden. Tesch (1992) beschreibt dieses Prinzip als Suche und Markierung der für die Forschungsfrage relevanten Dateisegmente, die dann noch mit einer Kurzbezeichnung (Code für eine Bedeutungskategorie) versehen werden. Tesch hat auch die "traditionellen" Vorgehensweisen gegenübergestellt, nämlich das Herstellen von Transkriptions-Ausschnitten im wörtlichen Sinn und die Übertragung relevanter Passagen auf Karteikarten. Beide Verfahren verursachen immensen "mechanischen" (im Unterschied zum unerlässlichen "konzeptuellen") Arbeitsaufwand. Computerprogramme helfen nun nicht nur die Datenmenge zu reduzieren, sondern auch die dabei sonst anfallenden mechanischen Arbeiten. Statt mit auf Kartenstapel verteilten Ausschnitten gelegentlich weitschweifiger Texte kann man die weitere Analyse mit den Codierungen dieser Daten-segmente selbst durchführen, d.h. man arbeitet nach der Datenreduktion mit Kurzbezeichnungen für die Kategorien und Angaben der Fundstelle der zugeordneten Bedeutungseinheiten in den Dateien weiter. Benötigt man gelegentlich den vollständigen originalen Dateiausschnitt, lässt sich dieser mit Computerunterstützung sehr rasch auffinden.

Wenn die Ausgangsdaten der Analyse als Video- oder Audioaufzeichnungen vorliegen, steht vor allen diesen Problemen üblicherweise noch die Notwendigkeit, die Daten erst einmal durch Transkribieren, also Umwandlung in eine Textversion, der genauen Analyse zugänglich zu machen. AQUAD 6 hilft Zeit und Geld zu sparen, da die Software direkte Reduktion der Daten auf Codierungen ohne den Umweg über Transkriptionen ermöglicht. Für kritische Inhalts-Sequenzen und ihre spätere Darstellung in Berichten, Aufsätzen usw. empfiehlt es sich, diese ausschnittsweise zu transkribieren. Dies kann mit der "Memo"-Funktion (s. Kapitel 8) in AQUAD durchgeführt werden.

Die kritische Frage in dieser Anfangsphase qualitativer Analyse lautet: *Wie und wo findet man Bedeutungseinheiten in den Dateien?* Für Anfänger der qualitativen Analyse ebenso wie bei der Einarbeitung in neue Inhaltsbereiche stellt sich diese Frage immer wieder neu. Weber (1985) referiert für Textdateien sechs gebräuchliche Möglichkeiten, generell Einheiten festzulegen, nämlich Wörter, Wortbedeutungen, Sätze, Thematiken, Abschnitte, den Gesamttext (möglich z.B. bei Schlagzeilen, Zusammenfassungen, kurzen Leserbriefen usw.) als Einheit der Analyse zu wählen. Diese Festlegungen können jedoch nicht mechanisch getroffen werden, sondern sie benötigen selbst wieder qualitative Vorentscheidungen. Bei Videos sind für die nicht-sprachlichen Inhalte zusätzliche Entscheidungen nötig (s. dazu die Codierung der Beispielfideos "Baby-1.avi" - "Baby\_3.avi". Bei non-verbale Daten, aber auch auf der Ebene einzelner Wörter, ganz offensichtlich aber bei den komplexeren Alternativen wie Wortbedeutungen, Sätze usw. als analytischen Einheiten, muss vorab eine Erkenntnis oder Hypothese stehen, dass die gewählte Einheit relevant im Sinne der Forschungsfrage ist. Weiter sind qualitative Zusatzentscheidungen nötig, beispielsweise darüber, welche Wörter synonym gebraucht werden oder welche idiomatischen Ausdrücke ebenfalls die gleiche Bedeutung für die Produzenten der zu analysierenden Texte haben.

Mit der Bestimmung der Bedeutungseinheit ist seit den Anfängen einer breiten sozialwissenschaftlichen Rezeption textanalytischer Verfahren ein wesentlicher Unterschied zwischen quantitativen und qualitativen Ansätzen festgelegt. Im gleichen Jahr, 1952, erschienen je richtungweisende Artikel von Berelson und von Kracauer. Während für Berelson die Inhaltsanalyse zur systematischen, objektiven und quantitativen Erfassung *manifesten* Kommunikationsinhalte dient, versucht die qualitative Inhaltsanalyse nach Kracauer unabhängig vom manifesten Textinhalt die verborgenen, *latenten* Bedeutungskategorien des Textes offenzulegen. Beide Autoren knüpfen mit gegensätzlichen Positionen an eine Auseinandersetzung an, die Thomas und Znaniecki (1918) mit einer inzwischen klassischen soziologischen Analyse der Briefe polnischer Emigranten aus Amerika 35 Jahre früher ausgelöst hatten. Die Frage nach der geeigneten analytischen Einheit - Wörter vs. Bedeutungen - ist offensichtlich mit der Zielsetzung der Textanalyse konfundiert.

Auf der Verfahrensebene schließen sich die beiden Zugangsweisen nicht grundsätzlich gegenseitig aus. Gerade mit Computerunterstützung kann man für qualitative Text- und Bildanalysen ver-

schiedene Vorgehensweisen des quantitativen Ansatzes mit Gewinn einsetzen, um die eigenen Interpretationsbemühungen zu unterstützen.

Grundsätzlich sollte man sich für die Auswahl der Bedeutungseinheiten bei qualitativer Datenanalyse immer vergegenwärtigen, dass unabhängig davon, ob man Erleben und Handeln von Menschen aus ihrer Sicht, d.h. aus ihren eigenen verbalen Darstellungen *verstehen* will oder ob man bestimmte ihrer aufgezeichneten Handlungsweisen *erklären* möchte, indem man sie auf das Bezugssystem der subjektiven Theorien dieser Menschen zurückführt, man die analytischen Einheiten *während* der Interpretation der Daten erschließen muss. Nur bei diesem induktiven Vorgehen ist gewährleistet, dass man Zugang zur subjektiven Weltsicht der Gesprächspartner bzw. der Beobachteten findet und nicht lediglich einige Teilaspekte davon, möglicherweise aus dem subjektiven Sinnzusammenhang gelöst, im analytischen Raster des Forschers hängenbleiben. Diese Strategie entspricht dem Ansatz der "grounded theory", der empirisch abgesicherten Theoriebildung bei Glaser und Strauss (1979).

Allerdings verlangt dieses Vorgehen beträchtlichen Arbeitsaufwand, sobald man mehr als einen einzigen Datensatz zu analysieren hat. Da man in einem weiter fortgeschrittenen Stadium der Analyse die Einzelergebnisse vergleichbar machen will, muss man - und das meist mehrfach - die Bedeutungseinheiten und ihre Codierung in allen Dateien auf Kohärenz untersuchen und unter Umständen modifizieren. Miles und Huberman (1994) haben deshalb einen Kompromiss vorgeschlagen, bei dem man vor der Lektüre der Texte/Sichtung der Dateien einen sehr allgemeinen, nicht inhaltsspezifischen *Orientierungsrahmen* für die Suche nach Bedeutungseinheiten festlegt, innerhalb dessen dann entsprechend der Bedingungen der einzelnen Datei über spezifische Einheiten entschieden wird. Dieser Kompromiss erscheint insofern vertretbar, als Datenreduktion auch schon *bei Beginn* einer Untersuchung, ja schon bei ihrer Planung anfängt.

### 5.3 Schema des Ablaufs der Datenanalyse

(1) Es ist unerlässlich, sich zu Beginn erst einmal mit dem untersuchten Gegenstandsbereich aus der Perspektive der Befragten oder Beobachteten vertraut zu machen - vor allem, wenn man nicht selbst in der Befragungssituation anwesend war, z.B. als Interviewer/in mit den Befragten gesprochen hat, oder wenn man nicht selbst bei Videoaufzeichnungen mitgewirkt hat. Das heißt anders herum ausgedrückt, man sollte nicht versuchen, schon mit dem ersten Datensatz ein differenziertes Kategoriensystem zu erarbeiten. Mit Sicherheit hätte es nur für diesen Datensatz, z.B. für das aktuelle Interview, Gültigkeit und müsste bereits bei der Analyse des zweiten Datensatzes grundlegend revidiert werden. Statt dessen empfiehlt es sich, aus einer größeren Zahl von Datensätzen entweder per Zufall 8-10 Transkriptionen oder Aufzeichnungen auszuwählen oder analog zur Fallauswahl bei Einzel-fallanalysen nach dem Prinzip des "theoretical sampling" gezielt einige wenige Datensätze aus dem gesamten Datenmaterial herauszugreifen.

Wenn wir beispielsweise Aufzeichnungen des Mathematikunterrichts an verschiedenen Schulen analysieren wollen, müssen wir vielleicht mit Unterschieden abhängig von Klassenstufe, Schulart und Unterrichtserfahrung der beobachteten Lehrpersonen rechnen. Nach diesen Kriterien würden wir einige wenige Datensätze auswählen und daran eine erste Orientierung erarbeiten. Vielleicht haben wir aber von vorne herein Klassen beobachtet, in denen unterschiedlich gelehrt und gelernt wird. Dann würden wir Datensätze (auch) nach diesem Merkmal zur Orientierung heranziehen.

Diese erste Phase modifiziert ganz offensichtlich die Empfehlung von Miles und Huberman (1994; s.o.), sich schon vor der Lektüre der Texte bzw. Sichtung der Dateien auf einen allgemeinen, nicht inhaltspezifischen *Orientierungsrahmen* für die Suche nach Bedeutungseinheiten festzulegen. Wir empfehlen hier, diesen möglicherweise aufgrund der Forschungsfrage in Umrissen vorab verfügbaren Orientierungsrahmen mit Hilfe einer Datenstichprobe zu überprüfen und zu differenzieren.

(2) Wenn wir Dateien nach dem Prinzip des "theoretical sampling" ausgewählt haben, mussten wir uns schon eingehend Gedanken über generelle Merkmale der Datensätze machen. Wir haben überlegt, wodurch das Profil der einzelnen Datensätze bestimmt wird. Im Fall des Beispiels der Daten aus "Unterrichtsbeobachtungen" also etwa durch das Alter der Schüler, Schulart, Berufserfahrung der Lehrkräfte, vielleicht auch durch Klassengröße, didaktische Orientierung der Lehrkräfte, usw. Diese Merkmale sollten wir in einem Memo (s. Kap. 8) festhalten und später als "Profilcodes" (oder "singuläre", d.h. nur einmal pro Datensatz zu seiner Gesamtcharakterisierung verwendete Kategorien - am besten gleich ganz am Dateianfang - zur Codierung (s. Kap. 6) verwenden.

Falls wir dagegen einfach nach Zufall einige Datensätze zur Orientierung herausgegriffen haben, wäre es eine gute Idee, sich jetzt schon einmal vorab zu überlegen, welche Merkmale der befragten/beobachteten Personen und der jeweiligen Situation im Zusammenhang mit der Forschungsfrage der Untersuchung als "Profilmerkmale" für die weitere Analyse bedeutsam werden könnten.

(3) Die Datenauswahl (siehe Punkt 1) lesen wir/sehen wir zunächst ohne aufwändige Codierung an und versuchen, den Gesamtzusammenhang zu verstehen, wesentliche Abschnitte zu unterscheiden, Gemeinsamkeiten und Gegensätze herauszufinden und so unseren Orientierungsrahmen abzustecken. Entscheidend ist, dass wir unsere Deutungen und Einfälle dabei in Memos (s. Kapitel 8) festhalten.

Am Ende dieses Arbeitsschritts empfiehlt es sich, jeweils eine kurze (!) thematische Zusammenfassung zu schreiben, am besten wieder als Memo. Darin notieren wir unseren "ersten Eindruck" davon, worum es in diesem Datensatz (Interview, Aufzeichnung, usw.) insgesamt geht, was wir zu diesem frühen Zeitpunkt der Analyse als zentrale Bedeutung dieses Datensatzes erachten. Außerdem sollten wir in einem eigenen Memo den - immer vorläufigen - Orientierungsrahmen skizzieren.

(4) Dann müssen wir eine wichtige Entscheidung über unsere grundsätzliche Interpretationsstrategie treffen: Wollen wir, wie in Punkt 3 durch die thematische Zusammenfassung schon angebahnt, nach möglichst umfassenden Bedeutungseinheiten suchen, diese codieren und sie anschließend schrittweise und in wiederholten Durchgängen differenzieren? Oder wollen wir gleich zu Beginn, eingeschränkt

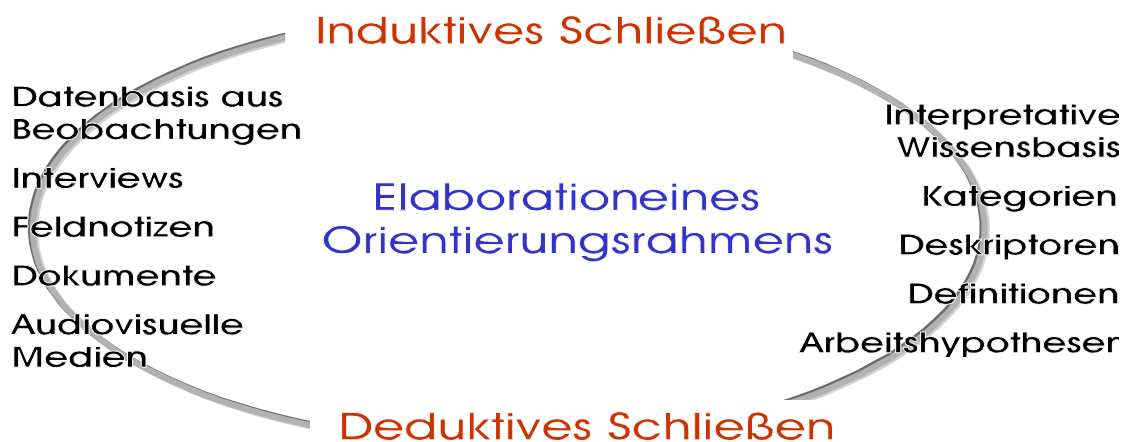
höchstens durch den generellen Orientierungsrahmen, auf *alle* Details achten, sie mit Hilfe von Codes markieren und dann in weiteren Codierungsdurchgängen die vielen Einzelheiten in generellen Kategorien zusammenfassen, um die Datensätze vergleichbar zu machen? Die erste Strategie bewegt sich von Allgemeinen zum Besonderen ("Differenzierung"), die zweite Strategie vom Besonderen zum Allgemeinen in den Datensätzen ("Generalisierung").

Es hängt natürlich auch von der Forschungsfrage und der Erfahrung der Forschenden mit dem Gegenstandsbereich ab, welche Strategie vorzuziehen ist. Grundsätzlich möchten wir aber Anfängern der qualitativen Datenanalyse von der Generalisierungsstrategie *abraten*, bei der sie zuerst auf alle spezifischen Merkmale achten und sie dann in weiteren Schritten zu abstrakteren, generelleren Kategorien zusammenzufassen versuchen. Damit hält man sich meist erfolgreich von tiefergehenden, am Anfang schwierigen, theorieorientierten analytischen Überlegungen ab und läuft Gefahr, sich lange Zeit in oft vordergründiger Geschäftigkeit zu ergehen. Man produziert Codes, man ist tätig und somit beruhigt - auch wenn man dann später feststellen muss, dass die Aktivitäten nicht unbedingt zielführend waren. In einem Seminar hat eine Gruppe von Studierenden beispielsweise Codes für Schlüsselwörter (nach Meinung der Studierenden) im Text vergeben, also im Grunde die Daten verdoppelt. In einem anderen Fall hat ein Anfänger für die Interpretation seiner Interviews rund 1500 (kein Druckfehler!) Codes erzeugt und dann natürlich vor lauter Bäumen keinen Wald mehr gesehen.

Daher erscheint es uns gerade für Anfänger aussichtsreicher, von Anfang an das Prinzip des "ständigen Vergleichens" im Ansatz der "grounded theory" (Glaser & Strauss, 1979) ernst zu nehmen und zunächst innerhalb der einzelnen, dann zwischen den Datensätzen nach Übereinstimmungen und Unterschieden zu suchen. Mit dieser Orientierung ist man weitgehend sicher, sich nicht in Oberflächendetails zu verlieren, sondern zielstrebig auf Einsicht in Bedeutungsstrukturen hinzuarbeiten. Die folgende Tabelle stellt die Ansätze nochmals einander gegenüber:

Strategie:	Differenzierung	Generalisierung
Codierung beginnt mit	Suche nach generellen Kategorien	Suche nach spezifische Aspekte
weiteres Vorgehen	Differenzierung zum Aufdecken spezifischer Unterschiede	Generalisierung zum Erkennen von Gemeinsamkeiten

Allerdings schließen die beiden Ansätze sich nicht aus, sondern bedingen einander wechselseitig. Die Interpretation qualitativer Daten ist eben kein linearer, sondern ein zyklischer Prozess:

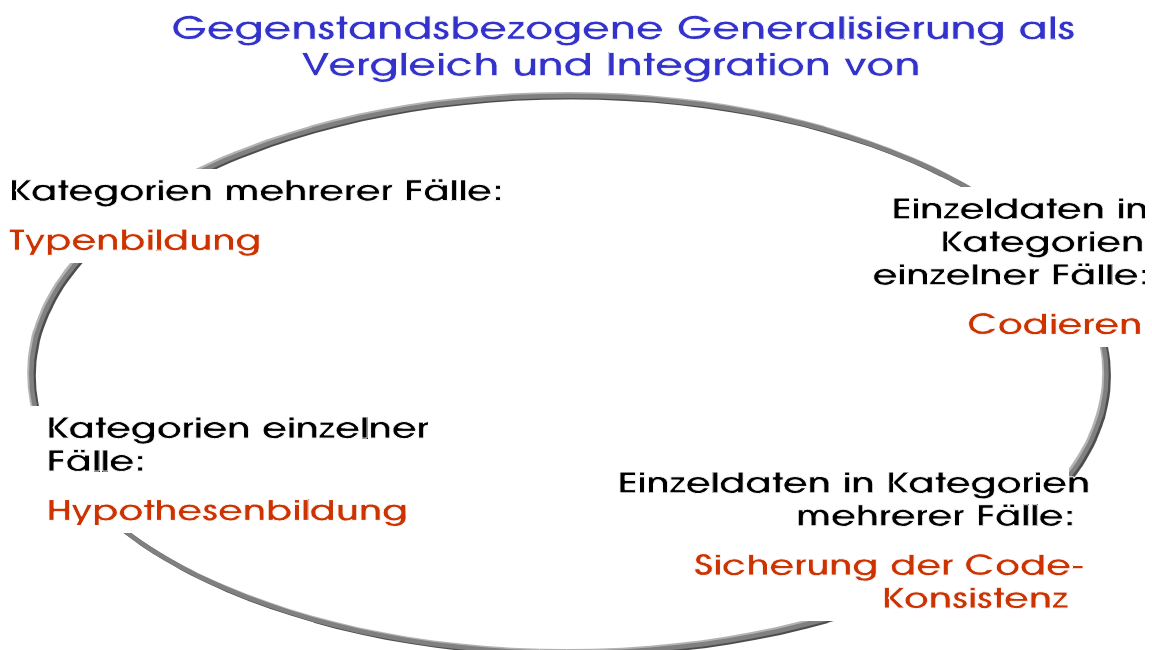


Shelly & Sibert (1992) haben unter Rückgriff auf Ausführungen Deweys zum Zusammenhang von induktivem und deduktivem Denken diesen Zyklus so beschrieben, wie die Abbildung oben es wiedergibt. Der Empfehlung für Anfänger, von generellen Kategorien auszugehen, ist unter dieser Perspektive allerdings entgegenzuhalten, dass man eher die Chance hat, etwas Neues in den Daten zu entdecken, wenn man kleine Datensegmente, sozusagen Zeile für Zeile interpretiert - allerdings auf die Gefahr hin, dabei das Generelle und Wesentliche zu übersehen.

Folgt man unserer Empfehlung, sich den Daten mit einer Strategie der allmählichen Differenzierung anzunähern, liegt der Akzent zunächst auf der "interpretativen Wissensbasis". Dieses Deutungswissen kann sich natürlich aus den Erfahrungen im Feld oder bei erster Sichtung des Datenmaterials induktiv entwickelt haben, beispielsweise in Form einiger sehr allgemeine Kategorien (Untersuchung in einem Kindergarten: "Die Erzieherin lenkt die Aufmerksamkeit der Kinder"; "Sie unterstützt bei Schwierigkeiten"; "Sie regt zu Verbalisierungen an" usw.) oder in Form von Arbeitshypothesen, die schon bei der Formulierung der Forschungsfrage maßgebend waren. Notwendigerweise differenziert sich die Datenbasis, wenn wir Aufzeichnungen von mehreren Personen und vielleicht aus unterschiedlichen Situationen analysieren.

Wendet man die Strategie der Generalisierung an, beginnt man die Analyse also mit spezifischen Bedeutungen, die unter interpretativer Perspektive in den Daten gleichsam aufscheinen, werden Wegweiser von spezifischen Interpretationen zu allgemeinem Deutungswissen besonders wertvoll.

Die Grundregel der "gegenstandsbezogenen Theoriebildung" (grounded theory), das "ständige Vergleichen" wurde als Leitfaden der Generalisierung charakterisiert. Auch diese Vergleichsprozesse verlaufen zyklisch, wie Shelly & Sibert (1992) betont haben. Angelehnt an ihre Einteilung lassen sich das ständige Vergleichen und seine Ziele im Fortgang der Analyse beschreiben, wie er in der folgenden Abbildung zusammengefasst ist:



(5) Nach der grundsätzlichen Entscheidung über das Vorgehen bei der Analyse werden die Überlegungen konkretisiert: man beginnt eine Stichprobe von Texten zu codieren. Dabei wird man notwendigerweise immer wieder zwischen dem Vergleichen und Integrieren von Einzeldaten im einzelnen Datensatz und dem Vergleich mehrerer Datensätze auf die dabei erzeugten Kategorien wechseln. Man bewegt sich also in kleineren, weniger ausgreifenden Analysezyklen innerhalb des zyklischen Prozesses, zunächst um die Konsistenz der Codierung sicherzustellen.

(6) Die dabei gewonnenen Codierungsregeln wendet man schließlich auf die insgesamt verfügbaren Dateien an. Dabei wird man regelmäßig auf Widersprüche und Ausnahmen stoßen, die zur Modifikation der Codierungsregeln anregen. In manchen Fällen kann es aufgrund solcher Einsichten notwendig werden, den Zyklus der Datenreduktion/-interpretation zu verlassen und in einer weiten zyklischen Bewegung nochmals in die Phase der Datenerhebung einzutreten.

(7) Memos zur Unterstützung der Hypothesenbildung und Typenbildung werden sich während des gesamten Prozesses ergeben. Gegründet auf solche Einfälle und die im Lauf des Codierens gewonnenen Einsichten konzentriert sich die Analyse in fortgeschrittenem Stadium vor allem darauf, systematische (Bedeutungs-)Zusammenhänge zwischen kritischen Kategorien zu erhellen. Dies wurde in Punkt 4 als Integration von Kategorien in "Hypothesen" beschrieben; z.B. "Wenn die Erzieherin E das Ereignis A oder B beobachtet, dann wendet sie die Handlungsmöglichkeit X an." Oder: "Wenn Lehrer L von mangelnder Motivation seiner Schüler spricht, folgt ein Bezug auf den Einfluss der Massenmedien und die Rolle der Eltern".

Die Erfahrungen in dieser Phase, beispielsweise von Widersprüchen in den Interpretationen, können ebenfalls zum Anlass für eine analytische Schleife zurück durch die Stufen des Codierens und der Konsistenzabsicherung (s. Punkt 4) sein. Auch nochmalige Datenerhebung mit modifizierten Fragen, Beobachtungssituationen etc. kann sich als notwendig erweisen.

(8) Eine ausreichende Zahl von Fällen vorausgesetzt, kann man abschließend versuchen, die einzelnen Fälle auf der Basis des Vergleichs und der Integration (theoretisch begründet) ausgewählter Kategorien in Typen zusammenzufassen. Als Ergebnis erhalten wir eine Unterscheidung der Fälle nach typischen Merkmalskombinationen.

(9) Ein genereller Hinweis soll hier nicht versäumt werden, der für den Prozess der qualitativen Analyse insgesamt gilt: Qualitative Analyse hat etwas mit der "Qualität" der in den Daten erfassten Ereignisse und Zustände zu tun. Die Codierungen dürfen also in der Regel nicht nur die Tatsache des Auftretens bestimmter Ereignisse neutral erfassen, sondern müssen die Qualität der Ereignisse mit repräsentieren. Beispielsweise würde es für Hypothesenbildung und Typenbildung in einer biographischen Interviewstudie mit Jugendlichen nicht ausreichen, lediglich mit einem Code "Eltern" die Dateisegmente zu markieren, in denen sie über ihre Eltern sprechen. Welche Jugendlichen kommen dabei nicht auf ihre Eltern zu sprechen?

Ein Merkmal, das ohne Variation bei allen gleich auftritt, kann nichts zum "ständigen Vergleich" beitragen, ist mithin nicht tauglich für weiterführende Unterscheidungen und Erklärungen. Wir müssen die Rede über die Eltern also von Anfang an qualifizieren: Wie reden die Jugendlichen über ihre Eltern? Wenig differenzierte Abstufungen, z.B. positiv / indifferent / negativ, reichen dabei zu Beginn der Analyse meist für den Zweck der Suche nach Gemeinsamkeiten und Unterschieden aus.

Es ist ein menschlich sympathischer Zug, wenn Anfänger der qualitativen Analyse sich mit solchen Wertungen insbesondere von personnahen Merkmalen oder Äußerungen zurückhalten, wenn sie vorschnelle Wertzuschreibungen ("positiv"; "negativ") vermeiden - aber genau um Qualifizierungen geht es bei der qualitativen Analyse. Spätestens bei der Prüfung der Konsistenz der Codierungen und der Festlegung von Codierungsregeln sollte man also "neutrale", nicht-qualifizierende Codes darauf untersuchen, ob sie Anlass dafür sein müssen, eine Rückwärts-Schleife im Interpretationsprozess zu durchlaufen und innerhalb und zwischen den Dateien zu vergleichen, ob Qualitätsunterschiede der damit markierten Datensegmente erkennbar werden.

---

#### 5.4 Wie findet man Codes für die Bedeutungseinheiten?

Wenn man nach Bedeutungseinheiten in den Dateien sucht, sollte man offen sein für alles, was an Bedeutungen und Sinnzusammenhängen auftaucht. Viele wünschen sich ein Rezept, das den Kategorisierungsprozess erleichtern und helfen würde, die Phase größter Unsicherheit während der ersten Schritte der Analyse abzukürzen. Andererseits könnte sich der Preis, der für Sicherheit und Zeitersparnis zu zahlen ist, als zu hoch erweisen: Konkrete Leitfäden führen nur zu jenen Bedeutungseinheiten, die vorhersehbar und in der Anleitung erfasst sind, während überraschende, außergewöhnliche aber vielleicht äußerst interessante Aspekte in Gefahr sind, von der weiteren Analyse übergangen zu werden. Die folgenden Hinweise sollten als Heuristiken verstanden werden, d.h. als allgemeine Hinweise, wie man Bedeutungseinheiten identifiziert und passende Codes findet, jedoch nicht als Regeln, die Schritt für Schritt zum Ziel führen. Es bieten sich drei Strategien an. Insbesondere bei Nutzung der ersten beiden Möglichkeiten und ihrer im folgenden dargestellten Varianten bietet AQUAD wertvolle Unterstützung.

1. Bei der Suche nach *Kategorien* geht man eine Datei daraufhin durch, ob sich darin enthaltene Ereignisse oder Aussagen über Ereignisse, Situationen, Personen sowie geäußerte Meinungen, Ideen usw. einem übergeordneten Konzept zuschreiben lassen.
2. Bei der Suche nach *Sequenzen* achtet man auf die Darstellung von Zusammenhängen in den Ereignissen oder Aussagen, sucht also wesentlich umfassendere Bedeutungseinheiten als bei kategorialer Suche.
3. Bei der Suche nach *Themen* schließlich ist die größte Abstraktion gefordert; u.U. wird ein gesamter Datensatz darauf reduziert, welche Thematik darin angesprochen ist.

### 5.4.1 Wie findet man kategoriale Codes?

Das Problem ist hier das gleiche wie bei der Suche nach Bedeutungseinheiten, denn bei dieser Reduktionsform decken sich Bedeutungseinheiten und Kategorien. Bei der Zuordnung von Codes ist eine zusätzliche Entscheidung nötig: Sollen die Kategorien bzw. korrespondierenden Codes *beschreiben*, *deuten* oder *erklären* (siehe Miles & Huberman, 1994, S. 56)? Bei dieser Entscheidung ist natürlich keine Computerunterstützung zu erwarten. Methodisch festlegen muss sich der Forscher oder die Forscherin schon in eigener Verantwortung. Der Computer hilft nur, alle Entscheidungen zu dokumentieren, zu ordnen, bei Bedarf ohne großen Aufwand zu revidieren. Für die Lösung des ersten Problems bieten sich drei Möglichkeiten an:

#### Verwendung vorgegebener Kategoriensysteme

Für den Fall, dass es in einer quantitativen Studie *nicht* auf gegenstandsbasierte Theorierekonstruktion ankommt, man also nicht aus den verfügbaren Daten subjektive Erklärungssysteme erschließen will, ist das Prinzip rasch umschrieben: Man kann dann auf bereits verfügbare Kategoriensysteme zurückgreifen und die eigenen Daten nach den in diesen Systemen enthaltenen Interpretationsschemata reduzieren. "Verfügbar" sind solche Kategoriensysteme beispielsweise aus eigenen früheren Untersuchungen; man findet sie auch in der empirischen Literatur oder in theoretischen Analysen des interessierenden Inhaltsbereichs.

Bedient man sich vorgegebener Kategoriensysteme, so muss man entscheiden, welche Abschnitte in den vorliegenden Daten welchen Definitionsbeispielen für die Kategorien des Systems entsprechen, und dann mit Hilfe des Computers Fundstelle und Codierung festhalten. Dies wird in Kapitel 6 genau erklärt. Da unweigerlich manche Äußerungen nicht eindeutig zugeordnet werden können und da gelegentlich im Interpretationsprozess Inkonsistenzen auftreten, bietet eine Suchfunktion für spezifisch codierte Dateiabschnitte ("*Suchen*"; "*Bestimmte Codes*") große Hilfe bei der Kontrolle der Datenreduktion. Nach Eingabe eines kritischen Codes liefert eine derartige Funktion rasch und mühelos alle zugeordneten Dateiabschnitte in allen bis dahin analysierten Datensätzen.

Allerdings findet man bei dieser Form der Zuverlässigkeitskontrolle von Codierungen zunächst nur die fälschlich einer Kategorie zugeordneten Datensegmente, nicht aber die irrtümlich *nicht* dieser, sondern einer anderen Kategorie zugeordneten Passagen. Um diesen Fehler zu entdecken, müsste man bei konventioneller Datenreduktion darauf warten, welche Inkonsistenzen sich bei der Überprüfung der anderen Kategorienzuordnungen herausstellen. Diese Fälle müssten dann neu codiert werden. Man könnte zur Kontrolle eventuell "fehlender" Datensegmente auch sofort zumindest die Menge der Datenausschnitte ähnlicher, für Fehlzugeordnungen "anfälliger" Kategorien durchgehen. Ist Computerunterstützung verfügbar, ist die Konsistenzkontrolle, wie gerade beschrieben, viel weniger aufwändig. Darüber hinaus kann man sich - nur bei der Analyse von Texten (!) - nun die Möglichkeiten der wechselseitigen Ergänzung der Definition manifester und latenter Bedeutungseinheiten zunutze machen. Drei unterschiedlich komplexe Strategien werden dabei verwendet:

- (1) Man definiert *Schlüsselwörter* als manifeste Indikatoren eines kritischen Bedeutungsgehalts und sucht im gesamten Text nach ihrem Auftreten. Man kann dann jeweils beurteilen, ob die Textpassage, in der ein Schlüsselwort gefunden wird, nicht doch der gerade interessierenden Kategorie zugerechnet werden müsste. Diese Strategie lässt sich auch mit den üblichen Textprogrammen realisieren.
- (2) Man fasst Schlüsselwörter in einem *Analyselexikon* oder *Wortkatalog* zusammen, d.h. man gibt eine Liste von Schlüsselwörtern zur Suche nach kritischen Textpassagen ein. Damit erhält man in einem Durchgang durch die Texte die gesuchte Information.
- (3) In AQUAD sind die ersten beiden Möglichkeiten vorgesehen, jedoch automatisch mit der dritten Variante, der Funktion *Key-words-in-context* (Popko 1980), verknüpft. Damit kann nach einem oder nach mehreren Schlüsselwörtern hintereinander automatisch in allen Texten einer Studie gesucht werden. Als Ergebnis erhält man alle Zeilen ausgedruckt, in denen das gesuchte Wort vor-

kommt. Die weiteren Entscheidungen über Umfang der Bedeutungseinheit und Codierung liegen dann wieder bei der Forscherin oder beim Forscher.

### Hypothesengestützte Kategorisierung

Bei der Bestimmung und inhaltlichen Spezifikation von Bedeutungseinheiten in den zu analysierenden Dateien lässt man sich von Hypothesen über den Inhaltsbereich leiten. Auf der Basis von Hypothesen versucht man, Kategorien und Codierungsregeln für die Daten festzulegen.

Das Vorgehen von Marcelo (siehe S. 17) bei der Kategorisierung von Interviews mit Anfängern/innen im Lehramt war beispielsweise hypothesengeleitet. Ein theoretisches Modell beruflicher Sozialisation diente als Gerüst für den Entwurf eines vorläufigen Kategoriensystems. Wachsende Vertrautheit mit den Interviewtexten und zunehmende Einsicht in die subjektive Weltansicht der Berufsanfänger führten dazu, dass einige dieser Kategorien als wenig geeignet verworfen, andere Kategorien dagegen erst entdeckt und neu in das System aufgenommen wurden. Mit fortschreitender Analyse wurden einige dieser Kategorien zu größeren Einheiten im Sinne subjektiver Theorien der Berufsanfänger verknüpft (vgl. Huber & Marcelo, 1992), die dann wiederum am gesamten Material überprüft werden mussten.

Am besten legt man schon beim Planen der Datenerhebung, etwa bei der Formulierung von Leitfragen für ein Interview sowie vor der Datenanalyse für jede der Hypothesen konkrete Kategorien fest (deduktiver Anteil). Mit diesen Kategorien verfährt man dann wie im oben beschriebenen. Andererseits entdeckt man mit Sicherheit während der Datenanalyse in den Dateien Segmente, die man nicht den vorab festgelegten Kategorien zuordnen kann. Man markiert dann diese Bedeutungseinheiten und entwickelt induktiv Kategorien, die sie mit den gegenstandsspezifischen Hypothesen verknüpfen. Natürlich sind bei diesem Vorgehen auch Modifikationen des ursprünglichen Orientierungsrahmens nicht auszuschließen. Die Freiheitsgrade und die Ansprüche an die eigenen Interpretationsleistungen nehmen bei dieser Strategie ebenso zu wie die Möglichkeiten, bei der Analyse den spezifischen Sichtweisen der Forschungssubjekte gerecht zu werden.

Schon bei dieser Strategie kommt man nicht ohne die "Methode des ständigen Vergleichens" aus, die für den Ansatz der gegenstandsbezogenen Theoriebildung (Glaser & Strauss, 1967, 1979; Strauss & Corbin, 1990) kennzeichnend ist. Wie oben beschrieben, besteht der zentrale Prozess beim "ständigen Vergleichen" darin, jeden induktiven Schluss von den Besonderheiten der Datenbasis auf allgemeinere Prinzipien, hier nun Kategorien der Datenreduktion, mit Hilfe deduktiver Rückschlüsse auf die Datenbasis zu prüfen.

Mit den Anforderungen an die Interpretationsleistungen des Forschers wächst bei diesem Vorgehen die Bedeutung von Computerunterstützung. Bei der Kontrolle von Zusammenhängen zwischen Kategorien stößt man auf die Grenzen einfacher Suchfunktionen. Von diesem Punkt an benötigt man für qualitative Analysen AQUAD, das auf der Basis logischer Programmierung deduktives Schließen ermöglicht (Tesch 1990; Shelly & Sibert 1992).

- (1) Die Möglichkeiten zur *Suche von kodierten Dateisegmenten* und - im Fall von Texten als Datenbasis - von *Schlüsselwörtern* (einschließlich Suchlexika und KWIC-Funktion) sind zur Konsistenzprüfung der Kodierungen innerhalb der Einzeldateien und über alle Dateien hinweg hier ebenso zu nutzen wie bei der Verwendung vorgegebener Kategoriensysteme (s.o.).
- (2) Aus den Hypothesen, die den Rahmen für die Entwicklung von Kategorien abstecken, ergeben sich Hinweise auf bestimmte *Relationen der Kategorien*. AQUAD stellt für die hypothesengestützte Datenreduktion Funktionen zur Prüfung solcher Relationen bereit. Damit lassen sich beispielsweise das *Über-/Unterordnungs-Verhältnis* einzelner Kategorien, *Sequenzen* bestimmter Kategorien oder *Cluster* bestimmter Kategorien ermitteln. Diese drei Relationen stellen die wohl am häufigsten zu kontrollierenden Zusammenhangsmuster von Kategorien dar. Gemeinsam ist dabei, dass man nicht nur eine Kategorie bzw. die durch sie repräsentierten Dateisegmente im Auge behalten muss, sondern zwei oder mehr Kategorien und den definierten Zusammenhang zwi-

schen ihnen - der auch Negationen einschließen kann! Auch die Nicht-Erfüllung der Vermutung, dass bestimmte Relationen zwischen Bedeutungseinheiten vorliegen, ist natürlich zu registrieren.

### Theoriekonstruierende Kategorisierung

Die anspruchvollste Form qualitativer Datenreduktion verzichtet auf alle vorab festgelegten Reduktionsvorschriften in Gestalt von Kategoriensystemen und auf Strukturierung des Analyseprozesses durch hypothetische Rahmenkonzepte. In Ergänzung dazu muss der Forscher/die Forscherin auf eigene subjektiven Erfahrungen, Vorannahmen oder Vorurteile in Bezug auf das Erleben und Verhalten der Forschungssubjekte achten und durch ständigen Vergleich der im Auswertungsprozess entstehenden Reduktionsprinzipien mit den Ereignissen oder Aussagen in den Dateien versuchen, voreilige Verfestigungen zu vermeiden. Sensibilität gegenüber Lesarten, die man selbst in die Daten legt, ist deshalb besonders notwendig, weil inhaltsanalytische Verfahren allgemein auf Daten angewendet werden, die oft in großer Distanz zu Zeitpunkt und Kontext ihrer Entstehung ausgewertet werden (Fischer, 1982).

Information über die Datenproduzenten (handelnde Personen, Sprecher oder Schreiber) und ihren Kontext muss aus zusätzlichen Daten oder aus der vorliegenden Datei selbst gewonnen werden. Dieser Zugang verspricht mit größter Wahrscheinlichkeit, die Forscher und Forscherinnen dem Ziel näherzuführen, die Welt der Forschungssubjekte mit deren eigenen Augen zu sehen und aus ihrer eigenen Perspektive zu verstehen.

Versucht man in diesem Prozess, die subjektiven Weltansichten nicht nur zu beschreiben, sondern mit Hilfe geeigneter Begriffe zu ordnen und systematische Zusammenhänge zu rekonstruieren, dann bemüht man sich um das, was Glaser und Strauss (1967; 1979) die *Entdeckung* einer "grounded theory", einer *gegenstandsbezogenen Theorie* genannt haben. Im Wort "Entdeckung" ist der wesentliche Unterschied zu Methoden enthalten, die gegebene Theorien bestätigen wollen. Strauss und Corbin (1990) stellen heraus, dass gegenstandsbezogene Theorien in der analytischen Auseinandersetzung mit dem Gegenstand entwickelt werden, weshalb man nicht mit der Theorie beginnt, um sie anhand gegenstandsspezifischer Daten zu beweisen, sondern mit den beobachteten, aufgezeichneten Phänomenen, die man besser verstehen und erklären möchte.

Bei diesem Ansatz sind die höchsten Interpretationsleistungen gefordert - weshalb er für Anfänger die größten Schwierigkeiten bereithält. Vorgegebene Kategoriensysteme enthalten implizit in der Auswahl von Beispielen für die Kategorien oder explizit in den Definitionen der Kategorien klare Interpretationsregeln; hypothesengeleitetes Kategorisieren kann sich wenigstens auf allgemeine Orientierung darüber stützen, wonach in den Daten zu suchen ist. Theoriekonstruierende Kategorisierung dagegen muss erst einmal entdecken, wovon die Daten handeln, wovon in den Texten die Rede ist. Anfänger neigen dazu, möglichst wenig Irrtumrisiko einzugehen und so wenig wie möglich zu interpretieren (s.o.). Im Extremfall führt diese Tendenz dazu, dass (die meist verwendeten) Textdaten auf relevante Wörter hin durchgelesen werden, vergleichbar der Verwendung von Schlüsselwörtern beim Kategorisieren mit fertigen Kategoriensystemen. Die Textpassagen, in denen das kritische Wort enthalten ist, werden dann als Fundstelle einer "Bedeutungseinheit" markiert und mit einem Code bezeichnet. Bei diesem Verfahren kann der Code aber kaum subjektive Bedeutung repräsentieren, sondern vor allem eben die Tatsache, dass im bezeichneten Textabschnitt ein bestimmtes Wort aufgetreten ist.

Diese Tendenz ist besonders ausgeprägt, wenn man noch nichts von den Möglichkeiten tentativer Interpretation und mühevoller Revision bei computerunterstützter Inhaltsanalyse weiß. Mit AQUAD wird man für "Spielen" mit den eigenen Einfällen nicht durch unzumutbare Revisionsarbeiten bestraft, falls man später Widersprüche findet, im Gegenteil, man wird zur Interpretation angehalten und ermutigt, da man von Anfang an Codes zur Bezeichnung von Textsegmenten benötigt und da Änderungen, Zusammenfassungen, Differenzierungen ohne große Mühe vorgenommen werden können (vgl. Tesch 1992).

Als *Faustregel* für theorieentdeckendes Kategorisieren kann man sich daher zu eigen machen: Suche in den Daten nach Bedeutungseinheiten, die einerseits *so groß wie möglich* sind, damit es überhaupt etwas zu interpretieren gibt, die andererseits *so klein wie nötig* sind, damit nicht inkon-

sistente Inhalte vom gleichen Code repräsentiert werden. Damit ist die oben empfohlene *Differenzierungsstrategie* umschrieben. AQUAD unterstützt diese Strategie, da es keine Einschränkungen bei der Definition von Bedeutungseinheiten macht; insbesondere dürfen sich mehrere Einheiten überschneiden und man darf umgekehrt die selbe Einheit mehrfach codieren.

Wenn man mit dem Gegenstandsbereich, der in den Daten angesprochen wird, und mit der Methode vertraut ist, kann man diese Regel auch auf den Kopf stellen, d.h. von kleinsten Bedeutungseinheiten ausgehen und eine *Generalisierungsstrategie* verfolgen. Mit diesen Voraussetzungen bleibt man erstens nicht so leicht an Details kleben und übersieht dabei nicht das Wesentliche der Daten. Zweitens weiß man dann, dass Detailcodierungen mit Hilfe von *Funktionen der Meta-Codierung* (s. Kapitel 7) computerunterstützt automatisch zu größeren Einheiten verknüpft werden können. Durch ständiges Vergleichen der codierten Dateisegmente bestimmt man dann induktiv die Merkmalsdimensionen, die sie umschreiben, und schließlich die übergreifenden Kategorien, denen sie zuzuordnen sind. AQUAD unterstützt auch diese Strategie, da man sehr rasch von einer spezifischen Codierung aus nach konträren, ähnlichen, abhängigen oder in sonstiger definierter Relation zu dieser Codierung stehenden Dateisegmenten suchen kann. Auf der Basis dieser Relation/en kann man anschließend versuchen, die Teilcodierungen einer übergeordneten Kategorie unterzuordnen.

#### 5.4.2 Wie findet man sequentielle Codes?

Bei der kategorialen Reduktion versuchen wir, Segmente der Dateien so voneinander abzuheben, dass ihnen klar zu unterscheidende, sich wechselseitig ausschließende Bedeutungen zugeschrieben werden können. Die eindeutige Zuordnung von Bedeutungen zu Kategorien impliziert allerdings nicht, dass auch die Dateisegmente voneinander unterschieden sein müssen. Abhängig von den Ausdrucksgewohnheiten der Datenproduzenten und Lesarten der Forscher überlappen sich unterschiedlichen Kategorien zugeordnete Dateisegmente häufig. Im Extremfall kann das gleiche Segment mehreren Kategorien zugeordnet werden.

Bei solchen Verknüpfungen in den Daten setzt die Strategie der sequentiellen Reduktion an. Sie sucht nach einem spezifischen Zusammenhang von Bedeutungen, markiert das Auftreten solcher Verknüpfungen und bezeichnet entsprechende Dateipassagen mit einem Code - wie bei kategorialer Reduktion. Der Code repräsentiert hier allerdings eine definierte Sequenz von Bedeutungen. Die Bedeutungseinheit in der Datei ist in diesem Fall die gesamte Passage, in der dieser Zusammenhang beschrieben wird.

Welche Strategien gibt es nun, über die Differenzierung von Kategorien und Unterkategorien hinaus, Bedeutungszusammenhänge in den Daten zu entdecken? Wir unterscheiden im folgenden zwischen einer Strategien der Suche nach *einfachen Sequenzen* sowie der Definition und Suche nach *komplexen Sequenzmustern*.

#### Suche nach einfachen Sequenzen

Neben der Suche nach *hierarchischen Sequenzen* von über- und untergeordneten Kategorien, die Strauss und Corbin (1990) zur sequentiellen Reduktion empfehlen, bieten sich mit Blick auf die grammatikalisch-linguistische Systematik eine ganze Reihe weiterer einfacher Sequenzen an. Welche davon für die Datenreduktion sinnvoll sind, hängt natürlich auch von der Forschungsfrage ab. In Texten wird häufig nach *kausalen Sequenzen* (auch mit Schlüsselwörtern wie "aufgrund", "weil", "denn" usw.) gesucht, nach *temporalen Sequenzen* ("während", "dann", "ehe" usw.), *konzessiven Sequenzen* (positive Einschränkungen wie "nicht ... jedoch" oder negative Einschränkungen vom Typ "zwar ... aber"), *konditionalen Sequenzen* ("wenn ... dann"), *finalen Sequenzen* ("damit", "zwecks", "um zu" usw.), *komparativen Sequenzen*, *modalen Sequenzen* und *Präziserungssequenzen*.

Weder die Aufzählung dieser Sequenztypen noch die dabei angeführten möglichen Schlüsselwörter erheben Anspruch auf Vollständigkeit. Sie sollen nur zu eigenen Versuchen anregen, der Forschungsfrage entsprechend die Daten nach Bedeutungssequenzen zu reduzieren. In diesem Zusammenhang ist die Funktion des Computers bzw. der Software als nützliches *Werkzeug* für die Interpretation,

nicht aber als Agens der qualitativen Analyse zu betonen. So hilfreich Funktionen zur Wortsuche sich bei diesem Vorgehen erweisen können, so eingeschränkt sind doch die Anwendungsmöglichkeiten. Wenn man nicht gerade sorgfältig formulierte Dokumente analysiert, ist der Ertrag von Suchlexika für spezifische Bedeutungssequenzen meist begrenzt. Im freien Interview lassen sich häufig schon die Satzstrukturen nur schwer festlegen. Wo ist der Hauptsatz, wo der Nebensatz? Dementsprechend werden die für die Definition der Sequenz kritischen Konjunktionen zwar vielleicht mitgedacht, aber nicht mitgesprochen. Darüber hinaus ist der Gebrauch der kritischen Konjunktionen und Redewendungen bei vielen Sprechern grammatikalisch nicht einwandfrei. Die Suche nach Schlüsselwörtern oder ganzen Lexika solcher Wörter ersetzt also nicht die einführende Interpretation, doch kann sie eine nützliche Heuristik bei dieser Arbeit darstellen.

### Suche nach komplexen Sequenzmustern

Bei der Interpretation von Daten kann man sich von strukturellen oder prozessualen Charakteristika des untersuchten Gegenstands anregen lassen und nach komplexeren Bedeutungssequenzen als oben skizziert suchen. Für die Rekonstruktion beispielsweise subjektiver Handlungstheorien der Datenproduzenten könnte man etwa nach Sequenzen von Situationseinschätzungen, Reflexion von Handlungsalternativen, Zuordnung von erwarteten Ergebnissen und der Einschätzung von persönlichen Folgen wie Zufriedenheit oder Enttäuschung suchen.

Wegen der Zahl unterschiedlicher Kategorien, die man gleichzeitig im Auge behalten muss, wenn man feststellen möchte, ob in einer Datei ein bestimmtes Sequenzmuster von Ereignissen bzw. Aussagen vorkommt oder nicht, ist diese Analysestrategie sehr anstrengend und irrtumsgefährdet. Allerdings stellen komplexe Sequenzanalysen nicht nur an die Dateninterpreten, sondern auch an die Computerprogramme erhebliche Ansprüche. Die Vielfalt an Strukturen denkbarer Kategorieverknüpfungen schließen es aus, fertige Deduktionsalgorithmen für alle möglichen komplexen Sequenzen in einem Programm zu implementieren, in das die Datenanalytiker dann nur die Codes, die er benutzt hat, als Variablen einsetzen müsste. Die Benutzer eines Programms zur qualitativen Analyse benötigen zur Lösung der hier beschriebenen Probleme daher Möglichkeiten, ohne Programmieraufwand die Regeln in die Analyseprozedur einzubauen, nach denen verknüpft Kategorie sequenzen in Daten erwartet werden. AQUAD bietet diese Möglichkeit (s. Kapitel 12).

#### 5.4.3 Wie findet man thematische Codes?

Den radikalsten Ansatz verfolgt man, wenn man versucht, die Codierung einer Datei auf eine zentrale Kategorie, d.h. auf ihre Botschaft oder ihr zentrales Thema zu reduzieren. Da quantitative Analysen als zyklische Prozesse ablaufen, kann man der thematischen Reduktion keinen festen Ort, etwa am Ende der Analyse zur Zusammenfassung der Befunde, zuweisen. Zwar erfüllt die Verdichtung der Daten in einer zentralen Kategorie an dieser Stelle eine wichtige Funktion, doch ist der Ansatz auch in anderen Placierungen sehr nützlich:

- (1) Bei relativ kurzen und homogenen Dateien kann die thematische Reduktion am Anfang des Prozesses der Datenanalyse stehen.
- (2) Bei langen, heterogenen, unübersichtlichen Dateien kann man die Datenreduktion mit dem thematischen Ansatz beginnen. Bei dieser Verwendung erfüllt die thematische Reduktion wichtige heuristische Funktionen. Man versucht zunächst, sich durch die Vielzahl vielleicht auch widersprüchlicher Details nicht verwirren zu lassen und den oder die Hauptgedanken herauszufiltern.
- (3) In der Regel steht die thematische Reduktion am Ende der Sequenz analytischer Schritte, oft erst am Ende mehrerer Zyklen von Datenreduktion, Rekonstruktion und Vergleich von Bedeutungssystemen, in denen die "Leitmotive" verschiedener Dateien allmählich herausgearbeitet wurden (vgl. Shelly & Sibert 1992; Strauss & Corbin 1990).

---

Strauss und Corbin (1990) empfehlen für die thematische Reduktion fünf Schritte. Diese Abfolge lässt sich mit jeder der drei beschriebenen Placierungen des Ansatzes im analytischen Zyklus vereinbaren und sie hilft, den Beitrag von Softwarefunktionen besser einzuschätzen:

- (1) Man beginnt damit, einen Leitgedanken, eine zentrale Kategorie zu identifizieren;
- (2) dann sucht man nach untergeordneten Kategorien;
- (3) oft wird es nötig sein, diese Kategorien zu differenzieren bzw. zu verknüpfen;
- (4) die hypothetischen Relationen zwischen Unterkategorien sowie diesen und dem Thema werden anhand der Daten überprüft;
- (5) Unstimmigkeiten, z.B. inkonsistente Kategorien, geben Anlass zu weiteren Analysezyklen.

## 5.5 Wie rekonstruiert man Bedeutungssysteme?

Bei der theoriekonstruierenden qualitativen Analyse geht es darum, aus den Ereignissen oder Aussagen in den Daten induktiv die Theorie über den Gegenstand zu generieren, die den Urhebern der Daten subjektiv zur Verfügung steht. Wir sprechen hier von der *Rekonstruktion von Bedeutungssystemen*. Das Problem, wie man die subjektiven Verknüpfungen findet, kann man induktiv, deduktiv oder durch Kombination induktiver und deduktiver Strategien zu lösen versuchen. Welchen Ansatz man bevorzugt, hängt vor allem von der Forschungsfrage ab.

Wenn man beispielsweise Aufzeichnungen oder Transkriptionen wenig strukturierter Interviews oder frei gestalteter Texte, z.B. Tagebuchaufzeichnungen analysiert, beginnt man im allgemeinen damit, einzelne Bedeutungseinheiten zu identifizieren und die entsprechenden Dateisegmente spezifischen Kategorien zuzuordnen. Dann wird man typische Sequenzen dieser Kategorien suchen. Schließlich versucht man, solche Sequenzen abstrakteren Kategorien, dem Thema oder den Themen des Sprechers unterzuordnen. Dabei fängt man *induktiv* an, wechselt aber spätestens bei der thematischen Reduktion zwischen induktiven und deduktiven Schlüssen ab.

Geht man mit einem spezifischen Erkenntnisinteresse oder einer bestimmten theoretischen Orientierung an die Daten heran, wird man von vorne herein nach bestimmten Zusammenhängen suchen. Man geht also von Hypothesen über mögliche Zusammenhänge aus und versucht, diese *deduktiv* anhand der Daten zu belegen. Unstimmigkeiten und Misserfolge geben Anlass zu induktiven Analyseabschnitten, mit denen wieder das hypothetische Orientierungssystem modifiziert werden kann.

Beim induktiven Vorgehen bemüht man sich, aus den Daten heraus Kategorien und ihre systematischen Zusammenhänge zu generalisieren. Beim deduktiven Vorgehen sucht man nach spezifischen Dateisegmenten, die allgemeine Vorannahmen oder Hypothesen über den Zusammenhang der Kategorien bestätigen können. Nach diesem Schema ist die folgende Übersicht der in AQUAD verfügbaren Techniken computerunterstützter Rekonstruktion systematischer Zusammenhänge gegliedert.

### 5.5.1 Rekonstruktion einfacher Bedingungen

Abhängig vom Fortschritt des Analyseprozesses lassen sich dabei wieder zwei Varianten unterscheiden:

(1) Suche nach *einfachen Bedeutungsabfolgen*. Man legt für die Suche einen Datenbereich vor und nach den Datensegmenten einer ausgewählten Kategorie fest, d.h. man bestimmt, wie viele Zeilen oder Zählwerkseinheiten (Frames bei Videos, Millisekunden bei Audios) vor und nach den Datensegmenten der ausgewählten Codierung abgesucht werden sollen. Innerhalb dieses Datenbereichs soll dann das Auftreten anderer Codierungen registriert werden. Treten bestimmte Codierungen häufig in der Nachbarschaft des kritischen Codes auf, kann eine mögliche Systematik hinter diesen Kombinationen überprüft werden. In AQUAD unterstützt die Option Suchen nach "*Codierungssequenzen*" im Modul "*Suchen*" diesen Rekonstruktionsansatz (siehe Kapitel 9).

(2) Suche nach *bedingten Bedeutungszusammenhängen*. Die "Bedingung" für das mögliche Vorliegen systematischer Zusammenhänge wird durch (mindestens) eine zweite Kategorie (in mindestens zwei Bedeutungsvarianten, z.B. "Geschlecht") definiert, der ein Datensegment gleichzeitig zugeordnet sein muss. Zur computerunterstützten Suche konstruiert man Codierungsmatrizen (Miles & Huberman, 1994) oder Codierungstabellen (Shelly & Sibert, 1992). Die zweifache Determination der Inhalte (Datensegmente) der Zellen einer solchen Tabelle legt die Interpretation der Befunde stärker fest als die bloße Feststellung einer Häufung räumlich-zeitlicher Nähe zunächst unabhängiger Kategorien. Andererseits sind zur Konstruktion einer Analysematrix mehr konzeptuelle Vorleistungen, mithin größere Fortschritte im Analyseprozess nötig. Das Modul "*Tabellen*" (siehe Kapitel 11, Abschnitt 2) steht in AQUAD für diesen Rekonstruktionsansatz zur Verfügung.

---

### 5.5.2 Rekonstruktion anhand der Prüfung bestimmter Relationen

Auch hier sind zwei Varianten zu unterscheiden, die den oben beschriebenen einfachen und komplexen Strategien des sequentiellen Codierens entsprechen:

(1) *Prüfung einfacher Codierungssequenzen.* Nehmen wir an, aus einem Elterninterview über familiäre Erziehungspraktiken ergibt sich der Eindruck, dass ein Vater sich besonders bemüht, seine Maßnahmen zu rechtfertigen, dann könnte man mit dem Modul "Verknüpfungen" (vgl. S. 10 und Kapitel 11, Abschnitt 3) gezielt nach Sequenzen relevanter Kategorien, z.B. nach finalen oder kausalen Codierungen suchen und so die Vermutung überprüfen.

(2) *Prüfung komplexer Codierungsmuster.* Für die Überprüfung von Codierungsrelationen, die über den Komplexitätsgrad der im Modul "Verknüpfungen" vorgegebenen Verknüpfungsstrukturen hinausgehen, steht in AQUAD eine "offene" Programmkomponente zur Formulierung komplexer, damit fallspezifischer Verknüpfungen von Codierungen zur Verfügung. Wie man seine Vermutungen über Zusammenhänge in eine von AQUAD prüfbare Form bringt, wird im Kapitel 12 anhand von Beispielen genau beschrieben.

## 5.6 Wie vergleicht man Bedeutungszusammenhänge?

Ständige Vergleiche der Interpretationen innerhalb eines Datensatzes (Text, Video, Audioaufzeichnung, Bild) und zwischen verschiedenen Datensätzen bilden den Kern qualitativer Analyseverfahren (vgl. Shelly & Sibert 1992). Schon bei der Datenreduktion *innerhalb einer Datei* bringt man ohne Vergleiche der Kategorien bzw. der Datensegmente sowohl untereinander als auch zwischen den übrigen Dateien keine reliablen Codierungen zustande. Nun möchte man aber in den meisten Untersuchungen mehr als mit einem für alle Dateien gültigen Kategoriensystem die Einzigartigkeit jedes Datensatzes bzw. der darin ausgedrückten Meinungen, Kompetenzen, subjektiven Weltansichten usw. erfassen. Irgendwann im Forschungsprozess will man generelle Konfigurationen *über die Dateien hinweg* herausarbeiten. Man muss also bei qualitativen Analysen mit der sprichwörtlichen Gefahr umgehen, dass man oft vor lauter Bäumen den Wald nicht sieht. Über den Unterschieden und ihrer Präzisierung darf man nicht das Gemeinsame aus den Augen verlieren. Dazu muss man über die Dateien hinweg systematische Zusammenhänge erkennen und vergleichen.

Das Problem des Vergleichens von Zusammenhängen besteht bei sozialen Phänomenen jedoch darin, dass eine Vielzahl von Bedingungen oft in verschiedenartigen, manchmal sogar widersprüchlichen Kombinationen gefunden werden. Bei vielen Untersuchungen steht man schließlich vor der zusätzlichen Aufgabe, die eigenen Befunde mit anderen Untersuchungen vergleichen zu müssen, d.h. im Grunde eine *qualitative Meta-Analyse* durchführen zu müssen. Auch wenn andere Forscher vergleichbare Fragen bearbeitet haben, werden sie dabei aber nur teilweise übereinstimmende Antworten gefunden haben. Bei der Suche nach Bedingungsbeziehungen für ein bestimmtes Phänomen werden in allen Wissenschaften, die sich mit der Komplexität natürlicher Systeme auseinandersetzen, über verschiedene Untersuchungen hinweg die unterschiedlichsten Konstellationen nachgewiesen.

Nehmen wir eine im Alltag von den Kontrahenten oft erbittert diskutierte Frage als Beispiel: "Besteht ein Zusammenhang zwischen Lungenkrebs und dem Rauchen?" Wer dies bejahen will, sieht sich regelmäßig mit dem Hinweis auf den 80-jährigen Großvater konfrontiert, der sich seit früher Jugend am blauen Dunst erfreut habe oder mit dem Hinweis auf ein Opfer des Lungenkrebses, das nie geraucht habe. Offenbar spielen viele Bedingungen eine Rolle. Empirische Untersuchungen liefern vielfältige empirische Argumente, aus denen man sich je nach persönlichem Interesse oder eigener Betroffenheit die passenden zur argumentativen Immunisierung herausuchen kann. Erst ein Vergleich der Befunde aller einschlägigen Untersuchungen oder wenigstens einer repräsentativen Stichprobe daraus kann Klarheit über relevante Bedingungsbeziehungen schaffen.

Ragin (1987) hat mit der Anwendung Boolescher Algebra auf qualitative Daten ein explizites komparatives Verfahren erschlossen. Das Verfahren beruht auf dem Quine-McClusky-Algorithmus der "logischen Minimierung". Dieses Verfahren erfüllt die Anforderungen an die gesuchte Vergleichsmethode in hohem Maße, zumindest schafft die Anwendung Boolescher Algebra ideale Voraussetzungen. Nach Ragin (1987, S. 121) kann man damit die Forderungen erfüllen,

- eine Vielzahl von Texten oder Einzelfällen zu vergleichen;
- komplexe Verknüpfungen von Bedingungen zu erfassen;
- falls erwünscht, "sparsame" Rekonstruktionen herzustellen;
- einzelne Texte sowohl in relevanten Teilen als auch als Ganzheit (vgl. die thematische Reduktion) zu untersuchen;
- konkurrierende Rekonstruktionen zu vergleichen.

Vergleichen wir dieses Verfahren mit variablen-orientierten Ansätzen, die von der Grundannahme additiver Aggregation einzelner Variablen ausgehen, können wir herausstellen, dass der fallorientierte Vergleich darauf angelegt ist (vgl. Ragin 1987, S. 51 f.),

- Invarianzen oder konstante Bedeutungszusammenhänge durch sorgfältige Vergleiche der Einzelfälle herauszuarbeiten;

- die Variabilität bedeutungsvoller Bedingungskonfigurationen stärker zu beachten als bloße Häufigkeitsverteilungen von typischen Fällen, woraus folgt, dass auch ein einzelner widersprüchlicher Fall Beachtung finden muss;
- Fälle als Ganzheiten zu erfassen, d.h. die Bedingungen des Falls in wechselseitiger Abhängigkeit zu sehen, die diesen besonderen Fall konstituiert, nicht aber die Einzelbedingungen in Abhängigkeit von einer Populationsverteilung;
- genau zu untersuchen, wie die insgesamt erschlossenen Bedingungen in unterschiedlichen Kontexten und in unterschiedlichen Konfigurationen zu unterschiedlichen Befunden führen.

Um die Regeln Boolescher Algebra auf Dateninterpretationen anwenden zu können, reduziert man in jeder Datei die gefundenen Bedeutungen radikal auf *Wahrheitswerte*. Man begnügt sich also mit der binären Notierung "Bedingung ist wahr" (d.h. gegeben) bzw. "Bedingung ist falsch" (d.h. *nicht* gegeben). Dabei spielt es keine Rolle, ob man es mit genuin qualitativen Bedingungen, also Interpretationen und der Zuordnung passender Kategorien, oder mit Messwerten eines quantitativen Merkmals zu tun hat.

AQUAD bietet die Möglichkeiten der logischen Minimierung im Modul "*Implikanten*". Wenn im Prozess der qualitativen Analyse Vergleichsoperationen nötig werden - und das kann aufgrund der Analysezyklen in jeder Phase sein - sollte man auch an die Möglichkeiten der logischen Minimierung denken.

Als *Heuristik* leistet sie schon bei der Generierung von Kategorien zur Interpretation gute Dienste, auch wenn erst wenige Dateien analysiert wurden, die Vergleichsbasis also noch schmal ist. Gegen Ende der Analyse, bei der *Zusammenfassung der Ergebnisse* oder wenn man Einzelbefunde gruppieren will, Typen von Dateien oder Sprechern unterscheiden möchte, Schlüsseltexte herausstellen möchte usw., erscheint das Verfahren der logischen Minimierung unverzichtbar. Denn, wie Ragin (1987, S. 51) feststellt, wächst mit jeder zusätzlichen Datei der Arbeitsaufwand geometrisch, mit jeder zusätzlich im Vergleich zu berücksichtigenden Kategorie oder Bedingung aber exponentiell an! Die Gruppierung oder Clusterung von Einzelfällen mit Hilfe der logischen Minimierung wurde in AQUAD besonders ausgebaut.

Schließlich kann man das Verfahren der logischen Minimierung anwenden, wenn man den *Vergleich mehrerer qualitativer Untersuchungen* beabsichtigt. Dies ist der methodische Schritt, auf den der Terminus "Meta-Analyse" üblicherweise angewendet wird. Wenn die Untersuchungsergebnisse quantitativ ausgedrückt worden sind, kann man auf eine Reihe hoch formalisierter Verfahren zur Meta-Analyse zurückgreifen. Im Bereich der qualitativen Forschung kann der Ansatz der logischen Minimierung die wünschenswerte Vereinfachung, Transparenz, Zuverlässigkeit und Dokumentierbarkeit meta-analytischer Vergleiche bringen. Mehr dazu erfahren Sie in Kapitel 13.

